BRIDGING THE SEMANTIC GAP: EXPLORING DESCRIPTIVE VOCABULARY FOR IMAGE STRUCTURE

Caroline Beebe

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
in the School of Library and Information Science,
Indiana University
August 2006

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Elin K. Jacob, Ph.D.

Alice Robbin, Ph.D.

Katy Borner, Ph.D.

K. Anne Pyburn, Ph.D.

Date of Oral Examination
August 4, 2006

Acknowledgements

A dissertation may be authored by one individual, but the process requires an immense amount of help and support. Many, many thanks are owed to the following people:

Dr. Elin K. Jacob, my advisor and committee chair, this task could not have been accomplished without your guidance, your insight, and your painstaking work training me to be a better writer. Dr. Jacob deserves credit for coining phrases descriptive of this work: Internal contextuality, semantic coherence, semantic integrity, semantic organizers, and the shared-ness measure.

Dr. Alice Robbin and Dr. Katy Borner, my SLIS committee members, you were each extremely enthusiastic, patient and supportive of my work.

Dr.Anne Pyburn, my Anthropology committee member, your faith in my abilities was more than I often had.

Dr. Howard Levie, my mentor in Instructional Systems Technology, who encouraged me in early theory building, and who passed on way too soon.

Dick Yoakam, School of Journalism, mentored me through the qualifying paper and is sorely missed.

Violet Wyld, the family cat, was my companion through the hardest and loneliest part of proposal creation, with her attempts at keyboard support, and her constant meows of agreement.

Genevieve Pritchard, my daughter, cheerleader, and support for data entry, you were ever ready with a can-do attitude and late night calculating.

Crystal Yoakam, my daughter, you were ever ready with the best consoling hugs; you stepped in to find Christian Kelly to format this document, and with your help this will get published.

Denise Schmutte, my statistics coach, who took the time to really understand my data and created the concept of a shared-ness rating.

Scott Schroeder, MBA student and Excel aficionado, took mere hours to create formulas that would have taken me days.

Cathy Spiaggia, my friend, and Tamara Morrison, my cousin, both of whom constantly encouraged me, over all these long years ... it’s finally done!

Bud Beebe, my Dad, and Mary BobYoakam, my mother-in-law, both would have been so proud.

Jane Beebe, my Mom, whose love and support has meant so much through the grueling finish.

Michael Yoakam, my husband, who stood by me through this seemingly endless process and to his credit did not let this exercise destroy our marriage; you have seen the worst of me and the best of me, you are the love of my life.

Preface

A picture is worth a thousand words: Which ones?

Describing pictures with words is like building trees with lumber.

“The study of form may be descriptive merely, or it may become analytical. We begin by describing the shape of an object in the simple words of common speech: we end by describing it in the precise language of mathematics; and the one method tends to follow the other in strict scientific order and historical continuity ... The mathematical definition of form has a quality of precision quite lacking in our earlier stage of mere description ... We are brought by means of it into touch with Galileo’s aphorism (as old as Plato, as old as Pythagoras, as old perhaps as the wisdom of the Egyptians) that the book of nature is written in characters of geometry.” (Sir d’Arcy Wentworth Thompson, “On Growth and Form,” 1917, p. 719)

Caroline Beebe

BRIDGING THE SEMANTIC GAP: EXPLORING DESCRIPTIVE VOCABULARY FOR IMAGE STRUCTURE

Content-Based Image Retrieval (CBIR) is a technology made possible by the binary nature of the computer. Although CBIR is used for the representation and retrieval of digital images, these systems make no attempt either to establish a basis for similarity judgments generated by query-by-pictorial-example searches or to address the connection between image content and its internal spatial composition. The disconnect between physical data (the binary code of the computer) and its conceptual interpretation (the intellectual code of the searcher) is known as the semantic gap. A descriptive vocabulary capable of representing the internal visual structure of images has the potential to bridge this gap by connecting physical data with its conceptual interpretation. The research project addressed three questions: Is there a shared vocabulary of terms used by subjects to represent the internal contextuality (i.e., composition) of images? Can the natural language terms be organized into concepts? And, if there is a vocabulary of concepts, is it shared across subject pairs? A natural language vocabulary was identified on the basis of term occurrence in oral descriptions provided by 21 pairs of subjects participating in a referential communication task. In this experiment, each subject pair generated oral descriptions for 14 of 182 images drawn from the domains of abstract art, satellite imagery and photo-microscopy. Analysis of the natural language vocabulary identified a set of 1,319 unique terms which were collapsed into 545 concepts. These terms and concepts were organized into a faceted vocabulary. This faceted vocabulary can contribute to the development of more effective image retrieval metrics and interfaces to minimize the terminological confusion and conceptual overlap that currently exists in most CBIR systems. For both the user and the system, the concepts in the faceted vocabulary can be used to represent shapes and relationships between shapes (i.e., internal contextuality) that constitute the internal spatial composition of an image. Representation of internal contextuality would contribute to more effective image search and retrieval by facilitating the construction of more precise feature queries by the user as well as the selection of criteria for similarity judgments in CBIR applications.

Table of Contents

Tables List

Figures List

Appendix List

CHAPTER 1
Introduction to the research

Content-Based Image Retrieval (CBIR) is an emerging technology made possible by the application of the binary nature of the computer that is being applied to the representation and retrieval of digital image resources. In its most fundamental application, CBIR image analysis algorithms create indexes of images by comparing ranges of colors, arrangements of colors, and relationships among pixels. There is a disconnect, however, between the binary code (physical data) of the computer and the intellectual code (conceptual interpretation) of the searcher – a "lack of coincidence between the information that one can extract from visual data and the interpretation that same data have for a user in a given situation" (Smeulders, Worring, Santini, Gupta, & Jain, 2000, p. 5). This disconnect is referred to as the "semantic gap" (Smeulders *et al*., 2000, p. 5; Stenvert, 1992). Identification of a descriptive vocabulary that is capable of representing internal visual structure of an image has the potential to bridge this gap by building a connection between the physical data and its semantic interpretation.

This research was designed to determine if there is a shared vocabulary for concepts and terms that can be used to develop a controlled vocabulary for representing complex elements in digital images. Two broad questions were investigated concerning the representation of images in the digital environment: Is there a vocabulary of natural language terms or shared concepts that can be used to describe the perceptual characteristics of images; and, if such a vocabulary can be identified, what terms and/or concepts constitute the vocabulary? Toward this end, the following research questions were posed:

1. Is there a shared vocabulary of terms that is used by subjects to represent the internal contextuality (e.g., the internal structural composition) of images?
2. Can the natural language terms used by subjects be organized into concepts?
3. If there is a vocabulary of concepts, is it shared across subject pairs?

1.1 Introduction to the problem

Soergel (1985) suggests that any electronic information system is dependent on the existence of data banks, whether collections of resource representations or full text documents, against which queries can be registered. Advances in electronic hardware, computing speed, display quality, storage capacity and network availability, furthered by continuing reductions in cost, are increasing the potential for development of image collections and delivery of non-text-based information in the form of static and dynamic images as well as sound. The World Wide Web provides an excellent example of the potential utility of systems that include both text and non-text-based information. Image data banks, however, are in their infancy. While museums and stock photo houses are providing Web access to their collections, the availability of electronic image collections in public or research libraries is limited, due both to the cost of development and to a lack of agreement on standards for both storing and describing images. Because electronic access to non-text-based collections is limited in any environment, it constitutes a growing area of interest for research and development. Though CBIR technology has been developing over the last fifteen years, Datta, Li, and Wang (2005) have identified tremendous growth in CBIR related publications, citations, and research directions in the last five years as interest in image access grows and the need for automated access tools becomes more apparent.

Searching for a resource is generally dependent on the existence of a verbal representation of the resource within the retrieval system. When the resource is an image, verbal description, or *ekphrasis* (Bolter, 1996; Elkins, 1999), is the usual method of representation. However, with the proliferation of graphic resources available on the Web, the dominance of text-based materials is challenged by an increasing emphasis on image-based communication. Eggleston (2000) cites industry estimates that "over 2700 photographs are taken every second, adding up to 80 billion new images each year" (p. 31). This estimate does not include the two billion x-rays, mammograms and CT scans that are generated annually. Assigning descriptors to images to support text-based discovery is time-consuming and often requires subject matter expertise. Documentation of the collection-building process (Besser & Snow, 1990; Besser & Yamashita, 1998; Eakins & Graham, 1999) indicates that it takes ten to forty minutes per image for an art history expert to add access terminology to a single image. Application of CBIR technology has the potential to cut the time spent indexing images by generating automatic representations based on pixel configurations.

Effective articulation of search queries depends on resource representation standards and on controlled vocabularies such as *The Art and Architecture Thesaurus* (AAT) (Petersen, 1994). According to Wellisch (1995), controlled vocabularies are natural language subsets "whose vocabulary, syntax (the way words are put together to form sentences or phrases), semantics (the meaning of words), and pragmatics (the way words are being used) are limited" (p. 214). Adherence to accepted standards such as *Anglo-American Cataloguing Rules, second edition revised* (Gorman & Winkler, 1998) and normalization of the language that people use in everyday communication provides for consistency and predictability in the indexing of resources. E. K. Jacob (personal communication, July 1998) presents a representation model consisting of physical description; conceptual description; and contextual description. *Physical description* includes both administrative data (e.g., access rights, object location, copyright holder, file type) and biographical characteristics (e.g., creator, title). The focus of *conceptual description* is both the naming of objects or *ofness,* (e.g., girl, artist, studio), and the interpretation of image reference or *aboutness,* (e.g., is it civil war or guerrilla warfare?). *Contextual description* provides a context for the referent of the image (e.g., the relationship of an object to other objects in an image).

CBIR indexes expand the notion of physical description to include the pre-semantic physicality of the pixel relationships, or the physical visual structure. However, this dimension of physical characteristics may be tied more closely to the conceptual description of an image than to either its biographical or administrative characteristics. The interrelationship between the physical characteristics of the image and its conceptual and contextual descriptions affects ekphrasis – the conceptual interpretation and verbal representation of the image.

A ten-level pyramid model introduced by Jaimes and Chang (2000; Jorgensen, Jaimes, Benitez, & Chang, 2001) categorizes methods of representation used with image resources (see Figure 1.1). The authors identify the first four categories of the pyramid with *Syntax*, or data-driven representations, and the remaining six categories with *Semantics*, or conceptual modeling. The semantic levels refer to the generic, specific, and abstract conceptual representations of the objects depicted in an image, or image ofness and aboutness. Although the model does not specifically address contextual description, the first syntactic level, *Type/Technique*, refers to image production, corresponding roughly to aspects of the administrative and biographical categories of physical description in Jacob's representation model. The other syntactic levels of *Global Distribution*, *Local Structure*, and *Global Composition*, however, build on analysis of the internal structural characteristics of an image using "low-level perceptual features such as spectral sensitivity (color), and frequency sensitivity (texture)" (Jorgensen et al., 2001, p. 940) as specified by CBIR methodologies. These low-level characteristics and relationships extend Jacob's conceptual and contextual descriptions to form the basis for representation of an image's *internal contextuality.*

Pyramid Model	Jacob Model
Syntax	
Type/Technique	**Physical description**
Global Distribution	
Local Structure	(not addressed)
Global Composition	
Semantics (Naming)	**Conceptual description**
Generic Objects	
Generic Scene	
Specific Objects	
Specific Scenes	
Abstract Objects	
Abstract Scene	
(not addressed)	**Contextual description**

Figure 1.1. Image Representation models. Pyramid model of image representation (Jaimes & Chang, 2000), left, as related to Jacob's model of resource representation, right.

Although Jorgensen *et al*. (2001) focus on testing the viability of their model for indexing, they argue that "such a structure can also facilitate searching by disambiguating among terms that could appear at several (syntactic) levels of the structure" (p. 941). Articulations of syntactic-level terminology exist in the various CBIR application interfaces and in much of the mathematical research that is generating methodologies, including terms such as *histogram*, *trigram*, and *local binary pattern*[1] that reference structural analysis methodologies. Extant controlled vocabularies, such as AAT, contain some non-Type/Technique syntactic-level terminology but it is minimal (Jorgensen, 1996) and organized for conceptual description rather than perceptual features.

Typically, CBIR applications use some form of query-by-pictorial-example (QBPE) as a search interface. QBPE methods rely on searcher identification of one or more images that are similar to the query. The CBIR system then analyzes the pixel

[1] For usage of these CBIR methods, also referred to by terms such as Bingrad/Trigram, Intensity/LBP, and Bingrad projections, see examples in the Leiden 19th Century Portrait Database (available at http://nies.liacs.nl:1860/).

configuration of the example image(s) and uses the result to identify images with similar pixel configurations. In general, however, the system is not informed of similarity criteria used by the searcher and the searcher is not supplied with an explanation of the similarity measures used by the system to generate the result set (Figure 1.2).

Figure 1.2. QBPE search results. Often result sets of images may have visually similar characteristics but no semantic similarity. These images[2] were retrieved from the Hermitage Museum using a QBIC color layout search query of 40% blue on top and 60% green on bottom.

An emerging direction in CBIR research builds on the assumption that it is more useful for CBIR interfaces to help users correlate semantics with perceptual clues than to have the computer automatically identify perceptual similarities and attempt to match them to semantic queries (Goodrum, Rorvig, Jeong, & Suresh, 2001; Santini & Jain, 1997, 1998; Smeulders *et al*., 2000). Research comparing perceptual judgments of image similarity produced by searchers and similarity ratings generated by CBIR methodologies indicates that there is a correlation between the metrics of image features and semantic

[2] Thumbnails are, left to right: *Ruins near the Mouth of a River*, Claude Vernet, 1748; *Monastery*, Unknown, 1st half of 19th Century; *The Great Theatre in St. Petersburg, Designs of the Façade and Sections of the Circles and Imperial Box,* Giacomo Quarenghi, Early 19th Centruy; *Cartouche*, Ottomar Elliger III, Between 1727 and 1735; *Portrait of Ivan Ivanovich Shuvalov,* Unknown, 1780s; *Christian Martyr on the Cross (St. Julia),* Gabriel Cornelius von Max, 1865. Retrieved October 3, 2002, from http://www.hermitagemuseum.org

description (Rogowitz, Frese, Smith, Bouman, & Kalin, 1998) and has become a growing research trend in CBIR (Datta *et al*., 2005). In similar fashion to Jorgensen's opposition of syntax and semantics, Chu's (2001) analysis of image indexing and retrieval research indicates that there are two approaches to representation, which he describes as *content-based* (or low-level) and *description-based* (or high-level). Low-level metric features of the content-based approach have yet to be related to high-level semantic descriptions of the searcher, creating a "sensory gap" between computational description and the real-world object (Smeulders et al., 2000, p. 4). To address this sensory gap, some CBIR research considers adding a "human in the loop" (Rui, Huang, & Chang, 1999; Rui, Huang, Methrotra, & Ortega, 1997) to what began as a fully automated, computer-vision, pattern-recognition research arena. In contrast, Chu suggests that these approaches are grounded in different domains – computer science (automatic processing of physical content) versus information science (manual production of verbal descriptions) – and he argues that collaboration and integration would greatly advance both approaches. Most recently Wu, Xu, Yang and Zheng (2005a) have suggested integrating vision cognition theory into the computing framework by selecting image areas based on a model of retinal attention, by identifying visual probabilities grounded in perceptual hypotheses, or by using "structured knowledge about the physical world to make inference and eliminate ambiguities of images" (p. 1266). Domain integration is also evident in the growing CBIR research trend applying text-based computational mechanisms for relevance feedback, as well as considerations of user feedback at the interface level.

Smeulders et al. (2000) indicate that the content-based and description-based approaches demonstrate an overlapping continuum between the development of semantic and syntactic representation methodologies (see Figure 1.3). The traditional text-based approach – a high-level, user-centered approach – begins with conceptual interpretation of the perceptual content of the image/object and moves toward broader or more general terminology. The computational approach used in CBIR digitizes the image/object and employs basic color physics to identify image syntax and object geometry, using the result to identify similar patterns in other images. Current CBIR research is concerned with the semantic categorizing of these patterns for subject identification, attempting to bridge the sensory gap. Text-based research, however, has yet to move past the perception of pattern similarity in results judging, resulting in a semantic gap with high-level features. Both approaches suffer from a lack of domain knowledge: CBIR does not understand user perceptions underlying conceptual semantics; and text-based approaches do not understand analysis based on visual syntax.

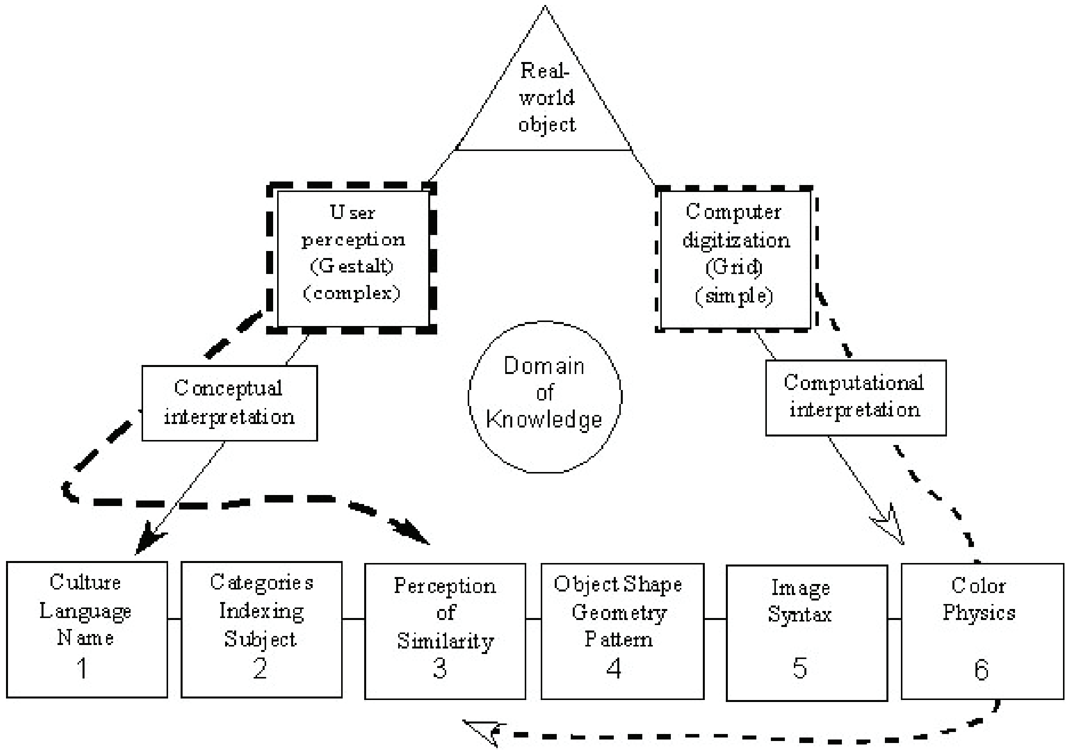


Figure 1.3. Two main approaches for developing representations of images. In an IR system user-centered and computational approaches both have a sensory gap and a semantic gap that exist between the user-centered conceptual approach to image retrieval (represented on the left from user perception to perception of similarity, 3) and the computational approach of CBIR (represented on the right from computer digitization to object shape geometry, 4). Traditional text-based (1, 2) image retrieval has yet to explore terminology for the perceptual characteristics of images (3, 4). CBIR research has focused initially on pixel differentiation and relationships (5, 6) and is just beginning to explore object identification (4) as perceived by the searcher (3) producing a terminological link (1, 2). (Interpreted from the ideas of Smeulders et al*.,* 2000.)

Current methods for automatic indexing using CBIR technologies can increase effective access to images but only by placing the onus of image description on the searcher and only if the searcher understands the dimensions of image physicality and can effectively close the semantic gap. There is the potential to enhance the precision of image retrieval if the searcher can be provided with a vocabulary for visual structure. For example, the vocabulary might describe both the color ranges of pixel groupings and their relative locations within the digital plane of the picture. The research presented here explores the possibility of developing a user-generated perceptual vocabulary that could be used by searchers during the search process, as well as CBIR developers, to represent the visual characteristics of a two-dimensional image, incorporating both image syntax (as the pre-semantic awareness of perceptual characteristics) and design characteristics (as internal contextuality).

1.2 The problem of words and pictures

Since the Renaissance, pictures have been considered subordinate to text. The relationship between language and visual images has been fertile ground for discussions focusing on both the similarities and the differences between these two mediums of expression. Arnheim (1969) begins his exploration into the nature of visual thinking with the observation that "the arts are neglected because they are based on perception and perception is disdained because it is not assumed to involve thought" (p. 3). In summarizing the writings of Cassier on the discrete yet interdependent realms of language, art and science, George (1978) states that:

> On a scale of abstract to concrete, objective to subjective, art occupies the end: primitive, concrete, creative; while language and science together are seen to occupy the extreme: advanced, abstract, critical. Language and science are abbreviations of reality; art is an intensification of reality. Language and science depend upon one and the same process of abstraction; art may be described as a continuous process of concretization. (p. 139)

George observes that language and science are more valued in Western culture because they are considered to embody the rational. Bolter (1996) acknowledges that images have historically been subordinated to words; but he argues that words, or text, are being challenged by the availability of electronic imaging, particularly in the sciences. He notes that the traditional role of ekphrasis, the verbal description of an artistic object or visual scene as the dominant referent of an image, is being reversed as images themselves become more prevalent in the electronic medium and are used to supplement or even supplant words. The art historian Danto (1998) argues, however, that ekphrasis is not "merely the equivalent of an image" (p. 10) because it not only describes how a picture looks but explains what it means, hinting at the tie between the internal structure of an image and its conceptual interpretation.

In art history, the tendency has been to emphasize the rhetoric of iconography (i.e., meaning) and the historical significance of style over the formal analysis of structure. Similarly, in discussions of image retrieval in information systems, the focus has been on verbal representations of conceptual content, that is, the ofness and aboutness of the image (Jorgensen, 1998; Layne, 1994; Rasmussen, 1997; Small, 1991). There are, however, inherent problems with the objective identification of conceptual content, as George's statement indicates, precisely because subjectivity is thought to be a characteristic of art and not of language or reason.

Foucault (1970) discusses these problems in his consideration of the painting *Las Meninas* by Velasquez (Figure 1.4). *Las Meninas* exemplifies, for Foucault, "the representation of classical representation" (p. 16) as the imitation of physical reality by photographic replication, which was the standard painting style up until the 16th century. Searle[3] (1980) observes that "the firm ground of pictorial realism (classical representation) begins to slip away" (p. 250) when the viewer considers the reflection in the mirror on the back wall of the painting. Thus, for the casual viewer, this painting seems to be *of* the artist painting a canvas and *of* the young girl with her attendants and her dog. But the canvas in the painting is not visible to the viewer and the mirror's reflection of the King and Queen of Spain suggests that *Las Meninas* may be *about* the royal couple. Searle comments that the surface features are indeed "the representation of classical representation" but that these features introduce another level that combines point of view with resemblance, the traditional visual aspects that imitate space: The artist has painted the scene the viewer sees, but the painting cannot be a mirror of reality because the artist himself is in the painting. Searle stresses the physicality of the picture's composition and concludes "there is no way to answer the question What is the picture a picture of? that does not include reference to the picture" (p. 258).

Historically, information storage and retrieval systems have been completely dependent upon *ekphrasis* for the representation of images. But Eco (1996) observes that, "even if it were true today that visual communication overwhelms written

---

[3] Searle (1980, p. 253) notes the paradox that we see the picture from the point of view of another spectator who also happens to be one of the subjects of the picture, like Escher's staircases that rise to end in another ascending star, or Steinberg's men drawing themselves.

Figure 1.4. Las Meninas, 1656, by Diego Valazquez, Museo del Prado, Madrid.[4]

communication, the problem is not to oppose written to visual communication; the problem is how to improve both" (p. 298). CBIR has the potential to provide a link between the conceptual description of an image and the physicality of visual communication.

1.3 Representation

In the electronic environment, the concept of representation is the foundation for building a more detailed understanding of the differences between words and pictures. Discussions concerning visual representation in the domains of aesthetics, semiotics, and information processing focus on the relationship of pictures to reality and on the requirement for resemblance versus the use of convention (Moriarty & Kenney, 1997). Even when the focus is limited to representation for retrieval and does not include the broader implications of knowledge representation, there is an obvious bias toward text; and text-based representation techniques are applied without consideration as to how they might be modified to accommodate the unique requirements of non-text resources.

Howard (1980) argues that the dichotomy between visual and verbal representations is both untenable and unrealistic. It is more reasonable, he suggests,

[4] Retrieved July 1, 2006, from http://museoprado.mcu.es/imenig.html

to view both visual and verbal representations as types of symbols that not only carry meaning but also may have multiple or overlapping functions. Following Howard, the term *symbol* is used broadly to refer to a representation that arbitrarily stands for something else and carries shared meaning based on association, convention or resemblance. Symbol is itself subsumed by the concept *sign,* which is also a representation but does not require shared meaning (Moriarty & Kenney, 1997; Williams, 1979).

Resemblance is neither sufficient nor necessary for representation. In *Languages of Art* (Goodman, 1968), described by Howard (1980) as a statement of "the single most elegant and rigorous theory of symbols" (p. 505), Goodman argues that different ways of presenting information can imply different ways of relating: What distinguishes different types of representation is not so much the visible marks on a surface as the different sets of rules that apply to the re-construction and interpretation of those marks. Howard contends that, to carry specific meaning, lexical referents operate with symbolic functions as syntactic structures operate with semantic functions. Thus, he establishes that resemblance is not required for representation and visual-verbal symbolic representations are in relationship rather than in dichotomy.

Howard (1980) identifies three methods of symbol analysis: lexical, logical and functional. Lexical (procedural) analysis identifies the various referents of a particular symbol or set of symbols. Lexical analysis of an image would thus include description of the visual structures identified during perception – the sense stimuli of pre-semantic awareness – that participate in the formation of meaning. Logical analysis, which encompasses both the syntactic and semantic aspects of critical judgment, provides the connection between syntax and semantics in the assessment of symbol system elements. Semiosis, for example, includes the logical analysis of how symbols "provide their meaning" (Howard, 1980, p. 504). In the digital environment, logical analysis of an image involves the analysis of the code of ones and zeros that form the mechanical representation, as undertaken by CBIR methodologies, and has yet to manifest a semantic connection. Functional (propositional) analysis identifies the specific referents of a symbol correlated with meanings such as expressions, examples, or denotations. In the digital environment, functional analysis connects the process of perceptual awareness in lexical analysis with the binary structures of logical analysis, thereby bridging the semantic gap.

Information science distinguishes three aspects of representation: the process of differentiation (physical and conceptual), the activity of representing (critical analysis), and the resulting object (Jacob & Shaw, 1998, pp. 146-147). Having been critically analyzed, the representation as object becomes the resource on which the information retrieval system operates. Given that the computer is a rule-based technology, digital representation systems focus on differentiation based upon physical rules and cues and not on the activity of critical analysis that translates perception into thought. But, for Goodman (1968, p. 227), the activity of representing is incomplete if it does not address lexical (i.e., physical) differentiation, identification of logical forms (the syntax), and the connection of the lexical perception and its logical form through functional meaning (the semantics).

In linguistics, psychology and philosophy, *representation* involves the broader arena of knowledge representation, including epistemology, semantic theory, and the primitives of knowledge structure. In contrast, *representation* in information science is concerned specifically with the creation of surrogates for resources. Lancaster (2003) observes that "The main purpose of indexing and abstracting is to construct *representations* of published items in a form suitable for inclusion in some type of *database*," (p. xx) This would include the specific case of digital databases.[5] Smith and Warner (1984) note that the representations used in information retrieval systems may refer not only to resources but also to objects, relationships, and processes. They identify five categories of representations: representations as surrogates; representations as queries; representations as resource and query relationships; representations as term relationships; and representations as resource relationships. Unfortunately, the mechanics of these representations are based on an examination of text, without consideration as to whether these categories could also apply to non-text-based representations. Because the surrogates in image databases are not text, CBIR introduces an additional type of representation in the form of query-by-pictorial-example (QBPE) pattern-matching. In information science, retrieval utilizing representations as relationships between queries and image resources is drawing greater attention (Chen, 2001; Enser & McGregor, 1993; Jorgensen & Jorgensen, 2005; Keister, 1994; Spink, 2001). Related research in CBIR is still focused on the similarity metrics of sequenced image selection by users during feedback iteration, but there is growing attention to multi-modal querying attempting to link verbal selections with signal features (Fan, Xie, Li, Li, & Ma, 2005; Santini & Jain, 1998).

The current art history perspective on *representation* is articulated by Gombrich (1960; 1972), Arnheim (1954) and Goodman (1968). Each rejects the classical notion of representation as the imitation of reality. Gombrich (1960) stresses the skills of the

[5] In digital information storage and retrieval systems, representation is concerned with the coding structures used inside a computer. However, the focus of the present discussion is not on the hardware, data standards or mechanics of converting analog data into ones and zeros, although these must be understood as fundamental characteristics of all digital representation systems (Mostafa, 1994).

artist in building structural relationships in the two-dimensional plane and the "faithful construction of a relational model" (p. 90) of the world of experience rather than simple identification of visual elements. While he rejects art as imitation, Arnheim (1954) theorizes about the close relationship between reality and the creation of structurally equivalent forms, exploring the human perceptual experience as a highly skilled process rather than a "passive imprint" (Golomb, 1994, p.19), such as a photograph. And Goodman emphasizes the human component in representation when he describes the need to acquire skill in "reading" pictures. In contrast, CBIR ignores the human perceptual experience and attempts to analyze at the structural level what the naïve user overlooks – or goes beyond – in the passive perception of pre-semantic awareness. Thus, the user must become aware of her perceptions in order to bridge the semantic gap between the structural processing of CBIR and the higher level of semantic querying. Text-based indexing, in contrast, totally neglects visual configuration and structural processing.

Representation in art criticism has been dominated for the last 60 years by *scrutiny theory*, as discussed by Wollheim (1993). According to scrutiny theory, the viewer confines herself to what can be known by looking at the marked surface, thus differentiating between representation (or understanding the conceptual composition of an image) and interpretation (or grasping its meaning or contextual description). Understanding involves rules or conventions and is concerned with imposed construction. In contrast, meaning is derived from the interpretive process, through discovery, based on intention. Moriarty and Kenney (1997) extend Wollheim's ideas into a theory of representation for the visual arts:

> … a necessary condition of R representing x is that R is a configuration on which something can be seen and furthermore one in which x can be seen. Sufficiency is reached only when we add the further condition that R was intended by whoever made it to be a configuration in which x could be seen. (Moriarty & Kenney, 1997, p. 240. Emphasis in original.)

They argue that visual representation involves not only the experience of the viewer, but also the fulfillment of the artist's creative intentions: An image is not necessarily "of" something, but always "stands for" something – some intention communicated by the artist/creator through a visual medium. Their theory of representation lays the foundation for representation as both a configuration (or object) and a planned visual communication (or intentional process): Representation in art can indicate either an object (a picture) or a process (the intentional act of picturing).

Representations as surrogates in retrieval systems are referenced as *resources*, which are defined by the World Wide Web Consortium as anything that can be identified by a Universal Resource Indicator (URI).[6] Resource representations in information systems are not addressed by any specific body of literature. Rather, there is literature dealing with knowledge organization that covers issues of representation, including classification schemes, thesaurus construction, indexing and abstracting. The terminology of this literature is often confusing, ambiguous, and imprecise: the term "classification," for example, is used to refer broadly to any process for representing and organizing knowledge (i.e., categorization) or, more narrowly, to the specific ordering of well-defined and mutually exclusive classes (i.e., classification, see Jacob, 1992). For the purposes of this discussion, the term *representation* will be used to refer to the entire range of indexing techniques that includes faceted and enumerative classification schemes, subject heading lists, indexing by assignment from controlled vocabularies, and indexing by extraction of keywords.

Traditional text-based surrogates include bibliographic records, descriptors (including classification labels), keywords, location codes, and abstracts. Though non-text resources such as pictures and video generally rely on text-based representations, in the electronic environment there is no clear distinction between the data record and the resource for which it serves as surrogate. The data record is an aggregate of bits that may point to a digital surrogate for a resource or to the resource itself. These documents might be books, which are considered textual even if they contain pictures, or they might be non-text objects such as videos, physical objects (e.g., incunabula or sculpture), or digital images.

The nature of visual representations as information retrieval surrogates requires close examination. According to Wartofsky (1979), in the physical realm, cultural artifacts are distinguished from naturally occurring phenomena. Natural phenomena may be selected as artifacts, however, as when stones are intentionally sited within a rock garden or shells are collected and arranged for display on a table. Nature may also be approached pictorially, as when a sunset is viewed as a "picture." Expanding on Howard's (1980) notion of representation-as-symbol, Wartofsky declares all artifacts to be representations in that they are symbols (in the broadest sense of representation) that embody intentionality: "… a spear is made for hunting and also represents the mode of action of the hunt; a spear picture is not made for hunting but is made expressly as a

[6] Uniform Resource Identifiers (URIs) identify Web resources (documents, images, downloadable files, electronic mailboxes, etc.). See http://www.w3.org/Addressing/

representation of a spear" (p. 282). The two artifacts may thus carry a similar message of intentionality although this intentionality is embodied differently. Artifacts defined as embodied intentionality may be visual or non-visual. And, while non-visual artifacts may be auditory or tactile, visual artifacts rely on the sense of sight and are not limited to pictures but include sculpted objects, physical movements or hand signs.

For Twyman (1979; 1985) intentionality involves mediation and/or creation. In line with Wartofsky's (1979) requirement for embodied intentionality, Twyman contends that the mere act of seeing is not sufficient to create an image. Though the eye filters visual information, it is often without conscious intentionality:

> It is estimated that 75 percent of the information entering the brain is from the eyes, and that 38 percent of the fibers entering or leaving the central nervous system are in the optic nerve. Current research indicates that the eyes have 100 million sensors in the retina, but only five million channels to the brain from the retina. This means that more information processing is actually done in the eye than in the brain, and even the eye filters our information. (Hanson, 1987, p. 39)

The brain then transforms the remaining information transmitted from the eyes into meaningful images, or messages:

> Studies indicate that vision is about one-tenth physical and nine-tenths mental. In visual perception, sensory input in the form of light patterns received by the eye is transformed by the brain into meaningful images. The interpretation depends on preconditioning, intelligence, and the physical and emotional state of the viewer. … The variety of our responses to visual stimuli is demonstrated by artists. Twelve people depicting the same subject – even from the same vantage point – will create twelve different images because of their different experiences, attitudes, interests, and eyesight. (Preble & Preble, 1985, p. 11)

For example, when a landscape is surveyed from the top of a skyscraper, the act of perceiving creates an internal, mental representation. According to Twyman (1979), however, this representation is not an image until it is either mediated, as in framing to produce an aerial photograph, or created, as in drawing or painting to produce a visual artifact.

When the definition of symbols as intentionally created visual artifacts is integrated with Howard's argument that the notion of resemblance implies shared meaning, a *visual representation* can be defined in this research context as an intentionally created two-dimensional (2D) artifact that stands for something else based on association and convention, but not necessarily on resemblance.

1.4 Definitions

> "… in the humanities as in the social sciences, there is often difficulty over labels for concepts. The same word may be used in different senses by 2 (sic) writers in the same field or a single concept may be identified by several different terms. This is also true of researchers in the field of communication who have latched onto convenient words from everyday speech or from foreign cultures and endowed them with precise, narrow definitions which suit their purposes – words such as sign, symbol, icon, code." (Morgan & Welton, 1992, p. x)

The domains of art, information science and computer science use similar terminology to express different perspectives on the concepts related to non-verbal phenomena and their manifestations in the digital environment. For example, the term *representation* has already been used here as the imitation of physical reality (art), as the ofness or aboutness of conceptual description (information science), and as binary content description (computer science). This section identifies germane cross-domain terms and defines their use in the present context.

In its current usage, *image* is defined in line with Twyman's (1979) definition of graphic language as intentional, two-dimensional, visible communication. In line with the arguments of Goodman (1968) and Howard (1980) that resemblance is neither sufficient nor necessary to define an image, the term *image* will be used to denote visual representations in general but does not include three-dimensional (3D) objects (e.g., sculpture), and will not denote reliance on resemblance. Defining the term *image* broadly so as to subsume the variety of mediated responses described by Preble and Preble (1985) – but not the viewer's preconditioning, intelligence, and physical or emotional states that may have led to a response – eliminates the need to consider "mental representations" and can account for image variability.

In the literature on non-verbal phenomena, the terms *image*, *graphic*, and *picture* are frequently used interchangeably or with topical specificity. While use of the term *image* does not imply resemblance, the term *picture* is used here to refer to an image that does rely on resemblance. When its use is confined to the digital domain, the term *graphic* may refer to text entities such as pictures of the alphabet; to any non-text entity[7]; to raster-based as opposed to vector-based[8] non-text entities; or, conversely, to an informational chart rather than a raster-based rendering. A *graphic* is defined here as an image that is less dependent on resemblance than a picture, moving toward the purely iconic, and may also apply to a resource that contains a combination of text and images, whether iconic or pictorial. The terms *image*, *picture* and *graphic* are not necessarily mutually exclusive; for example, many corporate logos, such as those of IBM and 3M, include alphanumeric characters but are commonly accepted as pictorial conventions (icons).

Due to the nature of the digital environment, an image can be either static or dynamic. Both video (and its technological precursor, film) and animation are examples of dynamic images. In research regarding representation issues in film, there is ongoing discussion (Barthes, 1985; Carroll, 1982; O'Connor, 1996) as to whether the basic coherent unit is the static frame (or still) or the dynamic shot (or scene). For the purposes of this discussion, the *frame* as static image will be considered the basic unit in the digital environment: Because they are produced by the medium, dynamic images are understood as a specific format in which static images occur.

There are numerous terms that describe the concept of form as applied to images. In everyday conversation, many of these terms are used interchangeably; but, within specific domains, individual terms are applied more carefully. According to Hurlburt (1977), the publishing activity of page layout exemplifies the notion of form defining structure through the characteristics and limits of the medium. In two-dimensional graphic design, the terms *form* and *shape* demonstrate specific differences in that *form* implies the three dimensions of length, width and depth, while *shape* indicates only length and width (Ragans, 1988). The AAT uses *form* to describe "the arrangement of visual elements such as line, mass, shape, or color" (Petersen, 1994). Using the concept

[7] In the digital environment, the definition of image as a visual representation that may or may not depend on resemblance accords with Twyman's (1979)definition of graphic language in that any distinction between verbal and non-verbal elements does not apply. To generate an image on a digital display, both verbal and non-verbal elements are indiscriminately represented as ones and zeros (the code) and the identification of displayed elements as text or image is not a representational issue.

[8] *Raster* graphics refers to the most fundamental mechanism of digital display. The image seen on the display (e.g., cathode ray tube or liquid crystal display) is made up of individual dots, or pixels. Pixels are square in shape resulting in a diagonal line having a stair-step effect when viewed at lowest resolution. *Vector* graphics is technically more sophisticated than raster graphics in that the image is stored in geometric format rather than as a set of pixels: The image is made up of lines that are adapted to the resolution of the display, rather than being stored at a specific resolution as a specific set of pixels. This minimizes the stair-step effect. Ultimately, however, to mechanically produce the visible image, vector graphics are dependent on the number of pixels available on an individual display device.

of form as embodied intention, Focillon (1948) contends that, in fine art, form is the "modality of life," a graph of activity that is inscribed in space and time.

In the present context, *shape* is retained in the sense of two-dimensional length and width, but the term *form* is defined broadly as embodied intention. As such, it has both primary and secondary features. *Primary features* (Belton, 2002) are those elements of form which in themselves are not a matter of semantic significance and include color, dimensions, line, mass, scale, shape, space, texture or value; *secondary features* are the relations of primary features with one another and include balance, contrast, dominance, harmony, movement, proportion, proximity, rhythm, similarity, unity, and variety. *Structure* consists of these secondary features and is synonymous with composition. Structure gives form its semantic significance.

The terms *texture* and *pattern* reflect closely related concepts and are dependent on color. In the present context, *color* is defined as a primary element and refers to all hues, including black and white, and their characteristics of purity (saturation) and tone (grayscale) – the basis of pixel differentiation. *Texture* is an important element in CBIR research (Lin, Chui, & Yang, 2001; Ma & Manjunath, 1998) but a difficult concept to grasp due to its application in different domains. The AAT defines *surface texture* as the "tactile and sometimes visual quality of a surface given to it by the size, shape, arrangement, and proportions of its minute parts" with attributes such as "roughness, smoothness, granularity" (Petersen, 1994). In the domain of graphic design, texture is "the representation of the structure of a surface as distinct from color or form" (McCreight, 1996, Texture, para. 3) and can be divided into three categories: "literal … remembered or implied … (or) a complex or vague pattern" (McCreight, 1996, Texture, para. 1). In the research literature of information science, definitions of texture focus on homogeneity: *texture* is defined as "a homogeneous pattern" (Smith, 2001, p. 971) or "the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity" (Rui et al., 1999).

Although definitions of *texture* used in CBIR and pattern recognition focus on homogeneity, there is confusion about the relationship between *texture* and *pattern* when considered in general. In the Morelli CBIR system, the basic 4x4 unit varies only by the organization of on/off pixels, yet these units combine to form patterns (Stenvert, 1992; Vaughn, 1992). *Pattern* is defined variously as "a spatial arrangement of pixels" (Smith, 2001, p. 971); "a periodic series of visual elements" (Razel & Vat-Sheva, 1986); or "a representative sample - a plan, diagram, or model" (McCreight, 1996, Pattern, para. 3). The AAT distinguishes between *patterns* as "design elements composed of repeated or combined motifs" and *motifs* "which are distinct or separable design elements occurring singly as individual shapes"(Petersen, 1994). Given that the digital image is a two-dimensional pixel representation of three-dimensional objects or scenes with surfaces, Caivano (1990; 1994) describes a system for ordering textures based on properties such as size, saturation, density and directional organization (see Figure 2.4). Technically, texture is created from patterns or motifs. For this reason, *texture* will be used here as synonymous with *pattern*, with *pattern* as the preferred term.

An *image domain* is defined as a collection of images that have either conceptual coherence or mediated visual similarity. A conceptually coherent domain is exemplified by any collection of images that share concepts: for example, a collection of airplanes, of dogs, or of landscapes. Examples of mediated domains include satellite imagery (SAT) (photography taken from a satellite and representing macro views of the earth); and photo-microscopy (MIC) (photography through a microscope and representing micro views of organic and inorganic materials.) [9] Fine art is a domain that can be based on mediated characteristics, such as drawing or painting, as well as having sub-domains of conceptual characteristics such as style. For example, abstract art (ART) is "essentially the result of the artist's unique, visual perception, given free form by his powers of selection" (Arnason, 1968, p. 163). More specifically, abstract expressionism represents "interacting color and line shapes from which all elements of representation and association seem to have disappeared" (Arnason, p. 173). The Web is not a conceptual collection but a mediated collection. However, as a mediated collection of resources, the Web frequently displays little or no basis for visual similarity and thus does not qualify as an image domain under the definition used here.

For the purpose of this study, the following definitions have been adopted. *Non-experts* are subjects with no special training in a specific image domain; and *special training* is defined as college level courses in art, architecture, photography or geographic mapping. A *natural language vocabulary* consists of words or phrases, generated by an information searcher without recourse to standardized or controlled vocabulary. A *word* or *phrase* is a natural language unit that represents a value.

A *controlled vocabulary* is a set of mutually-exclusive and non-overlapping concepts (properties and values[10] each of which exists in a 1:1 relationship with a linguistic label. A *term* is an authorized label that may represent a set of synonyms

---

[9] See examples in Appendix K.

[10] Many researchers refer to properties and values generically as *attributes*, but the preferred term used here is *property* with specific *values.*

and/or near-synonyms. A *stop word* is a word that is excluded from consideration as a preferred term due to lack of focus or specificity to the domain of the controlled vocabulary. A *property* (or *facet*) is a category that consists of a set of values grouped on the basis of similarity. A *value* (or *isolate*) is an instance of a property and may represent a single natural language term or a set of synonyms and/or near-synonyms. For example, *crimson*, *scarlet* and *rose* constitute a set of synonymous words that are represented by the term (value) *red* (preferred) which is an instance of *color* (property).

A *concept* is a category term. Within a controlled vocabulary, facets and isolates that represent more than a single natural language term are considered concepts, that is, categories of entities and/or properties. Terms may be introduced to the controlled vocabulary as *conceptual antonyms* when it is necessary to make existing concepts meaningful; for example, *negation* is meaningless without *affirmation*. *Conceptual organizers* may be introduced into the controlled vocabulary as superordinates to organize groups of related concepts.

1.5 Contribution to theory

Innovations in visual technologies point up the need to explore image representation methodologies to enhance information retrieval systems. Current CBIR systems (see Appendix G) make no attempt to describe or categorize the basis for similarity judgments generated by query-by-pictorial-example searches. In point of fact, CBIR research has yet to address the connection between the content or problem domain and the internal structure or organization of the image. MacDonald-Ross (1989) describes this possible connection as one of "the most profound and most important questions in graphic communications" (p. 149). The research project reported on here has explored the possibility of developing a user-generated vocabulary that can be used to describe the visual structure of images. Such a vocabulary has the potential to facilitate more precise feature queries and to capture criteria for similarity judgments when CBIR technologies are involved, whether alone or in combination with existing high-level, concept-based image retrieval systems. As such, this research contributes to current capabilities in support of more effective image search and retrieval.

CHAPTER 2
Literature Review

The literature review explores the interaction between written and visual communication: the combination of text-based and image-based representation languages that has the potential to make retrieval in image-based information systems more efficient and more effective. It discusses the relationship between form and function, reviews explorations of physical structure that have been undertaken by information system designers and researchers, describes current attempts to represent non-text based objects in the digital environment, and identifies research relevant to the creation of a vocabulary for visual structure.

2.1 Information retrieval systems

Because the primary purpose of electronic information systems has been the storage and retrieval of text-based resources, they have necessarily fostered little discussion of the relationship between the physicality of an image and its semantic attributes. Research has tended to focus on three basic and interrelated components of an information retrieval system: the users and their queries; the documents and their associated representations; and mechanisms for comparing user queries with document representations. The latter is described as a "match" system (i.e., a string-match as opposed to a descriptor based system).

Soergel (1985) describes an information system as involving the identification of problems; the acquisition of information resources to solve those problems; the development of a means to represent the resources; storage of the information resources and their representations; and the delivery of information and/or resources to those who request it. Thus an information system requires the "intertwined aspects of intellectual organization and technological implementation" (Soergel, 1985, p. 5). The technical side involves mechanisms for storing and retrieving resources and describes the mechanics of the system. The intellectual side involves the identification and representation of resources, presumably by a human, which will be processed by the system. From the technical side, resource representation is the necessary component serving as the bridge between the technical and the intellectual by providing a means for the mechanical matching of query to representation. The technical capability for matching a query term depends on both the existence of a digital surrogate, consisting of ones and zeros, and the assignment of text descriptors, which also consist of ones and zeros. The popular application of full-text indexing in the Internet environment is a purely mechanical operation that indicates the number of occurrences of words within a textual resource and involves no intellectual component. The focus of the current discussion is on indexes[11] as surrogates for intellectual and biographical content, rather than the mechanical indexes that are essential to the development of computing systems.

Much of the research on information systems reported in the scholarly literature focuses on human-computer interaction and system interfaces (e.g., Mostafa, 1994; Nake, 1994; Steuer, 1993); on the searcher, the intermediary and the query process (e.g., Belkin, 1990; Kuhlthau, 1991; Pappas, 1994); and on algorithmic methods for retrieving relevant resources (e.g., with respect to image resources, Besser & Snow, 1990; Busch, 1994; Caivano, 1990, 1994; Gecsei & Gecsei, 1989; Gouge & Gouge, 1996; W. Vaughn, 1992). Despite the integral relationship between the representation and its retrieval system, few attempts are made to link the document representation literature with the information systems design literature. Notable exceptions are Hirschheim, Klein, and Lyytinen (1995), who discuss data modeling as the extension of document representation, and Pollitt (1999), who equates relational database design with faceted classification. Perhaps one reason for the lack of such a linkage is that many text-based systems and databanks already exist and research problems involve re-design. Another reason, as suggested by Jacob and Shaw (1998), is reflected in the notion of anomalous states of knowledge (ASK)[12], an approach founded on the assumption that the "traditional classification and indexing languages seem not to be designed as good representations of either (need or text), but rather as available, convenient intermediate languages" that are "bound to fail" (Belkin, 1977, p. 189). Thus the focus of development and research has been on the mechanics of the retrieval system

[11] No matter how the index is achieved, the technical component of the information system uses an index as the basis of the match. The differences, however, between automatic and intellectually assigned indexes have an important impact on recall. An automatic index applies a mathematical formula to identify and count word occurrences to characterize a document. This is referred to as full-text indexing, natural language indexing, or the use of uncontrolled vocabularies. An intellectually organized index relates terms with similar meanings or alternate spellings. As Soergel (1985, p. 213) points out, using an automated index to search for *harbors* will not return documents that use the terms *harbour* or *port*. The searcher must remember to construct the search to include alternate terms, whereas an intellectually constructed index can relate those various terms for the searcher. QBPE/CBIR is also a purely mechanical process but the current focus is on how to link this process with verbal representations.

[12] "The ASK hypothesis is that an information need arises from a recognized anomaly in the user's state of knowledge concerning some topic or situation and that, in general, the user is unable to specify precisely what is needed to resolve that anomaly" (Belkin, Oddy, & Brooks, 1982, p. 89).

itself, including the interface, because the representational structure has proven to be too difficult a problem. Jacob and Shaw (1998, p. 147) point out the need to (re-)consider "the problem of representation from a different perspective." Although the argument for reconsideration of representational systems was originally directed toward text-based information resources, it can be applied to the visual realm, given the dependence of retrieval systems on the "convenience" of verbal representation.

2.2 Content-Based Image Retrieval (CBIR)

When images were first collected in digital repositories, the initial approach was to apply existing alphanumeric database technology to their management and retrieval. Two problems arose, however, as these collections grew larger. First, digitized images, particularly at high resolutions, are very large files; and, second, the nature of the image data structures did not work well with existing text-based retrieval methods that relied on string matching. It was necessary to identify new approaches to the indexing of images – approaches that were both high performance and efficient, given the size of the data being manipulated, while simultaneously providing support for image browsing.

Over the last fifteen years, there have been a number of proprietary or collection-centered attempts to develop visual information retrieval systems (VIR)[13], many originating with art and museum collections. Widespread access to the World Wide Web, in concert with the recent rapid advances in imaging technology, has focused attention on the development of more general systems containing a broad selection of images from disparate domains. Some of these systems are characterized by new approaches, including identification and application of retrieval properties that are made possible by content-based image retrieval (CBIR) technologies (see examples in Appendix G).

The goal of CBIR is to retrieve a manageable number of images for browsing by the user. CBIR refers to a set of pattern-based methods for pre-filtering queries in order to greatly reduce the number of target images that need to be searched. As implemented in many systems, it is a method for searching for images based on the similarity of global image properties (e.g., a single color value for an entire image). Such an approach generally starts with a rich set of low-level or generic visual properties and attempts to derive a higher-level semantics by applying domain knowledge provided by users or applications.

[13] Visual information retrieval systems were referenced in early 1990s literature as VIS, and later as VIR. VIS has been adopted by some groups to refer to the 'visualization' of information, which generally is not image information but an interface technique (Torres, Silva, & Rocha, 2003) that can be applied after CBIR technologies identify "similar" images (Rorvig & Hemmje, 1999). Most recently, the VIRAGE Company (a CBIR development company) has begun using the name VIR to reference a specific search engine.

Current CBIR methods of pattern-matching are similar to natural language, full-text searching in that there is no intellectual component. As such, content-based approaches are viewed as complimentary to keyword approaches and are not intended to replace them.

The low-level properties of a digital image are dependent on the digitization process that maps an analog image to a pixel grid. This technique is well established for verbal representations where sets of pixels are mapped to alphanumeric strings. Each letter or number code is individually manifested as an array of pixels[14] to create the image for digital display. Mechanically, each pixel is either on or off so that the spatial arrangement of pixels creates meaningful forms (see Figure 2.1).

Figure 2.1. The perception of pixel configurations as letters. The construction of digital letters is grounded in the gestalt psychology principle that similarity and adjacency of forms result in their integration or grouping (Beebe, 1986).

Before CBIR became a specific domain of research (Rui, Huang, & Chang, 1999; Smeulders et al., 2000), the Morelli system[15] (Stenvert, 1992; Vaughn, 1989, 1992) was proposed by art historians to explore image properties through reliance on analysis of form. Morelli matched, sorted and ordered pictures on the basis of visual patterns derived from the digitization process (Vaughn, 1992), as demonstrated in Figure 2.2.

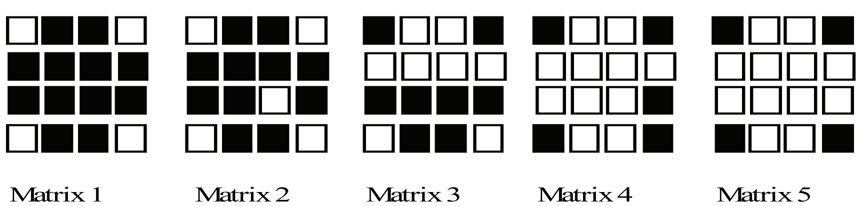


Figure 2.2. Pixel matrices typical of the Morelli system. Each specific image consists of multiple sets of basic pixel units. The four-by-four pixel matrix is the basic unit that the Morelli system used to compare digital images. Matrix 1

[14] At the most basic level, pixels are equivalent to picture elements because they conform to gestalt principles of structural relationship in order to form text and images.

[15] By *design,* the author implies composition, style and pattern (texture).

> was compared with matrix 2, then with matrix 3, etc., with each pair receiving a comparison score: Matrix 1 to 2 compares at 94%; matrix 1 to 3 compares at 50%; matrix 1 to 4 compares at 6%; matrix 1 to 5 compares at 0% (Vaughn, 1992, p. 13). Thus the results of a specific query-by-pictorial-example (QBPE) search in Morelli were based on the calculation of similarity percentages across aggregated sets of basic units. These calculations became increasingly more complex as multiple matrices were combined.

The Morelli system attempted to provide an objective means for describing and identifying pictorial characteristics, such as configuration, pattern, tonality and color, and for performing pixel comparison during the image search. The procedure was performed at ever more granular matrices, down to the level of 4x4 pixels. Each matrix recorded a value for configuration (edge clarity), for tonality (light), and for color. Form identification was thus based at the pixel level of object shapes (the parts) within the image rather than on relationships among objects (the arrangement). Stenvert (1992) described Morelli as an initial attempt to promote a scholarly, morphological study of images using a calculated abstraction based upon pixel arrangement:

> It must be remembered that the digital image does in itself perform an analysis on an image. It divides it into a series of units. As it happens, the categories used for these do correspond to the three basic areas of pictorial experience mentioned by Renaissance art theorists, namely design, color, and tonality (or chiaroscuro). This approach results in a means with which, for the first time, art historians would have ways of classifying pictures visually, rather than verbally. (Stenvert, 1992, p. 28)[16]

The Morelli system performed the process of pattern recognition without ascribing contextual relationships between the patterns: "it simply uses the means of evaluating position and tonal intensity that are established conventions of a digitized matrix" (Vaughn, 1992, p. 95). It is this rudimentary pattern matching that allowed the Morelli system to handle various types of images. Vaughn describes this rudimentary pattern matching as a kind of *visual syntax*[17] that operates in a fashion similar to full-text retrieval.

The Morelli database contained a very specific set of domain images representing

[16] By *design,* the author implies composition, style and pattern (texture).

[17] Vaughn bases this terminology in the art criticism tradition of formalism, which analyzes the relationships of the primary visual elements of a picture, independent of labels, conventions or meaning (see discussion by Feldman, 1967, pp. 457-460).

a collection of traditional paintings, and its associated retrieval system supported the juxtaposition of visual properties and verbal classification. Access to works of art is generally through the domain of art history, where representations of images are based on author, title, period, etc. Selection of an image in the Morelli system used this verbal access mechanism in that the process began with identification of an exemplar image retrieved by a semantic query. This exemplar was then used in a query by pictorial example (QBPE) search that visually matched the exemplar (or query) image to similar images. For example, a search on the term *portrait* would return a large set of half-length portraits (see Figure 2.3 for an example). A Rembrandt self-portrait selected as the query image might return a Rubens portrait that was compositionally very close to the Rembrandt. This association between the Rembrandt and the Rubens might otherwise have occurred only through chance browsing, as there would be no terminological overlap in the titles of the paintings (Vaughn, 1989).[18]

Figure 2.3. Sample of search results from the Leiden 19th Century portrait Database. The QBPE exemplar image is on the left, followed by four images from the results set (image numbers, left to right: c00140f-5, c001983f-2, c001984f-4, c005433f-1, c003506f-3). Reprinted with permission: Nies Huijsmans.

Although Morelli and CBIR were both topics of study in the early 1990s, work in art history with the Morelli system did not overlap with research on CBIR and pattern matching in computer systems. CBIR had its intellectual roots in two earlier research

[18] In addition, CBIR technology is often used in a similar way in medical applications. For example, recent advances in CAT Scan technology have made possible the acquisition of high-resolution images of the heart in which important anatomical features such as the coronary vessels are visible. The state of the art in MRI imaging, on the other hand, typically provides lower resolution but can facilitate the differentiation of viable (healthy) cardiac tissue from injured tissue. The registration of CT and MR images of the same heart permits the association of anatomy with function. As an example, given an injured region, the coronary vessels responsible for irrigating that region may be isolated and targeted for treatment (e.g., coronary bypass) (O'Donnell, Aharon, Halliburton, & Gupta, 2000, p. 790)

areas concerned with image retrieval: database management[19] and computer vision. Historically, the management of image databases has been a text-based problem in that images are assigned linguistic representations. Even though some databases contain the actual images, they are only accessible through linguistic descriptors that have been assigned by human effort and are thus open to individual subjectivity. Computer vision involves *image processing*[20], which aims, through automated processes such as applying curve smoothing, to compare images or to "convert one image into another such that properties that are most relevant for a certain *task* become more pronounced or more explicit," (Florack, 1997, Emphasis in original.) Although CBIR depends on techniques of database management, it has evolved from image processing to *image analysis* using techniques that seek to understand an image through the development of higher-level representations that can be linked, in turn, to visual semantics. The result of image analysis research has been the exploration of information elements (physical properties such as color); representation metrics (such as histograms); data models (based on mathematics); indexing methods (such as mathematical *topology*[21]); query formulations (such as browse) (Grosky, 1997); and, most recently, the linkage of linguistic queries to physical properties, data models, metrics, and topologies, (e.g., Goodrum et al*.,* (2001), Rorvig (1998), and Smeulders et al. (2000). These areas represent significant research agendas that must be addressed before CBIR can become a mainstream mechanism for the retrieval of digital images.

The forty-three CBIR systems identified by Veltkamp and Tanase (2001) in their survey are generally research-demonstration systems available from university and vendor developers. A few of these are production implementations that are available on the Web (e.g., the Leiden Portrait Collection search results pictured in Figure 2.3[22]). Several other Web-based production systems that use CBIR have been attempted but

---

[19] This is an area of VIS research for which the main focus is the system architecture and database structure regardless of the specifics of the feature sets. Most of this research assumes that manual terminological indexing of images is not scaleable; that textual descriptors are inadequate for features based on visual data; and that such indexing is ineffective in supporting unanticipated user queries (Huang, Methotra, & Ramchandran, 1997, p. 101). The general goal of this type of research is to develop effective multimedia information management techniques. These efforts include query languages such as SQL (Roussopoulous, Faloursos, & Sellis, 1988), computer vision (Brolio, Draper, Beveridge, & Hansen, 1989), image manipulation (Chang, Smith, & Meng, 1997), and system design (Borgdona et al., 1990; Huang et al., 1997). This area is beyond the scope of the present research.

[20] In education, *image processing* sometimes refers to *visual literacy.*.

[21] In mathematics, *topology* is the study of the properties of geometric forms that remain invariant under transformations.

[22] Other examples are listed in Appendix G which contains a bibliography of URLs for CBIR systems.

ultimately abandoned (e.g., Yahoo Image Surfer, Alta Vista PhotoFinder, and San Francisco Museum of Art [SFMA][23]). Second generation systems such as Yahoo! Images and Google Images are based primarily on surrounding metadata rather than CBIR technologies (Datta, Li, & Wang, 2005). The common feature of all CBIR systems is the extraction of "a signature for every image based on its pixel values and (the definition of) a rule for comparing images" (Wang, Jia, Chan, & Weiderhold, 1999). These signatures are composed of image properties that are defined within the domain of CBIR design and used as the basis for image resource queries.

There is, however, little consensus as to which properties are most appropriate in the retrieval environment. Initial attention has focused on color because it is the basic property differentiating pixels. Color is calibrated according to a color space[24], such as CIE and RGB, which assigns a numeric value to each color. The RGB color space is most commonly used, perhaps because it "was designed to match the input channel of the eye" (Smeulders et al., 2000). The MPEG-7 standard supports several color spaces and includes a color space identification tag for potential cross-walking between standards.

In a CBIR system, terms are rarely used to designate colors. A notable exception is the research of Mojsilovic, Gomes and Rogowitz (2004) who are prototyping a color naming metric. Generally, however, the only means for indicating color in a CBIR query is a visual color wheel that allows for point-and-click selection of a color by the searcher and the subsequent assignment of a numeric value by the system. The most common representation of image color in CBIR systems is the color *histogram* (Rui et al., 1999), which is used to generate a representation of the image as a whole. The histogram is a global property in that it calculates a single value for an image. Histograms can be calculated for various aspects of pixel assignments such as global color, localized color, color regions, or patterns formed by color intersections. Another approach to generating a histogram is to pre-coordinate color sets and match image color metrics to set ranges of colors. In either approach, all other data about the image is lost once the histogram has been calculated.

---

[23] Yahoo used Excalibur; SFMA used QBIC, and Alta Vista used Virage.

[24] A color space includes black through all the hues to white and all the gray levels (lightness, saturation or intensity). There are several different color spaces - for example, HSV, CIE, RGB, Munsell - each of which provides for the assignment of a numeric value to a color. The Munsell scales designate colors in a sufficiently standard way as to be recommended for use in general color description (Kelly & Judd, 1976). This standard consists of 13 basic hues and 16 intermediate hues, plus black and white, for a total of 31 basic color names. Three levels of lightness (saturation) then modify each: dark, medium, light. Note these basic colors are very similar to those identified by Berlin and Kay (1969) which have variations occurring in multiple blues versus Munsell's multiple greens.

Pattern[25] is dependent on color distributions. Several properties of pattern have been identified: repetitiveness (regularity), directionality, granularity (coarseness/roughness), contrast, line-likeness, and complexity. QBIC uses coarseness, contrast and directionality (Ventura, Gagliardi, & Schettini, 1998). Computational pattern representations based on the human perceptual system are attractive in CBIR because they are visually meaningful (Rui et al., 1999); but the similarity metrics for patterns remain behind the scene and descriptive terminology is not generally provided to the searcher. Caivano (1990; 1994) explores algorithmic descriptions of patterns in computer-aided drafting (see Figure 2.4). Pattern is also an essential characteristic in the analysis of satellite or aerial imagery: A CBIR-based pattern thesaurus has been shown to help both the searcher and the system recognize similar areas, such as orchards, highways, forests or deserts, by selecting a pattern of interest and having the system find similar patterns (Ma & Manjunath, 1998). Lin, Chiu, and Yang (2001) characterize their texture clustering algorithms using six linguistic terms (coarseness, contrast, directionality, line-likeness, regularity, roughness); but term application is hidden from the user by use of a QBPE interface.

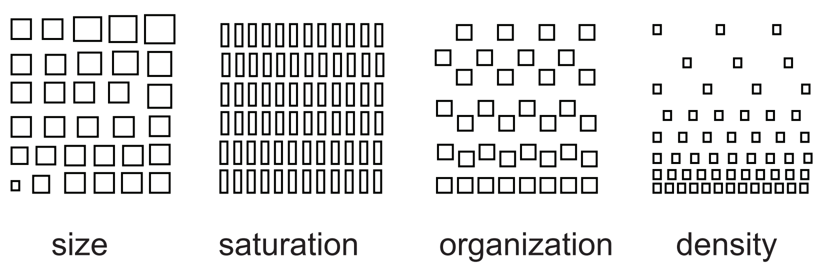


Figure 2.4. Pattern variables. According to Caivano (1994), patterns have four measurable variables: size, saturation, density, and directional organization. These variables are applied to the pattern unit or repetitive motif (e.g., the quadrilateral in this example). Caivano (1990) suggests that pattern formulas could be collected in a pattern atlas or visual thesaurus. Defined patterns would be assigned verbal descriptors for textural association: for example, smooth, rough, polished, abrasive, coarse, fine, soft, and hard. Additional descriptors could be used to identify psychological associations: for example, dryness, warmth, protection, desolation, and relaxation. (redrawn from selected examples provided by Caivano, 1994, p.76)

[25] The CBIR literature uses the term *texture*, the preferred term is *pattern*, which subsumes texture (see section 1.4 Definitions.

Color and pattern are relatively simple to quantify compared to the parameters needed for representing shape. Shapes can be either region-based through pixel clustering (areas) or contour-based through edge detection (boundaries) (Bober, 2001). But shape representation is still basically dependent on pattern matching and the identification of salient color regions based on image segmentation. One approach is to segment the image into a grid and identify a property based on the configuration of color pixels in each block of the grid. Another approach is to identify salient color regions of the image and to store the position of each such region in relation to the image frame using a set of *x, y* coordinates.

Matching shapes that are similar but not isomorphic is a major problem in CBIR. The theory and technique for mathematical analysis of spatial structures is called morphology, which includes both shape[26] and size. Two shapes may be conceptually similar but may have been morphologically translated, either by rotation or scaling: for example, the front view rotated to the side view of a horse. One approach to matching transformed shapes is to order them in terms of individual distance from a prototype shape. This process is exemplified by morphing, where the geometry of the control points in one shape is incrementally deformed to match the control points in another (Park & Ra, 1999; Sclaroff, 1997). Shape understanding or analysis of complex objects will be essential to object recognition in CBIR .

Each Web-based CBIR system currently available for evaluation makes use of different combinations of properties in the search and retrieval process. The weighting of individual images and the selection of a result set generally rely on the system's ability to measure the distance between various property values (e.g., between a *red square* and a *blue circle*). Two well-established commercial products are QBIC, which is used by the Hermitage collection, and Virage, which focuses on video retrieval and was used in Alta Vista Photofinder. QBIC computes global histograms for multiple color spaces, for pattern, for edge-based shapes (including user-generated sketches), and for associated text. Virage is based on both local and global histogram data that defines domain-specific, developer-created primitives and provides searchers with the ability either to sketch or to import images for query exemplars.

Query formulations in these CBIR systems are as varied as the system designs, but in first generation systems they all begin with the property color represented by a

---

[26] In the context of CBIR, the element of *form*, as a depiction of a three dimensional object, has been eliminated because the digital display is a two-dimensional medium and form uses shape and line to indicate the third dimension. In the CBIR context, therefore, *form* is subsumed by *shape* and *shape* will be used in this discussion to represent both 2D and 3D objects.

single-value histogram. This approach would work well, for example, if the search were on images from Picasso's Blue Period; it would not work well, however, if the search were for a work by Mondrian or Stella. Researchers appear to have made the basic assumption that the searcher will always seek an image that can be represented with a single overall color and that she is capable of describing that image by matching color need with a random image as an exemplar. Additional aspects of query formulation include: specifying relative importance (such as the *not*/*somewhat*/*very* matching criteria used in Blobworld or the *some*/*mostly* criteria used in Chabot); assigning color weights (e.g., 0-9 in SaFe) or percentages (e.g. QBIC); or selecting similarity calculation methods (e.g. *intensity*/*binary*/*gradient* in the Leiden Collection). Except when linked to keywords, all of these query types, as well as queries that involve select/provide/sketch a sample image (or part of an image), are variations on color values applied locally and/or globally.

As second generation systems have evolved, approaches to query formulation are more focused on local features. Recent versions of Blobworld (Carson, Belongie, Greenspan, & Malik, 1997, 2002) identify a small set of local pixel clusters in the QBPE image that have color and texture coherency, and allow the user to view the system's representation of the submitted image as a "blob" shape. A related approach is based on multiple objects: The searcher identifies a set of similar objects based on color-texture similarity and the system then "learns" the "shape," disregarding the color property, such as cars in multiple colors (see Rahmani, Goldman, Zhang, Krettek, & Fritts, 2005). Linking image-regions or objects to semantics is another area of rich research and there are several approaches. They include identifying text such as annotations, keywords, etc., associated with images (Fan, Gao, Luo, & Xu, 2004; Quack, Monich, Thiele, & Manjunath, 2004); selecting coherent image sets (landscapes, brain scans, fish) and training the system to match the term with similarity measures (Chen, Bart Jr., & Teng, 2005; Traina Jr., Figueiredo, & Traina, 2005); or asking searchers to provide annotations (e.g., selecting the faces in your family album and providing a name) (Zhang, et al. 2004).

Corridoni, Del Bimbo and Vicario (1998) apply an approach that maps the syntactic level, broadly defined as the numeric facts of the image, to the semantic level of color terms. The color terms are then mapped to a finite set of terms identified in Luescher's (1969) physio-psychologic research that relates colors to feelings (e.g., yellow through red are *warm* and yellow through green are *cool*). The method used by Corridoni et al. relies on the codification of numeric color representations according to the set of

180 color terms defined by Itten (1961) following his Bauhaus experience. The research of Berlin and Kay (1969) provides a set of 13 basic level color terms that is of interest to other researchers for mapping numeric color values to color semantics (Maillot, Thonnat, & Boucher, 2004; McDonald, Lai, & Tait, 2001; Mojsilovic *et al*., 2004; van den Broek, Kisters, & Vuurpijl, 2004). However, Maillot, Thonnat and Boucher and Mojsilovic et al. both, note the need for the added flexibility of modifiers (very dark, moderate light) and choose the 28 names for hues found in the National Bureau of Standards color dictionary (Kelly & Judd, 1976), which includes modifiers for lightness (very-dark, dark, medium, light, very-light), saturation (grayish, moderate, strong, vivid) and a combination of lightness and saturation (brilliant, pale, deep). Mojsilovic et al. are still concerned with "the lack of systematic syntax and the problem of several artificially constructed names (reddish-gray)" (2004, p. 92) and rename modifiers such as the "–ish" form for their color naming metric.

In order to use CBIR, users must adjust to its resource-centered rather than user-centered techniques because CBIR systems, like many systems in general, are "completely immune to the particular group of users and their queries" (Fidel, 1994, p. 575). Although there is the potential to link from low-level shapes to high-order semantic concepts, the ability of a searcher to judge overall image color or to describe a shape based on contour or color region has yet to be tested. Furthermore, the searcher is given no information about the algorithmic weighting that is the basis for relevance matching; and the results of a search are often displayed in random order or, if ordered, provide no information about the criteria used in presenting the results. As the most recent research trends begin to examine and build linkages between semantics and low-level features, the human-in-the-loop approach is receiving more attention, although often in the form of searcher profiling (Mountrakis, Stefanidis, Schlaisich, & Agouris, 2004; Yu et al., 2003) rather than interactive searcher feedback (Tsai, McGarry, & Tait, 2006).

In their critique of object recognition products, Forsyth et al. (1997) argue that these systems do not code "spatial organization in a way that supports object queries" (p. 122). Most products such as QBIC and Virage are based on exact modeling of object geometry as hierarchies of rectangles of internal color on a grid; however, application of this approach to model the side view of a horse would render the front view unrecognizable. The individual building a query statement must infer both the semantic and spatial relationships between color and content, such as blue on top (sky) and green on bottom (landscape) in a search for gardens. In the case of a color query for images containing a large amount of yellow, the result is "a collection of responses that is eclectic in content, (and) there is little connection between the response to this query and particular objects" (Forsyth *et al*., 1997, p. 125. See, for example, Figure 1.2.)

Forsyth et al. (1997) offer a promising approach to finding pictures of objects based on their spatial properties. Arguing that manual, text-based indexing is unrealistic for very large image databases, they propose a model that applies rules based on the defining criteria of an object domain. In their experiments, an object is modeled as a loosely coordinated collection of grouping and detection rules that reflect "both surface properties (pattern) and shape information" (p. 136). This approach works with objects whose boundaries are potentially variable (e.g., through rotation and sizing) since object identification is based on non-geometric cues such as color or the relationships between elements in the image. Thus the identification of specific object types and their domain-specific constraints is central. For example, to find images of naked people, the grouping rule "first locates images containing large areas of skin-colored regions; then within these areas, finds elongated regions and groups them into possible human limbs and connected groups of limbs" (Forsyth et al., 1997, p. 132). The object here is the human figure and the domain-specific constraints are skin color and the spatial configuration of arms, legs, torso, etc. Grouping rules are built from the local (selected skin color areas) to the global (size and shape of areas) and from the generic (limb shape) to the specific (limb and torso articulations). Although this approach might not be appropriate at the level of specific individuals, it should be effective at more general levels of identification and organization.

Rorvig (Rorvig, 1993, 1999a, 1999b; Goodrum *et al*., 2001) suggests a method for identifying isomorphic relationships between terms and image properties by statistically identifying similarities in vector clusters for both. Although his focus is on the selection of key frames of moving image documents from NASA that are to be used as abstracts in visual library systems, Rorvig's approach actually operates on the static frame. Extracted properties are hue, chroma, saturation, line frequency, line lengths, angle frequency, angles by degrees, and edge magnitude. In this approach, application of algorithmically-produced structural values introduces systematic error: for example, an analysis of light level focusing on high-contrast frames neglects all low-contrast frames. But Rorvig observes that such an omission can be explained to the searcher in a system applying this methodology, unlike a system that must rely on the individuality and inconsistency of human indexing decisions. Rorvig's initial ideas have been carried on in many new avenues of CBIR research such as similarity vectors and relevance ranking (Traina Jr. et al., 2005), combinations of algorithmic structure values (Li, Lee, & Adjeroh, 2005), and property extraction (Lakin, 1986; Mojsilovic *et al*., 2004).

With respect to CBIR technologies, it is reasonable to ask if there are shape boundaries beyond those tied closely to rudimentary pattern-matching or conventions of representation that could be employed in the query process. Twyman (1979) suggests that there are schemas of image elements – existing conventions of shapes and boundaries – that are important properties for domain objects and that these schemas could facilitate standardization in CBIR systems. The existence of such element schemas has been explored in CBIR mathematical research on spatial attributes (Cobb & Petry, 1998; Corridoni et al., 1997; Forsyth *et al*., 1997; Twyman, 1979) and through the development of taxonomies (Changizi, Zhang, Ye, & Shimojo, 2006; Cordier & Tijus, 2001; Lohse, Biolsi, Walker, & Reuter, 1991; Lohse, Walker, Biolsi, & Reuter, 1991; Twyman, 1979).

2.3 Visual language

> Spatial parsing is the first step in interpreting visual languages. It is the process of discovering the underlying syntactic structure of a visual communication object from its spatial arrangement so that it can be interpreted as a phrase in a particular visual language. Interpretation then produces a higher-level semantic representation. (Lakin, 1986, p. 36)

The term *visual language* can be defined in several ways. It can refer to a computer language used for programming or processing (S.-K. Chang, Ichikawa, & Ligomenides, 1986). It can refer to a language that points to the structure – the syntagmatic relationships – of the elements of an image: for example, the circle-square-triangle of the Bauhaus. It can also refer to a set of terms used to describe the visual characteristics (the paradigmatic properties and taxonomies) of an image; for example, the terms *line*, *color*, *pattern*, *graph* or *matrix* (Hortin, 1994). A visual language would have a lexicon of visual terms, a syntax or grammar used to recover syntactic structures and a semantics that establishes meaning under a system of interpretation.

In the category of computing languages, CBIR imposes logical – or binary – representation on visual objects and develops methods for processing the language of these representations.[27] Current progress in logical language representation is exemplified by the emerging MPEG-7 standard (Bober, 2001) that specifies a set of descriptors associated with content and is used to represent various types of multimedia information (Paek, Benitez, & Chang, 1999). There are four categories of visual descriptors used in the MPEG-7 standard: color, including descriptors such as *dominant* and *layout*; grid layout, including descriptors such as *view* and *coordinates*; texture; and shape, including descriptors such as *region* and *contour*. What is not provided by MPEG-7 is a standard vocabulary for referencing these descriptive categories in ways that are accessible to searchers: for example, the names of colors or the even more technical color-space designators such as *RGB*, *HSV*, and *monochrome*.

The logical approach to visual languages is based on the artificial reality of the mechanics of a computer system. These computer languages are dependent on the representation of structures that must be translated from the real world of the object to the binary world of the computer system. In information storage and retrieval systems, this process of translation is the traditional role of the indexer, who creates a text-based representation of a resource. With the development of image processing and computer vision capabilities, some aspects of the translation process can be handled automatically by the system itself: for example, the indexing of color or shape placement in an image frame. But, to be of relevance in the search process, any such automation must be informed by the relationship of the descriptive category to the real world.

---

[27] S-K. Chang et al. (1986) consider visual representation in the digital environment and provide for the separation of languages for visual and logical objects based on function. Research on visual objects in the digital environment can thus focus on: the iconic (ofness or aboutness); on the processing representations (binary processes); or, on the interrelationship of the two (the semantic gap between visual structure and meaning). All three approaches require a visual information language for the mechanical display of visual objects within logical representation systems (ones and zeros). S-K. Chang et al. identify four categories of visual languages based on the structures required by a computer system. These categories are first divided as logical versus visual objects and then described by their function: Logical objects with visual representations for 1) programming with visual expressions or 2) supporting visual interactions; and Visual objects with imposed logical representation for 3) presenting iconic visual information or 4) processing visual information languages. The first two categories have applications in image processing, computer vision, robotics, image database management, image communications, etc. The third category has applications in computer graphics, interface design, form management, and computer-aided design. HTML fits this category as a primitive text-based visual information language that attempts to control the size and position of text and images on the digital display. The SGML family represents syntax, which may only be a metaphor in images. Worth (1981) suggests that only individual artists develop individual syntax, e.g., the Diebenkorn grammar developed by Kirsch and Kirsch (1985).

The definition of a visual language as a set of terms that describes the visual, syntagmatic characteristics of an image reflects Howard's (1980) contention that lexical analysis be applied to a set of visual symbols. The lexical analysis of an image would include the identification of the perceptual parts – the visual elements – that participate in the formation of meaning reflected in the structure of the whole. In the design of visual stimuli, the Bauhaus movement tied functional analysis to the gestalt perceptual principles of form, not in the linear relationship indicated by the popular conception that form follows function, but in a relationship of codependency (Goodman, 1968).

Green (1995b) identifies two basic relationships of linguistic codependency: paradigmatic and syntagmatic. Paradigmatic relationships are based on category patterns in the language system: taxonomy (kind of), antonymy (opposite of), and partonomy (part of). These patterns are linked semantically through lexical relationships, such as: dog-canine (synonyms), hot-cold (antonyms), or nose-face (partonyms) (Green, 1995a, p. 317). When applied to visual information or graphic language resources, paradigmatic relationships are demonstrated by the relationship of the pixel to a pattern: for example, every instance of the letter *A* shares a similar configuration of pixels regardless of graphic style or font, thus exemplifying the most basic application of the gestalt principles of perception as applied to visual organization: **A A A**.

In contrast, syntagmatic relations hold "between the linguistic elements of the constructed form" (Green, 1995b, p. 366). In an English sentence, these relations are linear and reflect syntactical structures, as in knowing that adjectives generally precede nouns. In images, however, syntagmatic relationships may or may not be linear. For example, when a circle, triangle, and square are combined to form an icon, the three shapes can each be articulated structurally, through geometric formulations, or they can be part of a single, meaningful unit, such as the Bauhaus icon, which operates at a higher functional level in that the whole (gestalt) is greater than the sum of its parts. A more complex example is the mechanical definition of "naked human": finger connected to hand connected to forearm connected to torso, etc. (Forsyth et al., 1997), is a statement of the syntagmatic relationships that exist among the various components of the human body. Although this mechanical definition is computationally linear, a naked human is understood as a gestalt. Other demonstrations of paradigmatic and syntagmatic visual relationships can be identified in the works of the surrealists: Tanguy is a biomorphic surrealist who paints unrecognizable objects (paradigmatic confusion) in recognizable or syntagmatically-correct situations; Magritte is an anthropomorphic surrealist who paints paradigmatically-correct, recognizable objects in unrecognizable situations (syntagmatic confusion)[28].

Green and Bean (1995) argue that, in texts, simple concepts combine to form more complex concepts and that, when they do, "they operate as gestalts, that is, as integrated conceptual units" (p. 660). They explore the application and extension of conceptual syntagmatic relationships to topical information systems and identify four major relationships at work in information retrieval systems: syntagmatic, paradigmatic, hierarchical, and matching. They note that, in text-based resources, these relationships are distinguishable only on the basis of function, since structurally they are all based on combinations of alphanumeric characters. Syntagmatic relationships are most commonly applied to the phonology of language: the formation of word sounds and the combination of words into sentences, then paragraphs, etc. In images, these syntagmatic relationships are expressed by the Bauhaus in the function of the material and the gestalt principles of visual organization. Paradigmatic relationships are options available within a structural syntax (Morgan & Welton, 1992, p. 43): for example, the sets of terms that are identified in the facets of a classification. When creating an image, these options are the low-level elements, such as lines, circles, or dots (Van Sommers, 1984) that can be combined to form higher-order elements, such as specific patterns or objects.

When applying syntagmatic relationships to text, Green (1995b) explains how they can make retrieval more precise through the identification of specific relationships. For example, "libraries IN schools" is more precise than "libraries AND schools." Current application of syntagmatic relationships assumes that there is little overlap with the semantics of paradigms. This is demonstrated in information systems by the small set of Boolean – or syntagmatic – operators and the importance of paradigmatic relationships in indexing: for example, the generic relationships broader term (BT) and narrower term (NT). Although Green classifies syntagmatic relationships as syntactic phenomena, she stresses that this does not exclude them from being meaningful, as in the library example above.

Green's (1995b) argument that application of syntagmatic relationships in indexing languages can improve precision and recall in information retrieval systems points to the potential relevance of semantics in the mechanical analysis of syntactical relationships in the digital environment (Florack, 1997). At the foundation of Green's discussion is the assumption that it is relationships, such as those of space, time, or

---

28 See The Museum of Modern Art (www.moma.org) for examples: Yves Tanguy, "The Great Mutation," 1942; and Rene Magritte, "The Portrait," 1935.

authority, that govern the individual's understanding of the world and her activities within it (see Green, 1995b, p. 366, and; Mitchell, 1980, p. 276). In similar fashion, the semantic content of an image is the product of the relationships among visual elements in a two-dimensional world. The potential for syntagmatic linguistic relationships to constrain the result of a text-based information search points to the possibility that structural visual characteristics can be applied to the retrieval of images.

Green questions the notion that linearity is necessarily required in syntagmatic relations and suggests that the elements in some relationships may be perceived simultaneously. She argues that, in the English phrase "red rose" and the French "rose rouge," the color and the object are presented linearly but are not processed separately (Green, 1995b, p. 369). Similarly, the rapid sequencing of eye movements causes the individual to perceive the significant pattern in an image rather than its separate elements. In addition, there are complex situations where one activity may involve several entities and she suggests that this holds for the apprehension of visual scenes.

Paradigmatic and syntagmatic relationships are intertwined in images, as in the syntagmatic configuration of pixels that constitutes the letter *A* and the paradigmatic need for a visual expression of the concept *A*. Similarly, when images are formed, the higher-level perception is greater than the sum of the lower-level syntagmatic and paradigmatic parts, such that the semantic meaning is inherently tied to the structure. The entity, as a whole, becomes an independent and meaningful unit in itself, as in each frame of a comic book or the specific configuration of the Bauhaus icon. Syntagmatic relationships, whether applied to text-based or graphic language resources, demand methods of analysis that go beyond term co-occurrence – or pattern matching – into the identification of relationships and roles as well as the current focus on the naming of participant objects (Green, 1995b).

2.4 Image properties

Berenstien (1996, p. 24) describes image properties as both visual and non-visual. Non-visual properties refer to information that is not contained in the image itself but provides administrative or biographical information such as medium, access rights or file format. Visual properties include elements such as composition (overall structure, balance and placement of elements, or spatial relationships among elements); proportion (relative size of objects in relation to each other and to real world objects); contour; mass (weight and volume); color; contrast (degree of difference between colors); pattern; and orientation (landscape or portrait). In his synthesis of the theoretical foundations of visual learning, Hortin (1994) identifies the primary properties of a visual language as line, shape, form, size, color, and motion.

Bertin (1980; 1983/1967) has focused on identification of visual properties that characterize diagrams, networks and maps, also known as data-graphics, info-graphics or quantitative displays. He identifies eight properties of these non-moving graphic representational systems: the two dimensions of the plane (length and elevation) as well as size, value, texture, color, orientation and shape. In contrast, Caivano (1990; 1994) works with the single property of pattern, which he refers to as texture. He describes a system for ordering patterns based on quantitative properties such as size, saturation, density, and directional organization, which, he argues, can be used to develop systematic and algorithmic descriptions of patterns (see Figure 2.4).

In his matrix theory, Twyman (1979) categorizes all of these basic properties as *image encoding* options (see Figure 2.8). He seeks to isolate the image encoding options used in the design process and to build an analysis of their application. In art theory, this approach is identified with formalism and is considered structuralist, or process-based, since it considers only the visual elements and properties of an image and doesn't address meaning or interpretation.

Van Sommers (1984) notes that the deconstruction of images received widespread attention with the "picture grammars" of the 1960s, perhaps due to the developing capabilities of the digital environment. Application of the picture-grammar approach attempted to describe images as feature hierarchies and to identify universal principles (Machines, 1969). Rather than understanding "grammar" as a metaphor, this approach emphasized the analogical similarity between language and image design.

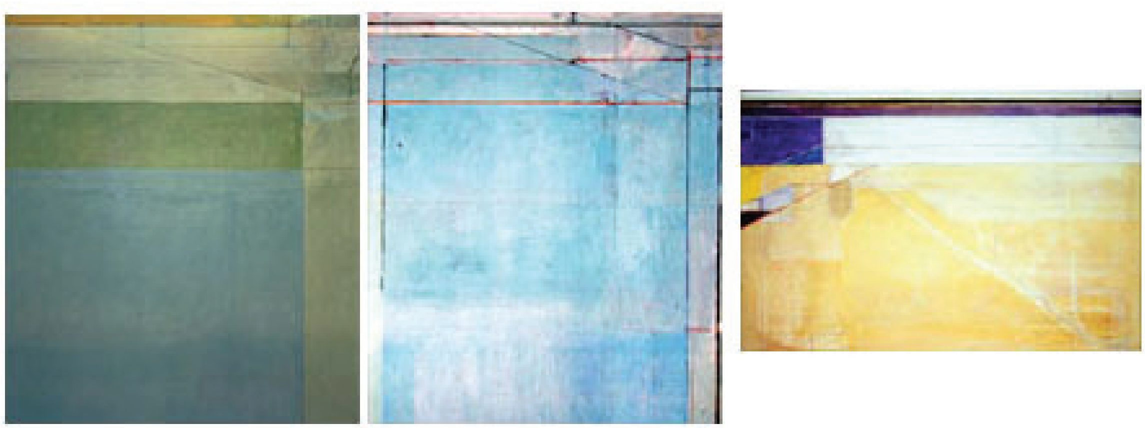


Figure 2.5: Examples of the Ocean Park series of paintings by Richard Diebenkorn. The Ocean Park series of paintings was deconstructed by Kirsch and Kirsch (1985) to create a structural grammar such that

a computer, given the Diebenkorn grammar rules, could generate pseudo-Diebenkorn images. R. Kirsch (1985) notes that "when Richard Diebenkorn saw this randomly generated picture he had 'the immediate shock of recognition', notwithstanding the unusual history of its construction" (p. 66). Oil on canvas paintings, left to right: Ocean Park #60 (1973), Ocean Park #66 (1973), and Ocean Park #132 (1985).

In contrast, the research of Kirsch and Kirsch (1985) identifies primitive image elements specific to the painter Richard Diebenkorn and codifies their compositional interrelationships. They use the concept of grammar for both image generation and deconstruction, applying the notion of syntactic structure to analysis of a static image (frame) in order to generate an algorithmic definition. In their theoretical explorations, they first deconstructed the Ocean Park series of Diebenkorn (see Figure 2.5), then wrote a computer program that would generate new images in the style of Diebenkorn based on the codified stylistic grammar. This work was pioneering in its extension of the elements of formal grammar employed in computer applications – elements such as parsing and language development – to the image domain through the use of pattern analysis. The fact that the Diebenkorn grammar could generate images of the same genre introduced the possibility that such grammars could be identified for other visual genres.

In visual arts education, the Agam Program teaches the child to analyze her environment into simple elements and to develop the "ability to choose the right level of detail in visual perception" (Razel & Vat-Sheva, 1986).[29] The program is based on a visual alphabet and 36 concepts of visual representation that include typical forms and variations, proportions, colors, dimensions and sizes, angles, trajectory (horizontal, vertical and oblique), symmetry, curved lines, and composition. The main object of the program is to develop the individual's visual language in order to comprehend the universe of forms, to decipher the basic elements of this universe, and to synthesize complex visual forms from simple basic elements. "An important component of this skill is the ability to see what things really look like … rather than to depend on one's predetermined verbal concepts" (Razel & Vat-Sheva, 1986, p.50), suggesting that the design of CBIR interfaces needs to guide the searcher to more basic element description so as to narrow the semantic gap.

Changizi, et al. (2006) suggest that the underlying configurations of human visual

[29] Levels of detail in visual perception can also be referred to as levels of *resolution* for both representation and interpretation, e.g., the perception of pixel groups resolves as the interpretation of letters, letters as words, etc. The level of resolution impacts interpretation.

signs take their form from familiar scenes that humans are naturally good at processing. Comparing Chinese symbols[30], non-linguistic symbols, and over 100 writing systems, they identified a set of 36 configuration types based on articulations of three or less segments (e.g., **L, T, X).** Calculating the frequency distributions of the configurations, they found significant similarity among the signs in the visual systems. To test the usage of these configurations in natural scenes, they selected images from three domains: photographs of ancestral people, photographs of rural scenes from National Geographic, and computer-generated architectural scenes. Again they found configuration distributions to be invariant across the diverse environments, suggesting underlying principles of shape. They emphasize, however, that topological shape (see Figure 2.1) takes priority over geometric shape since it accommodates paradigmatic variation and object junctions as an intermediate stage in recognition (Biederman, 1987). Indeed, the feature extraction metrics of Iqbal and Aggarwal (2002) for architectural objects successfully used the identification of **L** junctions and **U** junctions rather than image segmentation or object geometry. This supports CBIR metrics focusing on identification of feature thresholds rather than specific geometric definition and indicates the potential for identifying small sets of configuration types.

In an attempt to explain how humans mentally make sense of the images they see, Saint-Martin (Saint-Martin, 1990; Lester, 1995) has constructed a visual alphabet that consists of the graphic elements that make up a picture. For her, just as in CBIR, color is the basic element that gives shape and substance to the visible world. She calls her basic visual component the *coloreme*, which she defines as the smallest element within a direct or mediated image that can be "first differentiated in the perceptual process of vision" (Saint-Martin, 1990, p. 65). Lester notes in his analysis of Saint-Martin that "These physical (properties) of the image once noticed and identified find meaning through successive viewings" (Lester, 1995, p. 68). Saint-Martin lists five properties that accommodate the terms for properties identified by other researchers:

- *Color*: Value and contrast; hue and saturation.
- *Texture*: Pattern or surface.
- *Size*: Length, elevation, proportion, relation to others and to real world objects.
- *Boundaries*: Shape, including form, contour, mass and volume (e.g., block, cylinder, funnel), and line (e.g., angles, curves).
- *Direction or position in the frame*: Composition, including

[30] CBIR research is beginning to take note of the similarity in written language and visual language, exemplified most directly by Chinese characters for mountain and water evolving directly from the natural images (Wu, Xu, Yang, & Zheng, 2005a).

structural properties, whole structure, spatial relations, placement, balance, density, trajectory (e.g., horizontal, vertical, oblique), orientation (e.g., landscape or portrait), and motion.

When analyzing the language of physical structure in images, these five categories provide a way to organize the visual properties discussed by the other researchers.

In research on the image similarity judgments of children, Harrison (1992) notes that color is the primary grouping cue for young children but "developmentally color loses influence and is replaced by shape as the primary classification cue, with composition gradually becoming more important by the fifth and sixth grade" (p. 46). By college level, all properties have relative equality, with shape remaining slightly more influential. Hughley (1989) reviews a body of children's art education literature and highlights "the influence of structural properties on preference responses to paintings" (p. 74). Expanding on Harrison, he notes that older children generally make similarity judgments based on composition, style, and formal elements, while younger children rely on subject matter and color. The findings of other research studies (Mervis & Greco, 1984; Tversky & Hemenway, 1984) indicate that part-whole categorization differences are influenced by properties of the object, such as color, composition, degree of realism, and subject matter, as well as by viewing time.

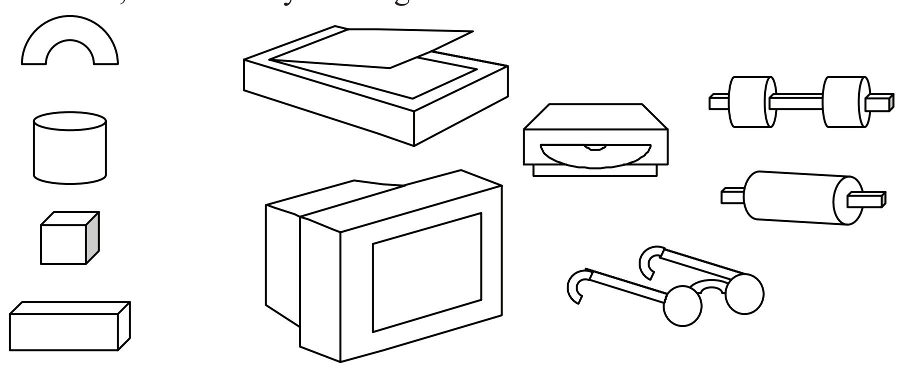


Figure 2.6. Examples of Biederman's (1987) geons. Geons, on the left, are used to construct objects, on the right. (Example created from similar figures by Biederman.)

Recognizing that every object consists of primitive shapes or parts, Biederman (1972; 1987) proposes a theory of visual perception called recognition-by-components (RBC) that is based on an alphabet of 36 "geometrical ions" or "geons" which he equates to the 55 phonemes used to construct verbal language. Examples of geons are volumetric terms such as block, cylinder, and funnel (see Figure 2.6). His experimental research supported the hypothesis that as few as three geons[31] are all that are required to uniquely specify an object and provide rapid recognition. In selecting the three most appropriate geons, texture, color, brightness, and contour are not required as the first three can vary without effecting the visual structure of an object and the fourth is embedded in the geons themselves. Biederman notes that the RBC theory of perception is supported by Tversky's (1984) model in which partonomies serve to separate entities into their structural components and to organize knowledge of function by the components of structure. Biederman concludes that "perceived part configuration, then, underlies *both* [italics added] perceived structure and perceived function" (1972, p.190). Similar objects, he notes, have a high degree of component overlap; and the representation of an object is "a structural description that [is] expressed [by] the relations among the component parts" such as their relative sizes, orientations and angles of attachment (1987, p. 118-119). RBC theory supports the underlying assumption of CBIR researchers that the perception of patterns in visual stimuli takes two steps: first, the recognition and classification of image elements; and second, the identification of patterns based on relationships among elements (Smoliar, Baker, Nakayama, & Wilcox, 1997).

Biederman's (1987) geons can be considered in concert with Saint Martin's coloremes (Saint-Martin, 1990). Coloremes are readily matched to the low-level aggregates of pixels in CBIR valuation, compared to the higher-level volumetric names of the geons. Agglomeration of coloremes into higher-level operations is governed by the syntactic rules of topological relationships (which are similar to geon partonomies); gestalt relationships (which describe their organizational structure); and the laws of color interaction (which provide visibility due to color position or juxtaposition). Topological relationships and spatial parsing can be used in CBIR as one means for generating similarity measures and extracting object features (Di Sciascio, Donini, & Mongiello, 2002; Wu *et al*., 2005a; Wu, Xu, Yang, & Zheng, 2005b; Xing, Liu, & Yuan, 2005).

Worth (1981) suggests that, instead of a formal syntax and lexicon, as in verbal grammars, images have form, structure, rules, and conventions of relationship and role. Specific structures become conventions when they are widely understood by decoders of a visual communication. Comic books use well-understood examples of conventions in visual communication, where conversation is text in balloons: symbols

[31] This is the same number of symbol segments used by Changizi *et al*. (2006). However, Changizi et al. used three segments based on the need to reduce complexity of computation and analysis.

and letter combinations represent sounds (e.g., ARRGGGHH!! and @#%); facial lines express emotions like pain or anger; and motion lines indicate action (Kindborg & Kollerbaur, 1987; McCloud, 1993). In graphic design, there are conventions that grow out of perception, such as the use of color as an attention getting device as opposed to color as a simple attribute of an object. And Panofsky (1972, p. 4) introduces connotation as the significance of a convention: for example, the act of lifting a hat in greeting is grounded in the customs of a cultural tradition. Thus, there is a relationship between the perception of a visual structure and the appropriate convention that facilitates image understanding or interpretation.

2.5 Indexing Languages

The identification of parts is essential to the development of concepts. According to gestalt theory, however, meaning is greater than the sum of the parts from which a concept is constructed. In semantics, Tversky and Hemenway (1984) propose that the perception of an object's parts underlies the perception of both structure (appearance) and function in basic level categories. The theory of basic level category terms asserts that "categories at one level of specificity in a taxonomy are psychologically and linguistically more primary than more general and more specific categories" (1984, p. 170). Tversky and Hemenway conclude that "parts are a better index of basicness than are other, purely functional or perceptual attributes" (1984, p.182):

> Basic level categories are most informative because, given our perceptual apparatus and the structure of the world, this is the level at which the natural correlations and discontinuities among features are most salient. Basic level categories are the most general categories having members with similar and recognizable shapes.[32] (Tversky & Hemenway, 1984, p. 170 -1).

Following experiments where subjects listed object attributes and then identified those attributes that were parts of the objects, Tversky and Hemenway suggested that knowledge about parts "underlies the superior informativeness of the basic level" (1984, p. 171) of object terminology and later demonstrated that naming body parts activates both perceptual and functional information (Morrison & Tversky, 2005). Like RBC

[32] Bird (not mammal or robin) is an example of a basic level category, a term that provides the most information for the least cognitive effort. Tversky and Hemenway have borrowed this from Rosch (1976).

theory, the concept of partonomy has potential application for CBIR development in that it references shapes that may hold across categories of images.

Until recently, developers of electronic information systems, including computer scientists and Web implementers, had argued that "the best way to search for information on the web was by using keyword searching" (Scheiderman, 1996), a purely alphanumeric, string-matching approach that relies on a computer-generated index. However, developers are beginning to recognize, just as Soergel (1985) demonstrated, that this method is inadequate and frustrating, particularly when the searcher must sort through thousands or hundreds of thousands of hits with limited means of narrowing the original query (Steinberg, 1996). The alternative to natural language searching is the use of a controlled vocabulary or indexing language.

Quinn points out that "the real purpose of controlled vocabulary is to enhance representational predictability" (1994, p. 140). This predictability is centered in the principle of one-word/one-concept – and its corollary principle of one-concept/one-word – which establishes conceptual relationships of equivalence or synonymy (Shank & Kass, 1988). For example, if the natural language terms *ship* and *boat* are used in the same resource collection, the controlled vocabulary might identify *ship* as the authorized descriptor or preferred term used to index resources, with boat as a non-preferred term pointing to ship. In a traditional classification scheme, each class exists as an authorized descriptor and the set of descriptors constitutes the indexing language of the classification structure. A classification system is thus the extreme example of a controlled vocabulary in that it requires the assignment of a data unit to a single class. In a categorization system, however, a resource can have more than one authorized descriptor assigned to it (Jacob, 1992). For example, in the Dewey Decimal Classification (DDC), a book is assigned to one and only one class, which is represented by a unique class number. Using the Library of Congress Subject Headings (LCSH), however, that same book may be represented by multiple categories, each of which is represented by a unique subject heading as descriptor.

Controlled vocabularies are most often developed for use in a specific domain. Thus the catalogues of art history image collections generally focus on such biographical attributes as creator, title, medium and date of creation. Abstract images and their often obtuse titles, such as those of Stella, Krasner, Pollack, de Kooning, Diebenkorn, Mondrian, Davis, Still, or Arp[33], are impossible to locate without knowledge of this

---

[33] For example: Frank Stella, "Flin Flon," 1968; Lee Krasner, "Composition," 1943; Willem de Kooning, "Composition," 1955; Jackson Pollack, "Autumn Rhythm (Number 30)," 1950; Richard Diebenkorn,

biographical data. The controlled vocabulary of the *Art and Architecture Thesaurus* (AAT) (Petersen, 1994) was originally developed to support the specific representational needs of art and architecture but is currently used in many other domains. For example, archaeology often uses the AAT even though it does not offer descriptive terms that are appropriate in all areas: for example, archaeologists may need to represent *horse bones*, but the AAT uses *horse* to refer to a sawhorse.

Faceted classification (Priss & Jacob, 1999; Yang, Jacob, Loehrlein, Lee, & Yu, 2004) is able to represent different views of data through the combination of facet isolates. Faceted classification "is based on small conceptual groups rather than lists … and has been used in the design of classification schemes, vocabularies, and information systems" (Iyer & Giguere, 1995, p. 89). As such, faceted classification is similar to the data modeling process in relational database management systems. Although there may be bias in the initial selection of isolates – or property values – in the faceted scheme, the primary source of bias is that imposed on the query by the searcher through the use of relational operators (Pollitt, 1999). A faceted design is therefore adaptable to the needs of multiple users as well as multiple resource types, unlike many traditional representational systems that are limited by the structure of the hierarchy (Hammond, 1993).

Unfortunately, practical and theoretical research dealing with the representation of non-text materials has concentrated on text-based approaches and ranges from the development of specific tools (e.g., AAT) to discussions of appropriate approaches to subject representation (Krause, 1988; Layne, 1994): For example, a painting *by* Raphael (physical representation) might be a picture *of* a woman and angel (content representation) that is *about* the Annunciation (concept representation) that is *similar-to* Giotto's Annunciation (contextual representation). Problems in adapting text-based representation to image collections are well documented (Austin, 1994; Jorgensen, 1996a; Krause, 1988; Layne, 1994; Lunin, 1994; Rasmussen, 1997; Small, 1991; Turner, 1993). In addition, because most of the images available digitally are *not* works of art, approaches to indexing methods used in art history are frequently neither appropriate nor adequate: Image seekers may have little or no interest in image creator or title, but will search, instead, on color, subject or visual constructs such as the silhouette of a boy jumping (Collins, 1998; Enser & McGregor, 1993; Jorgensen & Jorgensen, 2005). But, in point of fact, many digital images are not described at this level of detail (Stephenson, 1999).

Austin (1994) addresses the lack of adequate cataloging of images and, in order

---

"Untitled," 1970; Stuart Davis, "Blips and Ifs," 1963; Clifford Still, "Number 2," 1949; Piet Mondrian, "Tableau II," 1921; or, Jean Arp, "Objects Arranged According to the Laws of Chance or Navels," 1930.

to establish inventory control, proposes a minimum level of cataloging with nine access fields. Seven of these fields represent physical aspects of the image, including, for example, title, creator, date of production and current location. To these he adds an eighth field for conceptual access to the subject content and a ninth and final field for general description of the image. Austin advocates initial building of image databases that incorporate minimum-level physical detail as well as preliminary categorization, or naming, of image content with the representation of content interpretation to be added as standards for controlled vocabularies are developed in specific topical areas.

The foundation of subject indexing for images is frequently identified with Panofsky's (1972) three steps in the process of image analysis: pre-iconographic, or description of the image; iconographic, or analysis of what an image is about; and iconologic, or interpretation of the meaning of an image. Many authors (e.g., Enser & McGregor, 1993; Focillon, 1948; Shatford, 1986; Sonesson, 1988, 1989, 1995; Stenvert, 1992; Turner, 1993; Veltrusky, 1973) focus on these three levels of analysis as the foundation for image description.

Relying on Panofsky's first two levels of analysis – the pre-iconographic and iconographic – Layne (Shatford, 1986) develops a theoretical approach that distinguishes between ofness and aboutness in the subject representation of images. Layne points out that both of Panofsky's first two categories include qualities of ofness and aboutness:

> At the pre-iconographic level, the *Of* aspect is generic description of objects and events; at the iconographic level, it is a specific, or proper, appellation of those objects and events. *Of* words describe people, places, objects, conditions, and actions that have a physical manifestation. The *About* aspect is, at the pre-iconographic level, a description of the mood of a picture: at the iconographic level the *About* aspect is an identification of mythical beings that have no unique and verifiable concrete reality, of symbolic meanings and abstract concepts that are communicated by images in the picture. *About* words include those describing emotions (love, sorrow) and concepts (truth, honor). (Shatford, 1986, p. 45)

Although Layne acknowledges that images such as abstract art "may well have *About* qualities" (Shatford, 1986, p.45), she only addresses pictures, or representational images, presumably referring to aspects of mood and emotion. Interestingly enough, however, is the fact that, while abstract art does have ofness qualities, controlled vocabularies designed to represent works of art do not accommodate representation of the basic physical aspects of an image such as those privileged by the Bauhaus (e.g., circle, square, triangle).

Layne (Shatford, 1986) suggests that a faceted approach to representation of ofness and aboutness will address the questions Who?, What?, Where? and When? These aspects are applicable to both specific elements and the overall composition or set of internal relationships between elements in an image. Layne (1994; Shatford, 1986) proposes four categories of properties[34] that participate in the indexing of images: biographical properties, which describe the image's provenance; ofness and aboutness subject properties, which represent the intellectual content of an image; exemplified properties, which refer to type (e.g., portrait) and medium (e.g. oil on canvas); and relationship properties, which link an image to other images, textual works, or even objects. Applying Jacob's levels of document representation,[35] Layne's biographical and exemplified properties are aspects of physical representation; subject properties are equivalent to conceptual representations; and relationship properties are aspects of contextual representation.

Although Layne does not go so far as to suggest that relationships between objects within an image may affect the aboutness, she does suggest that it may well be important to identify relationships as well as to indicate their nature. Both Layne (Shatford, 1986) and Turner (1993) question the feasibility of iconologic indexing – indexing of aboutness – as potentially too specific to the intrinsic meaning of the work and the subjective response of the observer, reinforcing the notion that representation of the aboutness of an image is not purely conceptual, but overlaps with the contextual. The implication, then, is that interpretation occurs at the intersection between the levels of conceptual and contextual representations. Thus the first step in representation must involve analysis of the ofness elements of an image within the context established by the internal structure of the image itself, that is, examination of internal contextuality – the internal visual structure.

Relationships between elements are highlighted in Jorgensen's (1995a; 1995b; 1996a; 1996b; 1997a; 1997b; 1998) research in which subjects were asked to verbally describe a set of images. Her analysis of subject responses identified twelve categories of representation that were most typically applied in subject descriptions. Ranked by frequency of subject use, the category *location* was the

---

[34] Layne uses the term *attribute* to refer to a representational property of an image. The term *property* is used here in accordance with its definition in Chapter 1.

[35] See discussion in Chapter 1, and Figure 1.1.

fifth most typically applied by subjects. Jorgensen defines the notion of location as including “attributes relating to both general and specific locations of picture components” (1997a, p. 210), where general locations refer to the geographic place of the image while specific locations referred to the internal orgnization of objects within the image frame, such as *behind*, *next to*, *on*, or *against*.

Jorgensen (1996a) then compared this system of twelve categories to three established indexing systems: Library of Congress Thesaurus for Graphic Materials (LCTGM), a list of topical terms for subject access to images; Thesaurus Iconographique (TI), a descriptive system for conceptual representation of images; and Dewey Decimal Classification (DDC), a general subject classification scheme. To determine to what extent the terms in each of the three systems represented the categories generated in her research, Jorgensen took a random sample of terms from each system and coded them into her twelve categories. From the numerical distribution of terms to categories in each system, she concluded that no current system was comprehensive as a set of descriptors for a generic image collection. More importantly, she reports that LCGTM, TI, and DDC do not contain any terms for specification of element location in the internal structure of an image. She argues that this implies the need for “a different mechanism … to adequately represent the conceptual basis of location terms” (1996a, p. 194).

Ohlgren (1980) points to the lack of terminology standardization in iconography; but Turner (1993) questions whether a single standard is realistic. Both of them suggest that a general use vocabulary might have to include all potentially visible objects and that this could cause confusion when applied in different conceptual domains: For example, the AAT term *horse* that refers to a carpenter’s tool (i.e., a *sawhorse*) rather than the animal.

Application of the faceted approach in the construction of a controlled vocabulary has significant implications for the development of an image retrieval system that could integrate current approaches to verbal retrieval techniques with CBIR’s ongoing work with retrieval based on the internal structure of images. It is important to emphasize that the internal organization of the image is not the same as the physical description of the item, but considers the internal context of elements in the image. Thus, while an indexing language that combines natural language searching of names and titles with a controlled vocabulary is generally considered the most effective approach to retrieval (see, for example, Lancaster, 1998), there are strong indications that a vocabulary should be developed to represent the internal visual structure of an image as well as its non-visual physical properties.

2.6 Typologies

Smith and Warner (1984) observe that, "In developing a taxonomy, one must be concerned with both classification and nomenclature, where classification is the process of establishing and defining systematic groupings and nomenclature is the allocation of names to the groups so produced" (p. 123). While some would argue that taxonomies indicate mutually exclusive classes and that categories – or typologies – are more flexible structures, allowing overlap or choice in data assignment (Jacob, 1992), both approaches are tools for making generalizations about the data under consideration. In his exploration of the logical method of analysis, exemplified by semiosis, Howard (1980) notes that the result of such an approach is the proliferation of taxonomies and terminologies in the social sciences. Looking at research in psychology and noting the strong link between property and category, Cordier and Tijus (2001) propose a typology of object properties: procedural (actions), structural (part-whole), perceptual (surface or visible), functional (usage) and behavioral (related to living things). These provide a potential scheme for generalizing about property relationships as well as relationships between elements. Yet, when considering a potential vocabulary descriptive of image structure, a problem of categorical confusion is evident in this typology: an action such as cutting (procedural) would have to be visible as a surface characteristic (perceptual).

There is a small group of researchers who have attempted categorizations of various aspects of graphic language resources. For example, Rankin (1990) is concerned with the categorization of graph types, but considers only those graph forms which use an *xyz* coordinate space that is both explicitly stated and numerically scaled – those graphs which have a major use in scientific, engineering, medical, and commercial fields of communication (see Figure 2.7).

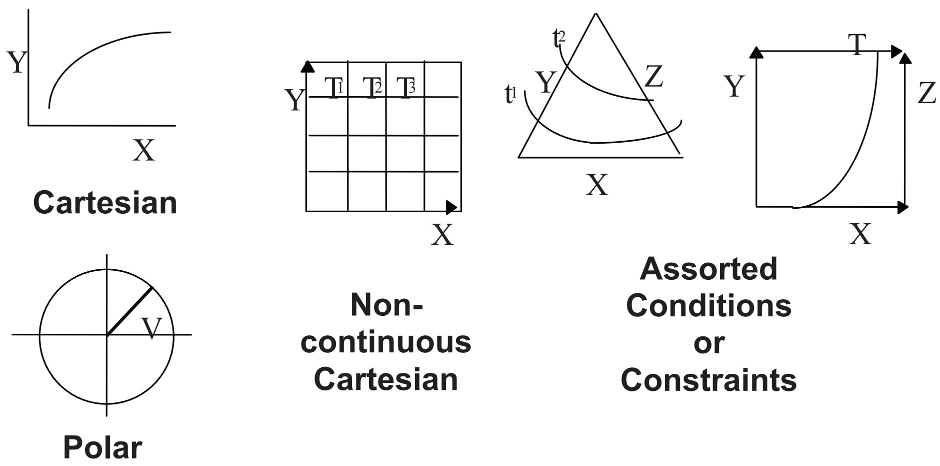


Figure 2.7. Four types of graphs based on variables. Rankin (1990) identifies four types of graphs based on the variables (or "data arrays") used to produce graphic illustrations, many of which "have received common names which are accepted throughout the scientific community" (p. 149). Common names for Cartesian arrays include line, trend, topographic, surface, conversion, and cylinder-like. Common names for non-continuous Cartesian functions include bar, clock, and matrix. Each of the four types then has subcategories based on the number of variables or dimensions displayed, for example, y=f(x) or z=f(x,y). In the examples illustrated above, Cartesian and Polar are two-dimensional while Non-continuous Cartesian and Assorted Conditions show three or more dimensions. (Redrawn from selected exemplars in Rankin, 1990)

In contrast, Goldsmith and Davenport (1990) explore structural composition[36] in their formal, mathematical definitions of graph similarity measures. Having defined the structure of a graph as the "organization of an object's constituent parts as viewed from the perspective of the whole object" (1990, p. 76), they measure the similarity between two graphs with common *xy* sets based on the structure of edge patterns. As in CBIR, similarity measures are then calculated based on the location of nodes or *xy* pairs, which are in either a relationship path or a relationship neighborhood.

The research of Lohse, Walker, Biolsi and Reuter (1991; Lohse, Biolsi, Walker & Reuter, 1994) investigates "informational graphics" – or directly conveyed "data structures for expressing knowledge" (1991, p. 37) – and analysis of the meaningful categories into which subjects grouped these images. Because Lohse et al. focus on the problem-solving nature of schematics, they differentiate between functional and

[36] The concept of composition structure in art subsumes the concept of spatial organization.

structural image types: structural images are "well learned and are derived from exemplar learning" (1991, p. 36) while functional images are characterized by the intended use. In their research, Lohse et al. dismiss the possible use of structure in the categorization task and focus solely on developing a problem-solving typology.

Lohse, Bilosi, et al. (1994) identify ten categories of images: graphs, network charts, process diagrams, tables, time charts, cartograms, structure diagrams, maps, icons and photographs. Unfortunately, however, the use of visual exemplars in this research confounds isolation of the information criteria used by subjects to identify these categories. Arrangement of elements and their relationships within the image emerge as syntagmatic, category-determining criteria and the category names themselves potentially reference various types of message construction. Thus, for example, graphs rely on object position and magnitude; network charts and process diagrams show component spatial relationships; tables are dependent on physical arrangement; and time charts and cartograms depend upon visual conventions.

Twyman, (1979, 1985) categorizes images by physical structure before considering meaning or intended use. He proposes a matrix that groups graphic languages, or visible communication forms, in terms of two properties: method of configuration and mode of symbolization (see Figure 2.8). Method of configuration is the graphic organization or physical structure of the image. His selection of symbolization and configuration as the basis of his matrix focuses on the relationship between these two approaches rather than the approaches themselves. Mode of symbolization refers to the representational form – the properties and elements – of the image, which he separates into three categories: verbal/numeric (i.e., numbers and other conventional characters such as typographic letters); pictorial; and the combination of verbal/numeric and pictorial. He further distinguishes the pictorial subcategory of schematics, which are "all purposeful graphic marks that are not words, numbers, or pictures" (Twyman, 1985, p. 247). This, Twyman admits, is a fuzzy division with the potential for significant overlap between the purely pictorial and the schematic, indicating that Twyman's matrix moves from a purely structural view of images to a more purposeful or use-oriented approach as suggested by Non-linear Directed Viewing (see Table 2.1).

Table 2.1

Categories of graphic language documents according to Twyman (1979).

	Verbal and/or numerical	**Pictorial & verbal and/or numerical**	**Pictorial**	**Schematic**
Pure linear	Minoan Phaistos disc	Bayeaux Tapestry, scrolls of continuous pictures	Trajan's Column, panoramas	route map, spectogram
Linear interrupted	text with word spacing and line breaks	comic strip	wall paintings of multiple scenes, Mayan codices	music notation
List	text in discrete semantic units (menu, order of service)	keys to maps	highway sign symbol list	(none)
Linear branching	family tree	tree diagram with pictures	pictorial tree structure	tree structure schematic of relationships
Matrix	tables	table with pictures	pictorial table	bar charts and line graphs
Non-linear directed viewing	advertising text	labeling, picture captions in page layout	non-random pictures (most common)	Network diagrams, subway maps
Non-linear most options open	concrete poetry, decorative text	pictures with parts labeled randomly	aerial photo	maps

Note: Twyman refers to the mode of symbolization or style (columns) and the method of configuration or structure (rows). Method of configuration moves from simple (pure linear) to complex (non-linear most options open) approaches to structure. He stresses that cells are neither mutually exclusive nor definitive. Table is a representation of Twyman's matrix and exemplars are interpreted from those provided by Twyman (1979).

Acknowledging that all visual resources carry a message, as discussed by Doblin (1980), Twyman (1979) observes that the internal graphic organization or structure of the image – the method of configuration – "influences and perhaps determines the searching, reading, and looking strategies adopted by the user" (p. 121). Twyman creates his own descriptive terminology based on the fundamental notion of linearity; but he fails to adequately define linearity and, like Lohse et al., the resulting categorization relies on his selection of visual exemplars. However, Twyman's categorization of these exemplars may be supported by work in visual search strategies that involves eye movement studies and applied gestalt principles (Arnheim, 1982; Guba et al., 1964).

With respect to the specific cells, or categories, in his matrix, Twyman (1979) suggests that the verbal/numerical linear-interrupted category, best exemplified by traditional texts, is the dominant form precisely because of mass production technologies (i.e., the printing press) and associated structural conventions. He also suggests that commonly-used forms, such as Linear Interrupted, work well by virtue of the simple fact that they are commonly used, implying the notion of convention as suggested by Riegl (1992/1893). This could change within the environment of the Web, however, as alternative methods of visual representation are more readily available to and easily adaptable by non-specialists.

The typologies devised by Rankin (1990), Lohse and his colleagues (Lohse, Bilosi, et al., 1991; Lohse, Walker, et al., 1994) and Twyman (1979, 1985) all suggest confusion between function and structure as is evident in their category definitions. Rankin initially defines his categories on the basis of physical *xyz* relationships; but he proceeds to identify subcategories based on intellectual content as represented by the number of variables. Lohse et al. intend to build a functional taxonomy of informational graphics, but the use of specific visual exemplars has underlying structural implications for the resulting system of categories. Twyman's observation that the method of configuration influences the search strategies of the user leaves unanswered the question of the relationship between conventions of configuration (syntax) and domain-level knowledge (paradigm). What strategies for extracting information does the viewer adopt once she has identified a method of configuration? And can the method of configuration be coded in the representation of a resource in a digital system?

2.7 Image perception

> The laws of organization, however, may still turn out to be gross but useful prescriptions for designing pictures so that they will be comprehended as we want them to be comprehended – although we should note that these 'laws' have never been adequately formulated or measured as objective and quantitative rules. (Hochberg, 1972, p. 52)

Text has to be processed sequentially before the message can be received. Although images are also processed sequentially, eye movement studies are necessary to demonstrate the spilt second linear activity of eye scans (e.g., Guba et al., 1964; Wolf, 1970). For this reason, visual perception is described as direct experience (Berenstein, 1996). Image perception begins with the individual observing that a marked surface contains information. This is normally followed by three types of activity: attention and scanning; interpreting significant figures and cues; and perceiving global meaning (Levie, 1984).

Levie (1984) identifies eight major theoretical approaches to the process of perceiving images:

1. Perspective theory (classicism)
2. Direct perception (structuralism) (Gibson, 1979)
3. Constructivism (Gombrich, 1960; Gombrich, Hochberg, & Black, 1972)
4. Generative theory (Hagen, 1979)
5. Gestalt theory (Arnheim, 1954, 1969)
6. Perception as purposive behavior (Gombrich et al., 1972; Hochberg, 1983)
7. Semiotic approach (Pettersson, 1998; Sonesson, 1989)
8. Theory of symbols (Goodman, 1968)

The first four theories are grounded in the notion of resemblance. Perspective theory combines the resemblance requirement of classical representation and a mechanical process of rendering images; but it fails to explain the perception of three-dimensional objects when rendered as shape projections in the two-dimensional plane. Direct perception theory also depends on classical resemblance but recognizes that the viewer's knowledge of object properties will remain constant despite the rendered perspective: For example, the viewer knows that the basketball behind the bread-box is a complete circle even though it is rendered as a semi-circle. Constructivism extends direct perception by focusing on the construction of meaning through cognitive understanding – the visual and cultural conventions of the viewer, such as the word-balloons used for dialog in comics. Generative theory attempts to reconcile the theories grounded in resemblance by suggesting that image meaning is generated by the interplay between the geometry of the visual stimulus and the cognitive understanding of the perceiver.

Gestalt theory approaches the image as a balanced and self-contained whole – a *gestalt* – whose component parts and their organization influences perception of the whole: A viewer first perceives the most significant structural patterns of an image, which leads, in turn, to a meaningful interpretation of the whole stimulus. Thus, gestalt theory assumes that "the most powerful conveyor of meaning is the immediate impact of perceptual form" (Arnheim, 1982, p. xi). In contrast, and similar to generative theory but without dependence on resemblance, perception as purposive behavior argues that the viewer cannot grasp the meaning of a whole image in a single glance, but builds a mental schema through a process of sequential visual attention or successive fixations on focal areas of the stimulus.

The final two theoretical approaches to perception revolve around the study of signs rather than analysis of the perceptual process itself. Signs have two aspects: the concept being represented or depicted (the s*ignified*) and the physical entity that expresses or conveys that concept (the *signifier*). In the present context, the signifier is the image and the signified is the intended meaning – the ofness or aboutness of the image.[37] Semiotics explores the culturally constructed rules and codes that link signifier and signified, addressing Howard's (1980) logical method of symbol analysis, either within or between systems. Symbol theory focuses on the organization of signs and their referents. According to Levie (1984), signs serve three major functions – denotation (ofness), expression (of feelings), and exemplification (abstract example). These three functions are subsumed by Howard's functional method of symbol analysis.

The common thread between all these theories is the perceived entity – the stimulus geometry, structural pattern, signifier or image – and its role in communicating meaning. In her guide to online image searching, Berenstien (1996) refers to the two aspects of images as the visual – the entity – and the non-visual – the meaningful "information not contained in the picture but relevant to it" (p.23). Berenstien notes that, in order to find images, the searcher must be able to describe the image in accurate detail, translating what is perceived (the visual) "into language that a person or computer can understand" (p.23) (the non-visual).

While the image as a medium of expression (communication) is frequently explored in terms of sense stimuli, it can also be explored through application of gestalt concepts relevant to graphic design. *Gestalt* is a German word without a direct English counterpart. Loosely identified with form (i.e., composition or configuration) and shape, the notion of gestalt reflects the concept that "'wholes' [sic] are experienced as such and not as the sum of their parts" (Krampen, 1994, p. 290). Grounded in the

---

[37] Howard's (1980) "lexical analysis," which identifies various meanings of symbols, is correlated with Panofsky's (1972) iconographic and iconologic description, ofness and aboutness, as well as ekphrasis. Howard also discusses how the terms symbol, language, and representation are often equated and confused.

epistemological assumption of Descartes and Kant that the mind possesses "innate ideas about form, size, and other properties of objects" (Rock, 1984, p. 11), the concept of gestalt entered discourse in psychology and visual design early in the twentieth century. In psychology, gestalt theory was applied to perception and sensory organization to explain how the individual sees the world and to identify the relationships that exist between physical patterns in the world and human physiological activity. Kohler (1947) adopted gestalt theory to counter behaviorists who rejected the study of direct experience as a physiological activity. He argued that "when discrete entities unite in a group, the part which equality (or similarity) plays in the unification cannot be explained in terms of learning" (p. 84). For example, the unification of similar lines into a pattern of stripes or of similar circles into a pattern of polka dots illustrates the gestalt principle that similarity and adjacency of forms result in their visual grouping. In contrast, the behaviorist approach would require a new grouping concept to be learned whenever the component form changed.[38]

The principles of gestalt specifically address the perception of form and the organization of units in a composition. These principles were suggested by experiments, summarized in texts such as those by Rock (1984) and Goldstein (1980), that demonstrated both the physiological and psychological foundations of perception. These studies of perceptual organization indicated that the human visual system exploits a range of regularities in image structure. While the principles cannot predict what is perceived, they can be used as a starting point for considering the underlying physical structure of images.

The figure-ground relationship was considered fundamental for perceptual activity (Kennedy, 1985) and included principles of simplicity, good-continuation, closure, proximity and similarity (see Figure 2.8). Simplicity, or good figure, suggests that figures appear to be simple and symmetrical even when they are not. Good-continuation refers to the linearity of form and allows the individual to follow a line of text across a page; it is exemplified in Twyman's (1979) Pure Linear category. Closure is related to continuation in that it reflects the tendency of the viewer to fill incomplete areas or close gaps in what is seen. Just as it is psychologically more rewarding to achieve closure through

[38] The application of gestalt was carried to the extreme, however, in the study of optical illusions by the Berlin school, which held that there was a "direct and lawful connection between physical stimuli and their sensory perception" (Krampen, 1994, p. 291). Cataldo (1966) points out that, over time, differences in the application of gestalt to perceptual psychology were resolved by focusing on the principle of an isomorphic correspondence between the structural characteristics of visual form and the observed phenomena in human behavior.

the completion of activities, closed figures are viewed as more stable. Proximity, or adjacency, leads to visual grouping of forms based on the nearness of their parts, while similarity provides that visual parts that resemble each other will tend to be seen as a group, as exemplified in Twyman's Linear Interrupted category[39]. Similarity can involve shape, size, color, direction or even a time period. For example, similarity in both time and place characterizes the dynamic of film.

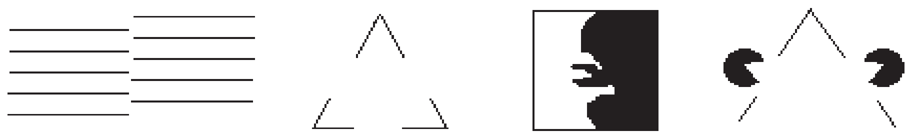


Figure 2.8. Gestalt principles. Examples of some basic gestalt principles: closure, figure ground, proximity, and similarity.

In the Bauhaus, in the early decades of the twentieth century, gestalt principles were applied in education for graphic design and visual communication and formed the basis of stimulus design. Later, Arnheim (1954) extended this work to explore the dynamics of forces and tensions, which he identified as the central pull-and-push of image design. Today, the basic tenets of functional document design, as discussed, for example, by Berryman, (1979), Hurlburt (1977), Kepes (1965), Lupton and Miller (1996, 1993), or Ragans, (1988), are in accord with the basic perceptual principles of gestalt and recognize that the transfer of information to the viewer is through the medium of vision.

As Kennedy (1985) points out, critics of gestalt argue that figure-ground is a pictorial effect – nothing more than lines and contours on flat surfaces creating edges but not depth. Kennedy contends that, "In principle, perception is simply a means of grasping the underlying mathematics of forms" (p. 35), a position that supports the structural typologies of Rankin (190) and Twyman (1979). The mathematics of form to which Kennedy refers are exemplified in the digital environment, where fractal geometry is used to create or simulate animate and inanimate forms. For example, the simulation of a landscape may be perceived as coastlines, clouds, hills or deserts, but these simulations

[39] There is not a one-to-one correspondence between the Gestalt principles and the Twyman categories. For example, in the Linear Interrupted configuration of verbal/numerical symbolic mode, the letters and numbers form a line (good continuation) to be read. They are at once similar in form and in close proximity allowing them to be read. A typographer could lay out a paragraph as if it were a matrix of letters and numbers, but it would fail in good continuation such as early forms of boustrophedon writing (Lupton & Miller, 1996). A paragraph with a different typeface for every letter would fail in the similarity principle and would therefore be more difficult to read (see Twyman, 1979 and Table 2.1).

are based on geometry and the variations of lines as edges.

Another argument criticizing gestalt derives from the theory that individuals hold in memory schema for types of objects such as noses, houses or even scene arrangements (Gombrich, 1960). The gestalt response is that individuals privilege basic forms and perceive by constructing objects out of the visual building blocks of circle, square, and triangle (Arnheim, 1954; Dondis, 1973; Lupton & Miller, 1993). From this perspective, gestalt is grounded in perceptual simplicity and relies upon the principles of basic geometry for visual analysis. Indeed, the position that there are basic privileged forms implicit both in Plato's theory of Ideal Forms "of which the world of material objects apprised by the senses was an imitation" (Collinson, 1987) and in schemata theory assumes structured knowledge held in memory and activated for comprehension (Bothamley, 1993).

If the views of privileged forms manifested in the Bauhaus movement are rejected, then the task of perceiving shape would be unconstrained and boundless, given the infinite variation of nature. This is not to say that recognition is based first or foremost on the perceptual forms of basic geometry. Rather, gestalt theory contributes the concept of an image mediated through the perceptual principles of simplicity, continuation, similarity, proximity, and closure. These principles form a geometric basis for visual analysis of the physical world that informs the interpretation of the viewer. Because coding in the digital environment is itself grounded in mathematics, it should be possible to develop coding systems that represent the basic geometric forms of perception. Developing a vocabulary for visual structure would focus attention on the geometric properties of an image. Gestalt theory and the Bauhaus movement provide direction for such a language of properties – of rules and cues (Kirsch & Kirsch, 1985) – for image analysis and description based on critical judgment and the visual differentiation of physical elements in the image.

2.8 Form and function

> … the nature of the medium of communication itself is only considered in the Design Arts. There it is theorized but not generally tested. Substantiation is borrowed from perceptual psychology and learning theory and non-visual communication studies. But the principles of image construction remain generally untested. (Morgan & Welton, 1992, p. x)

Focillon[40] (1948) argues that "a work of art is an attempt to express something that is unique, it is an affirmation of something that is whole … but it is likewise an integral part of a system of highly complex relationships" (p. 1). If *image* were to be substituted for *work of art*, Focillon's statement would reflect the immediate challenge confronting the design of digital image retrieval systems: the naming or categorization of non-verbal materials for subject access (Jorgensen, 1998; Layne, 1994; Small, 1991). All mediated images involve, at a minimum, a selection process that produces a two-dimensional entity. Responses range from the selection of a plot of ground to produce an aerial photograph, to the codification of a message into a logo, to the selection of what to display within the confines of the medium's frame, or viewport, be it a book page, television screen or computer display. All images are at once complex but nonetheless simple in their wholeness: That is, all images integrate the complexity of that which has been selected within the framework of the selector's unique perspective.

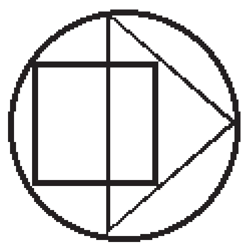


Figure 2.9. The Bauhaus icon.

The Bauhaus proposed a language of vision and a fundamental visual grammar whose central elements were the triangle, the circle, and the square (see Figure 2.9). One major approach to the process of selection and mediation has its roots in the Bauhaus. Lupton and Miller (1993) acknowledge the mythic proportions of the Bauhaus as the foundation of modern design in its search for a unified work of art centered around a system of signs that are both natural and universal. But Hurlburt (1977) contends that "no movement has been so misunderstood" (p. 38).

Misunderstandings of the Bauhaus can be traced to several sources. In the first place, the language of vision was understood as speaking directly to the eye, following gestalt theory in its most extreme application. Secondly, although the term "graphic" was acknowledged as relating to both writing and drawing as well as to conventions of data display in the sciences, its relationship with verbal language was only analogical.

---

[40] Some material in this section was previously published in Beebe and Jacob (1998).

Nonetheless, subsequent research has focused on the ability of visual languages to replicate the forms and functions of verbal languages (Carroll, 1982; Knowlton, 1966; Kolers, Wrolstad, & Bouma, 1980). These efforts to link visual and verbal forms and functions continue in the arena of digital systems with research that strives both to devise algorithmic grammars for the geometry of image processing and to match alphanumeric strings with the binary data that form the image display (Caivano, 1994; S.-K. Chang *et al*., 1986; Cobb & Petry, 1998).

The popular legacy of the Bauhaus is obvious in the clichés that purportedly originated with two of its proponents: "less is more" and "form follows function." "Less is more," attributed to Ludwig Mies van de Rothe, directly addresses the concept of gestalt: The idea that the whole is greater than the sum of its parts. It specifically references the principle of good-figure and applies it as a guideline for design simplicity.

"Form follows function," attributed to Louis Sullivan, has evolved into a starting point for the design process. Sullivan was an American architect who advocated the design principle that "a building should publish to the beholder the purpose which it subserves – what it is for, what it is about, why it is as it is and not otherwise" (Bragdon, 1922, p. 5). Bragdon's accounts of his interaction with Sullivan demonstrate that, for Sullivan, "function" meant the natural characteristics of the materials themselves and not the sociocultural application or utility of the final product. Thus, for example, a building using steel-framed construction materials of the industrial era should not be given the appearance of a solid masonry structure. Rather, the form should follow the functionality of the constituent materials and of the techniques employed in constructing with those materials. This concept is highlighted by Van Sommers (1984) who distinguishes between modern objects and those from an earlier period when function was more evident than form. For example, "a steam locomotive did not conceal its crew, its fuel, and its combustive processes in the way a modern diesel does" (Van Sommers, p. 126).

The notion that function should determine form was intended to address the need for architecture to take into account the unique characteristics of the construction materials themselves. Because Sullivan's idea that form follows function has been removed from its original context, its true meaning has consequently been lost. Today, "form follows function" is often taken to imply a linearity of process, where function is misinterpreted to mean purpose and not the characteristics of the materials that determine the form of the product. Popularization of "form follows function" has led to misinterpretation of Sullivan's intent, privileging the functional utility of the end-product and thus neglecting or ignoring the characteristics of the constituent materials.

This misinterpretation of the relationship between form and function is analogous to the relationship between the verbal representation and the image in digital retrieval environments: The graphic form of the image – the material and its structure – is neglected in deference to meaning and verbal description.

Although spatial dynamics – the geometry of the gestalt involved in interpreting the physical world – provides for a relationship between individuals and the representations and creations of objects in the world in which they live and move, they are frequently taken for granted. Bragdon (1922) points out that, in the microscopic world of nature, the constructing units tend to arrange themselves with relation to simple geometrical forms – the Bauhaus's visual building blocks of circle, square, and triangle. These forms give rise to unity and simplicity in the face of complexity: the honeycomb, the snowflake, the proportions of the human figure, the triangular arrangement of features in the human face. These complex forms reflect Mies van der Rohe's notion that "less is more," Berger's (1989) images made up of elements that are themselves images, and Carroll's (1982) innate structure. Boulding (1968) suggests a typology of images beginning with the simplest level of organization, called static structures, which are exemplified by the statue, the picture, or, more abstractly, the atom. Successive dimensions of organization are identified as mechanical, homeostatic, biological, botanical, animal, and human; but, because Boulding's typology is hierarchical, each subset in his organization includes the characteristics of all its superordinate levels. Similarly, indexing practice subsumes the conceptual dimension of representation when creating verbal contextual representations in the form of abstracts. The internal structural relationships of the image (i.e., internal contextuality) may equate to Boulding's static structures or may constitute a new dimension of contextual abstracting, yet it is not considered in either case.

Barthes (1985) explores the relationship of form to meaning in a discussion of a single frame in Eisenstein's film *Ivan the Terrible*. This specific frame shows two courtiers showering the young tsar's head with gold. Barthes distinguishes three levels of meaning in this frame: the informational or communication level; the symbolic level; and a third that he identifies as the obtuse level. On the informational level, Barthes finds meaning in the setting, the costumes, the characters, and their relationships. On the symbolic level, the shower of gold conveys the theme of wealth and a more general signification of exchange. The obtuse level encompasses the physical features of the image itself such as the density of the courtiers' makeup, the delicate line of one courtier's eyelid or the shape of the "stupid" nose on the other (p. 42). These are the structures of the image – the forms or shapes selected by the filmmaker – that contribute to the gestalt of the frame. Furthermore, the obtuse meaning is not semantically situated but "remains suspended between image and its description" (Barthes, 1985, p. 55). The implication is that there can be no linguistic description of this meaning, suspended as it is between the visual and the verbal. However, the principles of gestalt might be brought to bear to explicate the relationship between communication and structure – between form (the internal structure of the image) and function (image as communication). Focillon (1948) captures the essence of Barthes' obtuse level in his statement regarding the principle of form: "the sign bears general significance, but having attained form, it strives to bear its own individual significance; it creates its own new meaning, it seeks its own new content, and then endows that content with fresh associations by the dislocation of familiar verbal modes" (p. 5).

In his delineation of the three steps in the process of image analysis (pre-iconographic, iconographic and iconologic), Panofsky (1972), identifies two aspects of pre-iconographic analysis: the *factual* and the *expressional*. Panofsky describes factual analysis (the ofness from Shatford, 1986) as the identification of pure forms, that is, "certain configurations of line and colour, or certain peculiarly shaped lumps of bronze or stone, as representations of natural *objects* such as human beings, animals, plants, houses, tools and so forth" (1972, p.5. Emphasis in the original.) These forms are referred to as artistic *motifs*. Expressional analysis identifies the emotive qualities of pose or gesture, such as *happy* or *sad*, or the atmosphere of a scene, such as *peaceful*. The expressional and the factual combined constitute Panofsky's primary subject matter and "an enumeration of these motifs would be a pre-iconographical description of the work of art" (1972, p.5).

Focillon (1948) expands the notion of pre-iconographic motif by focusing on the configurations of line, color and shape. He describes three categories of motif: variation of form with the same meaning; variation or multiplicity of meaning with the same form; and form devoid of meaning. Doblin (1980) illustrates the first two categories when he describes encoding the same message in three ways (e.g., three o'clock, 3:00, or ) and multiple messages using a single form (e.g., **I** interpreted as I the speaker, the numeral one, or an image of a pole). Focillon illustrates the third category – form devoid of meaning – with the history of the interlace form which has its origins in the symbol of entwined snakes in the caduceus of Aesculapius, the

Roman god of medicine and health. He points out that the symbolic association with medicine was lost as the interlace form was incorporated into the ornaments and monumental architecture of Islam and East Christian communities and ultimately took on a life of its own (see Figure 2.10). Collections of motifs are thus *forms devoid of meaning* from which artists select those to be used in a new context, either to take on new meaning or as a variation of form or meaning.

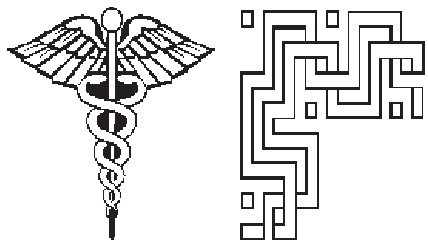


Figure 2.10. The interlace form. Focillon (1948) suggests the intertwined snakes of the caduceus were the historical foundation of the interlace form. The interlace form lost its symbolic association with medicine and took on new associations as it was incorporated in new visual applications or ornaments. The interlace form developed many variations of meaning through new applications and new associations, thus exemplifying Focillon's notion of "form devoid of meaning."

Focillon's notion of *form devoid of meaning* extends Panofsky's pre-iconographic category beyond object identification to include the physicality of the image itself. It suggests that forms are first perceptually organized or recognized and then imbued with meaning, either through identification (pre-iconographic ofness of form) and higher level categorization (the iconographic) or through interpretation of meaning (the iconologic). However, Focillon's *form devoid of meaning* is not addressed at all in verbally-based information retrieval systems.

Focillon (1948) does not address the possibility that, within a single representational system, a form may have only one meaning and a meaning only one form. In information retrieval systems, the indexing of images is facilitated by this principle of one form with a single meaning in that it involves the controlled representation of entities, parts and properties. This process serves as the basis for indexing languages in most information retrieval systems, including image-based systems. But, if objects are represented solely by linguistic interpretations of conceptual meaning generated within a verbally-biased context, then the flexibility inherent in Focillon's expansion of Panofsky's (1972) pre-iconographic categories is lost;

and objects residing in both public and private collections are at risk of becoming irretrievable.

2.9 Relevant Research

2.9.1 Queries and searchers

Studies of searchers and their queries have been limited by narrow focus on specific applications, collections, or systems. Of the small number of studies that exist, many assume the art historian to be the typical user of image databases, although a study at Pennsylvania State University found the largest user base was in the earth and mineral sciences (Pisciotta, Dooris, Frost, & Halm, 2005). Unfortunately, most studies look only at the specific digital collection being developed and their findings cannot be generalized. Furthermore, system developers are frequently focused on the implementation of a particular metric rather than its appropriateness for a user need or its suitability for a particular interface style (i.e., browse versus target search). There are no studies that compare search requirements across domains or address image searcher behavior as opposed to that of the text searcher. Additionally, the demographics of image searchers in the digital environment has yet to be addressed.

Enser and McGregor's (1993) study of the needs of image searchers was constrained by its narrow focus on requests placed with a single news photo archive. They determined that 69% of the requests sought unique entities (such as "paddlesteamer Medway Queen") and 34% non-unique entities (such as "paddlesteamers"). Although significant refinements to non-unique requests – refinements such as time, location, action, event, or technical specification[41] – were clearly stated, they were difficult to satisfy. In a similar study with art historians, Hastings (1999) analyzed user queries, grouped them according to type of query – who? where? when? and what are? – and then attempted to map the queries to the most appropriate access mechanism. Hastings concluded that the more complex the query, the greater the need for image browsing and content-based retrieval criteria (e.g., color, shape, texture). H-l. Chen's (2001) study also focused on art history and validated the query categories of Enser (1993) as well as the indexer categories of Jorgensen (1996b, 1997a). Text-based image searches of art history students were analyzed by three judges, who reported little confidence in the effectiveness of retrieval strategies for non-unique-refinement queries as well as for Jorgensen's categories for

[41] Examples of technical specifications include silhouette, aerial-view, and image-orientation (portrait or landscape).

location and visual elements. Chen suggests that traditional, text-based access points are relatively ineffective and that training in query formulation reflecting non-unique-refinements, locations or visual elements is a potential development arena for retrieval systems.

In summarizing user query research, Jorgensen and Jorgensen (2005) note that requests for specific objects dominate over general topics, categories or abstract concepts. Yet, in their own analysis of more than 1400 queries in the search logs of a web-based image provider, they found that unique term searching was less common than indicated by earlier research and that descriptive queries were more frequent. Although 50% of the unique terms were nouns, only 5% were proper nouns. Descriptive adjectives accounted for 19% of the terms and verbs for 11%. Only 1% of the terms were prepositions, which they claim reinforces the rare use of visual construct queries. However, as Landau and Jackendorff (1993) point out, there are a surprisingly few number of prepositions (they estimate 80) in comparison to the number of different kinds of objects.

CBIR search types include target, category and browse and are closely related to searcher goals. Target searches are precise. In a target search, the searcher may want to find the picture of a young girl having her portrait painted with her attendants. The canvas and artist are in the foreground, and there are two people in the doorway in the back of the room. The searcher doesn't know any factual data about the image (*Las Meninas* by Velasquez), but could perhaps draw a sketch of the scene (see Figure 1.4). In a category query, the searcher may want an image with specific basic characteristics, like a garden or a sunset, but does not have a specific image in mind. Forms of category search are query-by-pictorial-example (QBPE) and "more-like-this" searching (see Figure 1.2 and Figure 2.3.) which can identify resources in the image database that match the searcher's exemplar. On the other hand, the searcher may have no idea what she wants and a general browse approach, or search-by-association, will help her to focus. Some CBIR interfaces give the searcher control over the thresholds set on image properties, such as level of histogram or degree of contrast. The searcher, however, must understand the meaning and implications of histogram control before these techniques can be viable tools.

In addition to searcher needs, various demographic variables have been suggested or applied in query research with human subjects. The research of Kolers and Smythe (1979) who experimented with subject memory for images and found that people remember not because of pictorial qualities but because of personal skill and experience in using a system of symbols. They concluded that experience constitutes an important aspect of symbol use, which is itself a form of knowledge. In the image encoding process, Twyman suggests that choice of category for creating graphic language documents (see Table 2.1) may be affected by whether the designer is a specialist or non-specialist in a given visual domain. Lohse and his colleagues (Lohse, Walker et al., 1991; Lohse, Bilosi et al., 1994) reached a similar conclusion when they screened subjects for their experiments in graph typology development. They reported that subjects with graphic art training were more consistent in their categorization of images than subjects without graphic art training. And Fabun's (1965) work suggests that there may be a cultural domain component as well:

> Americans tend to see the edges of things and the intersection points of crossing lines, and to attach importance to them. Thus our streets are normally laid out in a grid pattern and we identify places by their proximity to intersections. Europeans and Orientals, however, are more inclined to attach importance to an area, thus a French street or avenue may change its name every few blocks; and houses in Japan may not have street numbers but be identified by name and area or the time at which they were built. (p. 25)

Giardetti and Oller (1995) identified factors that might play subtle roles in the subject's differentiation of photographic meaning: education; training in photography (judged by self-reports); ethnicity; gender of viewer; and photographic qualities (such as color versus black and white.) Drabenstott (1999) suggested that the most successful users were those without subject expertise but with knowledge of the structure and content of the catalog. When reviewing research relevant to learning from visual materials, Fleming (1984) noted that gender differences play a significant role in various tasks concerned with the perception of images. These findings indicate that the individual differences of users could affect the query process and need consideration when developing query mechanisms.

2.9.2 User studies in CBIR

CBIR systems have focused on the development of specific technologies for feature extraction, data management, and retrieval mechanisms and have generally failed to engage in systematic user studies in practical applications (Rui et al., 1999). Most system evaluation is either informal or based on some form of benchmarking (e.g., comparing the results of one method against those of another) with success judgments established by the system's designers. As a benchmarking example, Ratan and Grimson

(1997) constructed low-resolution templates (models of color and spatial arrangements) using color-space attributes from user selected target areas. They compared the images retrieved from their templates to those retrieved by systems using color layout and histogram methods to generate a benchmark for system precision. Typical of CBIR evaluation, they assumed that benchmarking was an acceptable surrogate for the judgments of real users.

PicHunter experiments (Cox, Ghosn, & Miller, 1997) compared similarity judgments provided by users working on a system with low-level, non-verbal properties (e.g., color, contrast, brightness etc.) to those provided by users working with a newer hybrid version of the system that included broad semantic attributes (e.g., sky, hill, city, person etc.). In the test situation, semantic attributes were kept hidden from users in order to evaluate the performance of the system itself. These experiments showed that users working with the hybrid system required fewer searches in the target-seeking task than those working with the unannotated, low-level system. The more interesting finding, however, was that the number of searches was reduced even further when the method of similarity judgment was explained to the user. It was also interesting that users' similarity judgments relied on a small number of relevant features, suggesting the potential for targeting salient features as a means of focusing similarity comparisons.

Typical of much CBIR research, user testing of the MARS-1 relevance feedback interface actively sought to keep user input to a minimum (Rui et al., 1997). The focus of user trials was on the viability of the feedback computational model. Feedback to the user consisted of a display of thumbnails that were visually similar based on color, shape and position and provided what the researcher described as "an improved retrieval result" (p. 88). Researchers indicated that the user was not required to specify the exact values of his query but instead would only need "to mark the retrieval returns that he thinks is (sic) relevant" (p. 88) to provide feedback to the system. However, the researchers used their own similarity judgments to interpret the feedback the users provided to the system. Rui et al. designed their system based on the express assumption that it would be advantageous to keep user input to the query process coarse and minimal. It is ironic, then, that they should have concluded that "retrieval performance improves significantly by incorporating humans in the retrieval process" (p. 82).

CBIR research has typically demonstrated a tendency to make unwarranted generalizations about user behavior. For example, Ravela and Manmatha (1997) stated that "relevance, for users of a retrieval system, is most likely associated with semantics" (p. 67), an assumption about users that appears to be grounded in the privileged status of text. Similarly, Santini and Jain (1997, 1998) claimed that users reason on a semantic level “in which objects, not perceptual clues, are the main concern” (1997, p. 3) and that “preattentive similarity judgment is done without focusing attention on any part of the image” (1998, p. 2). In developing their interactive interface, they assumed that “the meaning of an image will be revealed to the database by interpreting the sequence of queries posed by the user” (Santini et al., 2001, p. 337). This position is similar to that of Belkin (1982) in that, similar to his notion of anomalous states of knowledge (ASK), it assumes the naiveté of the user in the query process. Thus the interface that Santini and Jain developed included little semantic interaction between the user and the system but depended on the ability of the system to identify visual similarities based on its computational indexing capabilities and on the ability of the user to guess the basis for the system’s identification of similarity between two images.

Development of the WebSeer system (Athitsos, Swain, & Frankel, 1997) was based on statistical observation of system logs of user queries. Based on a test period in which there were over 80,000 queries, 82% of which specified whether images should be photograph or graphic, the researchers generated a set of metrics intended to differentiate the characteristics of graphics from those of photographs. While this may be a viable approach, user feedback is needed to determine whether the choice was somehow imposed by the structure of the system or if the differentiation of other image types might be a high priority for filter development.

In a recent study on color-wheels in user interfaces, CBIR researchers were interested in minimizing the number of colors in the CBIR interface, which typically presents the searcher with millions of color choices. The researchers (van den Broek *et al*., 2004) found that the use of color-wheels was overwhelming for subjects and they were too naïve in the field of color-spaces (e.g., RGB, CIE) to be able to use that method effectively. The amount of color variation available was judged as unnecessary.

Other assumptions of CBIR developers involved QBPE as the simplest method for user interfaces. Wang (1999) claims that, “to complete a query, a user only needs to specify the query image” (The SIMPLIcity Retrieval System, para. 7) thus providing the user with a simple query interface. In their review of CBIR research, Smeulders et al. (2000) commented that user-system query interactions attempt to balance “between obtaining as much information from the user as possible and keeping the burden on the user minimal” (p. 20) and further observed that “the simplest form of (user) feedback is to indicate which images are relevant” (p. 20). However, because searchers have traditionally been limited to verbal forms of access to images, access based on image properties – the internal contextuality of the image – has not been examined.

Fortunately, there is a growing tendency to include the user in the evaluation process[42] (Rui, Huang, Methotra & Ortega, 1997; Rui, Huang, & Chang, 1999). With increased interest in automatic annotation and feature matching, developers are asking users to annotate images to test against system-generated annotations (Tsai et al., 2006) or, more typically, to enable the system to model user preferences (Mountrakis et al., 2004). Mojsilovic et al. (2004) were the first to look at the possibility of natural language queries. Trying to map semantic categories to feature metrics, they asked subjects to group thumbnails into a small set of developer-determined categories; they then asked subjects to provide names for the categories and write a brief category description. Unfortunately, the developers did not see the advantages of constraining the resulting terms in the form of a structured vocabulary. The result was a set of category-specific modifiers rather than a vocabulary of cross-category attributes.

2.9.3 Image description

Experimental research that reports on text-based image description is sparse. Lohse and his colleagues (Lohse, Walker et al., 1991; Lohse, Bilosi et al., 1994) had subjects sort images into visually-similar categories. Giardetti and Oller (1995) explored thematic categories mapped to prototypical images. And Jorgensen (1995a, 1996a, 1996b, 1997a, 1997b, 1998, 2001) used verbal description methods to collect terms about images.

Jorgensen's research (1995a; 1996a) is by far the most consistent and comprehensive but it focuses on developing indexer tools rather than on user needs. She reports on the properties of images that she identified using the constant comparative technique for theory generation. The goal of the constant comparative technique was to identify image attributes in subject-generated verbal and written descriptions, to define attribute types, and to build superordinate categories of attributes. Based on the similarity of the sorting task and the decision making necessary in classification, Jorgensen began with an image-sorting task to generate a vocabulary of properties. Consistent with the methodology of Glaser and Strauss (1967, pp. 101-115) Jorgensen allowed the categories to emerge from the simultaneous coding and analysis tasks, as there were no pre-existing schemes identified. Using 77 images selected from an illustration annual, subjects talked aloud as they sorted the test images into categories for their own use at a later time.

---

[42] In CBIR research this is referred to as "human in the loop."

The results identified 47 different subject-generated properties that were subsequently organized by Jorgensen into 12 categories. The first six categories were perceptual properties (*objects*, *people*, *color*, *visual elements*, *location*, and *description/number*); the next five were interpretive properties (*people relationships*, *art history information*, *abstract concepts*, *story/content*, and *external relationships*); and the last was the *viewer reaction* property (e.g., conjecture, uncertainty, etc.).

To confirm her original categories, Jorgensen (1996b, 1998) expanded the task to describing, searching and memory. Eighty-two subjects were given two minutes to write natural language descriptions of six different images. Half of the subjects were instructed to describe what they "noticed" in each image, while the other half were instructed to describe each image "as if it were an image which they hoped to find in a collection" (1996b, p. 209). After four weeks, the subjects who had described what they noticed were asked to describe what they remembered of the six images. Jorgensen concluded that all property categories were confirmed with *literal object* the most common, followed by the groups *color*, *people*, *location*, *story* and *visual elements*.

Jorgensen (1996b, 1998) then tested an indexing template derived from her twelve property classes. The describe task was repeated with 48 new subjects. In this new research, however, subjects were given a template derived from the twelve categories along with brief category descriptions and were asked to place each descriptive term in the appropriate category. Results indicated that, as a group, subjects were inconsistent in the assignment of their terms and phrases to the twelve categories. Jorgensen concluded that some categories were not well understood and that the use of descriptive phrases had further confounded category assignment. For example, location of picture components (e.g., *above* or *on*) was misunderstood as location of picture (e.g., *cave* or *Japan*). Jorgensen interpreted this inconsistency as a result of both the loss of visual information in the template terms and the misinterpretation of category names and suggested that the subjects were naïve and needed more detailed instructions. Although she did not control for the effects of visual training or experience,[43] she concluded that the set of image description classes had been re-confirmed and that the importance of perceptual properties had been validated.

Jorgensen's (2001) more recent testing of her indexing template as well as the pyramid model of image representation methods (see Figure 1.1) introduced by Jaimes and Chang (2000) used the same describe, search and memory tasks as in previous

[43] One source refers to the amount of visual training and experience as the level of "visual literacy" (Pacey, 1983).

research. Seven hundred images were selected from the Web and four of these were randomly presented to each of 41 naive users and 22 students of indexing. Results indicated that "when naïve participants were asked to describe images more formally in a *retrieval* context, we see that (properties) then occur at these lower syntactic (perceptual) levels … as with descriptions generated by *indexers*" (p. 942). She interpreted these findings as supporting the applicability of the Pyramid model for representation of visual properties as well as its utility as a crosswalk to her template.

Jorgensen's (1996a) research indicated that type of task (describing, sorting, or searching) has a major impact on the distribution of image properties assigned by subjects. The description task resulted in the identification of perceptual properties, while the sorting and searching tasks provided interpretive and reactive properties (Jorgensen, 1997b). Based on this and on her previous work, Jorgensen (1997b) concluded that a variety of properties are needed to facilitate retrieval of images and highlighted the value of *color* and *location* as perceptual categories.

Findings from the referential communication research of Fussell and Krauss (1989, p. 511) are consistent with the findings of Jorgensen et al. (2001) that image retrieval descriptions provided by users frequently include properties at the syntactic – or perceptual – level. Fussell and Krauss's research indicated that subjects describing nonsense figures for identification by others in a shared communicative environment tended to use literal descriptions that relied on geometric elements. In addition, Hughley (1989) found that the verbal responses of older children (generally 11 to 15 years old) to works of art are dominated by composition, style, degree of realism, and formal elements, while the responses of younger children (generally 2 to 11 years old) are dominated by subject matter and color.

In their work with still photography, Giardetti and Oller (1995) noted that pictures have different levels of communicative potential and propose empirical research of the author/message and audience/message processes. Giardetti and Oller identified six broad categories (transportation, communication, production, psychomotor, cognitive, and affective) in order to determine how prototypical photographs might be used to illustrate themes. They selected eleven *National Geographic* photographs for each of the six thematic categories, for a total of sixty-six images. Thematic typicality was evaluated by two panels of expert judges: one group consisted of three professional photographers and teachers of photography; and the other included ten volunteers participating in a national conference on photographic education. Photographs were assigned to categories by majority consensus across both groups. For each image, Giardetti and Oller then assigned a unique picture caption that referred to its category in very general wording.

Using a scale of 1 (least typical) to 10 (optimally typical), the judges rated all photographs in each thematic category for typicality. This allowed the statistical selection of the single most typical photograph in each category plus standard deviations for the next four typical photographs, resulting in a ranked test set of 30 images. These images were tested with 20 subjects, who were asked to sort the thirty pictures into categories of their own construction. Using factor analysis, photographs in five of the six categories were demonstrated to be significantly typical of the category. Giardetti and Oller suggest that there is a kind of "language" aspect to visual literacy (p. 143) and that the cross-subject results of the typicality judgments imply the influence of socio-cultural factors as explicated by control factors (e.g., age, gender, and education).

Lohse and his colleagures (Lohse, Bilosi et al., 1991; Lohse, Walker et al., 1994) used cluster analysis methods to develop a structure-based categorization of images. Relying on the work of Tufte (1983) and Bertin (1983/1967) to provide prototypical exemplars, he selected 60 images with no attempt to avoid bias or provide for random selection. Sixteen subjects were subsequently asked to complete three tasks with the set of 60 images: naming, rating and sorting. The naming and sorting tasks were similar to the describing and sorting tasks of Jorgensen.

In the earlier research project of Lohse, Bilosi et al. (1991), the naming task generated descriptive keywords for each of the images used in the study. Through analysis of keyword frequencies, he identified binary dimensions of ten properties: spatial/non-spatial; temporal/non-temporal; concrete/abstract; hard to understand/ easy to understand; continuous/discrete; attractive/unattractive; emphasizes whole/ emphasizes parts; numeric/non/numeric; static structure/dynamic process; conveys a lot of information/conveys little information. In the rating task, judgments were made by each subject for all image items using these ten properties. Lohse, Bilosi et al. conceded that the rating task might have biased the subsequent sorting task where subjects were asked to sort images by similarity after having named and rated them.

Although the research of Lohse and his colleagues (Lohse, Bilosi et al, 1991; Lohse, Walker et al, 1994) represents a rare experimental approach to the question of graphic taxonomies, it is most notable for its failure to define document type and for an obvious bias toward graphs, including only token photographs and icons[44] and no graphic art. Thus, when subjects were asked to create a hierarchy,

[44] Lohse, Walker et al.(1994) use the term "icon" in the sense of a practical code.

three categories fell out easily: *graphs/tables*, *maps/diagrams*, and *network charts*. In contrast, *icon* and *photograph* formed outlier categories. Additional problems with the approach of Lohse and his colleagues are evident in his findings, which tended to confirm the original six categories – graph, table, map, diagram, network, icon – but added five new categories (time charts, cartograms, process diagrams, graphic tables, and pictures [photographs]), which were actually sub-divisions of the images in the original categories. This raises a question as to whether there would have been additional categories if the stimulus set had included more examples of different types of graphs, as identified by Rankin (1990). Furthermore, the single category grouping for all photographs suggests similarity judgments based on the single characteristic of media type rather than any internal properties of the images.

With the possible exception of Jorgensen, none of the research discussed here addresses image collections of the breadth and size available when accessing the Web. Giardetti and Oller (1995) used experts to select 30 exemplars from *National Geographic* and focused on photographs only. Lohse and his colleagues (Lohse, Bilosi et al, 1991; Lohse, Walker et al, 1994) focused on informational graphics – images used to solve problems – which were selected, in large part, from the work of Bertin (1983/1967) and Tufte (1983). Jorgensen's early research (1997a) used images selected from an illustrators' annual, thus limiting potential descriptors to the domain of illustration as opposed, for example, to photographs. In a later study, Jorgensen (2001) did select a sample of 700 Web images, thus addressing breadth of collection, but images that were not at least pictorial in part were specifically eliminated. CBIR systems such as QBIC and Virage begin to address both size and breadth of Web-based image collections but remain experimental, with little real-world application, due, perhaps, to the lack of any serious approach to usability testing.

2.10 Closing the semantic gap

The semantic gap reflects a disconnect between the binary code of the computer and the intellectual code of the searcher and is demonstrated by two interrelated semantic problems in CBIR. First, in their two-dimensional displays, CBIR systems represent real-world objects as hierarchies of polygons or as geometrical functions for which the searcher has no form of reference. Second, because the searcher's request may be ambiguous or indistinct, there is a gap between the user's conception of her needs and the CBIR image features that will best fit it since the system has no way of relating high-level verbal concepts to representative geometry.

One approach to closing the semantic gap is to quantify the fuzzy fit between a request and a description through QBPE system learning techniques (Crehange, Foucault, Halin, & Mouoaddib, 1989). This approach would aid the system in identifying user need based on the visual choices the user makes among retrieved images: "the images he chooses play the role of (positive) examples and those he rejects play the role of negative examples" (Crehange et al., 1989, p. 626; see also Quack et al., 2004). A variation of this approach has the searcher interactively mark the area of interest in a given example or result set (Rahmani et al., 2005).

Another approach to the semantic gap is to build a transitional vocabulary, for which there is an early precedent in the work of Papadias (1999). Using similarity measures to compare pixel positions based on direction and distance from each other, researchers assign reference terms for direction (e.g., *left, right, up,* and *down*). These directional terms are assigned by comparing the binary encoding of each pixel (e.g., R100000000 is upper left, R11000000 is lower right) and then computing distance directly from the codes. In query-by-sketch, these distance-direction codes are used to create the query. When the query is expressed in SQL, "linguistic terms may be used instead of bit-strings, e.g., meets-north (v0, v1) instead of R001111100-001100000" (Papadias et al., 1999). These directional terms are not conveyed to the searcher; and it has not been demonstrated that these terms would be either understood by or of use to the searcher.

Nonetheless, Cobb and Petry (1998) have proposed just such a vocabulary for the description of features in GIS systems. They suggest that the "manner in which spatial data is modeled affects important aspects of its use, querying capabilities and relationship inferences" (p. 253). Using satellite imagery, they propose to define shapes and directional relationships between topological features (e.g., forest, desert) that can then be used for querying. Topological features are bounded within a minimum bounding rectangle (MBR) – a rectangle that is as small as can contain the object. They then apply a finite set of basic directional relationships that can occur between MBRs within an image frame. These basic relationships have two sets of characteristics that are applied on both the horizontal and the vertical axes: 1) *disjoint*, *tangent*, *overlapping*, *contained*; and, 2) *before*, *meets*, *during*, *overlaps*, *starts*, *finishes*, and *equals* (see Figure 2.11). This system of relationships provides the groundwork for tying the directional features of objects in an image to their semantic categories and/or descriptors. A sample query might be finding a hosue (H) with a garage (G) that meets but comes before H, which comes before and overlaps a pond (P), which occurs before and during a forest (F). Although their domain of attention is the features in geographical information systems (GIS), the same concept could be applied to other domains once the features are clarified. Zhang (1998) suggests a similar approach in an iconic indexing technique "to index pictures by pictures" (p. 122). He proposes the use of symbolic pictures – shapes such as circle, square, rectangle, triangle – to index the original. To define the spatial relationships between objects, he has devised a location vocabulary of basic spatial relationships: *in*, *left-of*, *below*, *in-front-of*, *inside*, *outside*, *overlaps*, etc.

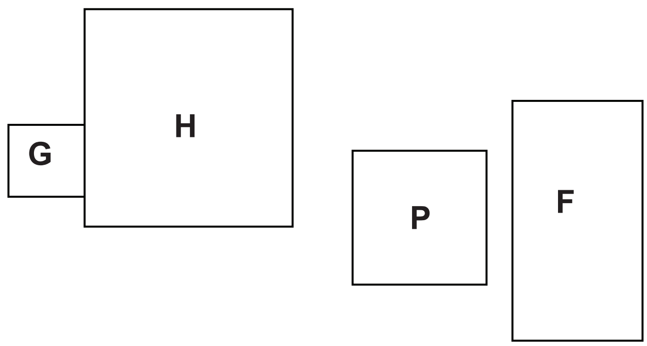


Figure 2.11. Minimum bounding rectangle (MBR) relationships. MBRs are based on a set of 85 possible directional relationships.[45] In this example created from the ideas of Cobb and Petry, 1998, p. 257, the relationships of *before, overlaps, during,* and *meets* are demonstrated within or pertaining to a single plane. In the horizontal plane, moving from left to right, G (garage) comes *before* H (house), G *meets* H because they share a common boundary, H comes *before* P, and P (pond) comes *before* F (forest). In the vertical plane, G occurs *during* H, H *overlaps* P, and P occurs *during* F.

Rogowitz et al. (1998) reported that user studies on psychophysical scaling experiments of image similarity judgments suggested potential identifying semantic categories that correlate with image processing descriptors: "For example, images with indoor scenes tend to be brownish, have low light levels and many straight edges" (p. 589). Santini et al. (2001) proposed an interface in which searchers directly manipulate pixels in image spaces, arguing that the system would subsequently learn image similarity

[45] When two MBRs are compared, the seven relationships (before, meets, overlaps, starts, during, finishes, equals) form 49 base relations and 36 inverse relations (7x7) + (6x6) because the inverse relationship of *equals* duplicates its initial definition.

through searcher interaction and grouping. However, their proposal provided no semantic approach to the representation of image properties.

Transitional vocabularies are also being explored through various data mining techniques: matching keywords to metric features (Quack et al., 2004); automatic annotation (Tsai et al., 2006); pre-sorting by image genres, such as natural scene and photograph, etc. (Mojsilovic et al., 2004); and pre-defining objects based on their structural forms (Wu *et al*., 2005b). Most promising is the work of Belkhatir, Mulhem, and Chiaramella (Belkhatir, Mulhem, & Chiaramella, 2005), who have proposed a visual semantics facet that organizes the semantic concepts encountered in previous transitional vocabulary studies. They mapped feature metrics from an assemblage of previous research results to each concept and its location in the image frame: for example, sky (generally blue) in the top region, houses in the middle region, and green grass in the lower region. They then validated their results by running similarity measures against the image clusters formed by query results, even though this approach is recursive since the concepts were derived from these metrics. Adding human input in addition to a well-structured vocabulary might increase the viability of this approach.

In a technique similar to that used for document clustering, Barnard and Forsyth (2001) combined text with image property clustering to facilitate image browsing, since testing of their statistical model indicated that image-only clusters lose semantic similarity and text-only clusters lose image similarity. Subjects were asked to identify coherent clusters of images from a set containing those generated both by image-text processing and by random assignment. “The definition of coherence was left up to the subject, but the subject was aware that both semantics and visual content could play a role, and that the occasional anomaly did not necessarily mean that a cluster should be rated as a non-cluster, if the overall coherence was adequate” (Barnard & Forsyth, 2001 p. 411). Subjects identified clusters as coherent with 94% accuracy. The findings of Barnard and Forsyth appear to support use of a general-level browsing environment with prototypical thumbnails representing various clusters of images. Although this approach makes the more-like-this search more accurate, it still fails to pinpoint either searcher criteria for similarity or coherence factors. In addition, the searcher does not understand how image properties are identified by the system.

Barnard and Forsyth (2001) suggested that *auto-illustrating* and *auto-annotating* are possible by linking similar properties with associated term clusters (object recognition) or term clusters with properties (naming objects). They acknowledged that some words and properties, such as omelet, could never be recognized or named by the system due to the specificity of concept and property differentiation. However, they failed to consider that training the user as well as the system might enhance precision in the use of a descriptive terminology.

Di Sciascio *et al*. (Di Sciascio et al., 2002) explored building shape descriptions for objects, based on Biederman's (1987) RBC theory and structural decomposition. Similar to the minimum bounding rectangle concept (Cobb & Petry, 1998), they created basic object shapes – circle, square, triangle – around image objects, then noted their location relationships. Descriptions could then be approached either from the bottom up, shape to relationship, or top down, relationship to shape. Their greatest challenge will be the computational intensiveness needed to build such feature indexes.

Finally, there is the sketching interface used in the Hermitage collection, which would also benefit from training the user. Using this type of interface, the searcher draws the basic shapes of the target image (including location and color) essentially building the description from bottom up –shapes to relationships.

2.11 Conclusion

The privileging of text over image affects the development of CBIR systems and their effectiveness. Because analysis and understanding of texts is the focus of most educational systems, the lexical and syntactical structure of text is well understood, as is the match of terms or synonyms through Boolean-type queries in text-based retrieval systems. Searchers are not experienced, however, in making exemplar sketches of the salient features of an image, in numerically generalizing visual characteristics, or even in applying verbal description. If non-linguistic information (e.g., binary structures of digital objects) is to be represented only through text, then current approaches will continue to focus on retrieval methods designed for text.

Research on digital pattern matching (see, for example, Barnard & Forsyth, 2001; Leung, Hibler, & Mwara, 1992; Lew, Sebe, & Huang, 2000; Smeulders *et al*., 2000) points to the feasibility of an image-specific approach to CBIR system design through application of the notion of the internal contextuality as an approach to visual abstracting. Through pattern abstracting, for example, a portion of one digital image can be matched against that of other digital images and patterns can be identified and compared (Caivano, 1994), thus establishing a mechanism for similarity judgments.

Unfortunately, however, the notion that pattern matching "is really the visual equivalent of the 'word search' that is a standard feature of every word-processing and database package" (Vaughn, 1992, p. 2) has led to a focus on the application of traditional, text-based algorithmic approaches without consideration of how alternative methods of representing non-textual data might be achieved. Gouge and Gouge (1996) questioned the value of pixel pattern matching as a viable retrieval mechanism; and Hibler, Leung, and Mamock (1992) explored the potential for assigning entity-attribute-relationship terminology to both the image and the query in order to identify location relationships as entity proximities in the vertical and horizontal planes, as demonstrated by Cobb and Petry (1998). This type of description could be considered analogous to the indicative abstract, which is not interpretive but "contains descriptive information on purpose, scope and methodology in a given document … (and) is intended simply to indicate – or describe – what the document is about" (E. Jacob, 1996, personal communication). Thus there is the possibility for querying particular patterns within an image, such as pixel structures that form stripes, to identify aboutness at its most primitive level – an image about stripes.[46]

In the digital realm the visual appearance of an image (form), whether created digitally or serving as a digital surrogate, is determined by the capabilities (functionality) of the materials used in its construction – the pixel – and this should not be neglected in deference to verbal description. In gestalt theory, the identification of parts is essential to the development of concepts. The principles of gestalt, manifested in the Bauhaus as implications for structural relationships, could be used as a starting point for considering the internal contextuality of images.

Image creation is the manifestation of a message within the style of the creator –the image form. Can this creative process, manifested in the mediated product, be captured and adapted as a retrieval tool through such techniques as color indexes or the mathematics of shape? Can the internal contextual structure of an image be translated into a verbal expression of image content? To do so, it will be necessary to identify ways of linking the mathematics of pattern matching, which explores the structural unity of the image, to verbal translations of the actual process of perceiving images as mediated products.

[46] For example, see the paintings of Kenneth Nolan, "Graded Exposure," 1967; Gene Davis, "Moon Dog," 1966; or Morris Louis, "Moving In," 1961.

CHAPTER 3
Questions and Hypothesis

This research was designed to determine if there is a shared vocabulary of concepts and terms that can be used to develop a controlled vocabulary for representing complex elements in images.

This research investigated two broad questions concerning the representation of images for retrieval in the digital environment. Is there a vocabulary of natural language terms or shared concepts that can be used to describe the perceptual characteristics of images? And, if such a vocabulary can be identified, what terms (values) and/or concepts (facets) constitute the vocabulary?

Toward this end, the following research questions were posed:

1. Is there a shared vocabulary of terms that is used by subjects to represent the internal contextuality (e.g., the internal structural composition) of images?
2. Can the natural language terms used by subjects be organized into concepts?
3. If there is a vocabulary of concepts, is it shared across subject pairs?

The research employed a quasi-experimental equivalent materials design in order to collect words and phrases used by subjects to describe the physical structure of images. The test materials consisted of images available on the World Wide Web. These images were selected from three domains: abstract art (ART), satellite imagery (SAT), and photo-microscopy (MIC). Given that the natural language terms collected from subjects could be organized into a structure of related concepts, the following null hypothesis was tested on the basis of frequency of term occurrence:

> Subject pairs do not share a vocabulary of concepts for describing test images from the three domains of abstract art, satellite imagery and photo-microscopy.

Assuming that a shared vocabulary would be indicated if subjects used conceptually similar words in their descriptions, interrater reliability measures were computed to test the null hypothesis to determine whether the set of concepts was used by subjects across the three image domains.

Vocabulary was generated using a model derived from referential communication studies (Fussell & Krauss, 1989a, 1989b; Galantucci, 2005; Krauss & Fussell, 1991;

Lau & Chui, 2001; Nohara-LeClair, 2001). In this model, a subject must formulate a message that allows a listener to identify the intended referent. Various objects have been used, such as nonsense figures, landmarks, flags, and computer simulations. However, these studies frequently use nonsense figures because they do not represent conventional concepts: "in communicating them a speaker must pay careful attention to the common ground he or she shares with the message recipient" (Krauss & Fussell, 1991, p. 176). Krauss and Fussell (1991) asked subjects to describe each figure in such a way as to be able to identify the figure from a large group at a later date. Two weeks later they were asked to select figures based on someone else's descriptions. Results indicated that "coordinated communication requires interlocutors to create and use a shared body of knowledge" (p. 197). In the Nohara-LeClair (2001) study using a set of 25 flags, subjects worked in pairs separated by a visual barrier. The task for the describer was to communicate to her partner (the matcher) which flag had been selected from the set of 25. Taking turns, each subject described ten flags. Matchers were encouraged to ask for clarification or confirmation. Results indicated subjects developed shared knowledge over the course of the interactions.

For this study on vocabulary for representing internal contextuality, images from the content areas of abstract art, satellite imagery and photo-microscopy were selected for lack of readily identifiable semantic objects or scenes. Twenty-two pairs of subjects were asked to generate oral descriptions for 14 images: one subject would provide an oral description of an image that would allow her partner to produce a drawing of it. Subjects took turns performing the describe and draw tasks over the course of 14 trials, with seven descriptions per subject. Subjects were allowed time to discuss the description and drawing after each trial. The resulting vocabulary was normalized for word variance, term identification, and conceptual organization. Frequency occurrences were tabulated for every word, every term, and every concept, and for the distribution of use across the three image domains.

In order to determine if subject pairs did or did not agree on a shared vocabulary of concepts for describing the test images across the three domains, an interrater reliability test was used. Concepts were the focus of the evaluation because of the specificity occurring at the term level. Frequency counts and subject pair usage data could not be subjected to a significance test because concept distribution across domains did not form mutually exclusive categories: for every subject-pair, a concept could be used in 1, 2, or 3 domains. For this reason, the results were assessed on the basis of overall consensus and consistency in concept usage.

Frequency scores and subject pair distribution can identify concepts with the largest usage scores as well as concepts used by all subject pairs, but cannot indicate if a concept was used in all domains. To identify a shared vocabulary across subject pairs, it was necessary to determine if subject pairs agreed on concepts for usage in all three domains. This is an interrater reliability problem: Interrater reliability refers to the degree of agreement between judges asked to rate or score something, such as behavior, performance or open-response items on tests.

In this research, each subject pair was treated as a judge generating natural language descriptions in an "open-response" situation. These natural language descriptions consisted of terms (and concepts), each of which could be interpreted as a rating. Since each image has a pre-determined domain designation, each concept can be identified as being used in a specific domain when used to describe a given image. For each concept used (rated) by a subject pair (judge), a shared-ness score was calculated to reflect its use across image domains. The shared-ness scale used in this analysis is a scoring rubric similar to that used when judges are knowingly performing a rating task. It represents the distribution of concept use across image domains by a subject pair during their whole describe-draw task session. The scale was constructed as a continuum extending from 0, indicating that the concept was not used by the subject pair, to 3, indicating that the concept was used by the subject pair to describe images in all three domains.

Using this scale for concept shared-ness across domains, higher scores indicate a greater degree of cross-domain use for a concept. While the concepts shared across all three domains could be identified through analysis of frequency data, the shared-ness scale takes into account those concepts which were used in varying levels across subject pairs (e.g., a concept that was not used by all pairs but was used in all domains by one or more subject pairs).

To determine the existence of a shared vocabulary of concepts, the Spearman's rank correlation coefficient for interrater reliability was computed to determine consistency across subject pairs and is reported as the mean correlation of subject pairs agreeing on use of a concept across domains. Although most discussions of interrater reliability focus on consensus (see, for example, Cohen, 1988), Stemler (2004) contends that "consistency estimates of interrater reliability are based upon the assumption that it is not really necessary for two judges to share a common meaning of the rating scale, so long as each judge is consistent in classifying the phenomena according to his or her own definition of the scale." In this study, the question was whether subjects (judges), who are not necessarily aware of an image's domain, are assigning concepts (rating scale) to images based on a shared understanding of concepts. Agreement across subject pairs would indicate the presence of a shared vocabulary of concepts that was applicable across image domains.

Chapter 4
Methodology

4.1 Data collection

The objective of this research was to determine if there exists a shared vocabulary of concepts that can be used for verbal description of the internal contextuality of images. If there is such a set of concepts, it was theorized that it would be manifested through repeated use of the same or similar terms across the description of multiple images. Application of a controlled vocabulary of concepts and terms derived from natural language descriptions of images could be expected to contribute to the development of more effective image retrieval systems.

A group of forty-four subjects, working in randomly-assigned pairs, was asked to generate oral descriptions and corresponding drawings for fourteen randomly-ordered images under the condition that one partner generated a drawing of an image from the oral description provided by the other partner. Images were selected from three content areas: abstract art (ART), satellite imagery (SAT) and photo-microscopy (MIC). Six of the fourteen images were identical across all pairs of subjects. The other eight images were unique for each pair of subjects. The existence of a shared vocabulary for representation of the internal contextuality of images was assessed on the basis of term occurrence in the descriptions provided by subjects. Individual vocabulary words were examined to identify potential value sets; and these value sets were then analyzed to identify the concepts, if any, used across subject pairs. The process involved identification of a normalized vocabulary, which was then organized into a hierarchical faceted vocabulary of concepts and terms and subsequently analyzed to determine applicability across image domains.

4.1.1 Design

The research used a quasi-experimental equivalent materials design (Campbell & Stanley, 1963). This one-group, repeated-X design assumes equivalent samples of materials: because the effects of individual treatments are expected to be long-term, repeated treatments require the use of different, non-identical but equivalent content sets. This repeated-measures design can be represented as

$$M_a X_1 O \quad M_b X_0 O \quad M_c X_1 O \quad M_d X_0 O \quad M_e X_1 O \quad M_f X_0 O \quad \text{etc.}$$

where M indicates the specific materials used and the sample M_a, M_c, M_e is equivalent to the sample M_b, M_d, M_f. The importance of sampling equivalence across sets of materials is indicated by the following, for each subject pair:

Materials Sample A X_0 O
Materials Sample B X_1 O
Materials Sample C X_2 O

4.1.2 Materials

In their report on the Agam Program for visual skill development, Razel and Vat-Sheva (1986) point out that it is important for visual learners "to see what things really look like rather than to depend on one's predetermined verbal concepts" (p.50). To encourage research subjects to focus on the visual structure of an image rather than its conceptual content, test images were selected for lack of dependence on identifiable semantic objects or scenes. Such images would highlight the syntactic levels of Global Distribution, Local Structure and Global Composition identified in the Pyramid model presented by Jaimes and Chang (2000) and Jorgensen et al. (2001). (See Figure 1.1)

The domains from which images were selected are satellite imagery, photo-microscopy and abstract art. CBIR developers refer to satellite or microscope imagery as "non-specific images" which they consider non-application specific because such images exhibit ofness but little aboutness (Martinet, 2003). Abstract art, which is a subset of fine art, was selected because of the success of Kirsch and Kirsch (1985) in generating a structural grammar from examples of abstract art. Examples of these three domains include:[47]

Satellite imagery (SAT): deserts, cities, or bodies of water
Photo-microscopy (MIC): cell or crystal structures
Abstract art (ART): the artwork of Kandinsky or Monet

SAT was used by Cobb and Petry (1998) to develop a vocabulary that describes the structural relationships between image components. Satellite images used in this research were selected on the basis of two criteria: the absence of commonly recognizable buildings, bodies of water or coastlines and, when possible, availability as public-domain or copyright-released materials.

The organic forms represented in photo-microscopic images (MIC) were used for the description of visual structure in the architectural theory of Sullivan (Bragdon, 1922). They were also used in William's (1979, p. 20) geometric study of natural structures.

[47] Samples of each image type are available in Appendix K.

Photo-microscopic images were selected on the basis of two criteria: lack of recognizable verbal concepts, such as snowflake or insect antennae, and, when possible, availability as public-domain or copyright-released materials.

Examples of abstract art images (ART) were selected with the help of Eileen Fry, slide librarian for Indiana University's School of Fine Arts. Selections were determined on the basis of two criteria: Fry's difficulty in assigning terminology for the internal physical description of an image and copyright ownership by the IU Art Museum.

MIC and SAT images were digital photographs; ART images were digitized photographic surrogates of abstract art rendered in various media (paint, pastel, textile, etc.). Each image was no larger than 480 pixels in height and width, averaging approximately four inches by six inches, although proportions varied slightly. All images were collected from the Web, downloaded, and reproduced in the center of 8.5 x 11 white paper. Image pages were backed with card stock and placed in protective plastic sleeves.

Images were both color and black and white across all three domains. Color is a surface description concept, which has been shown empirically to be of less importance than edge description in object recognition tasks (Biederman, 1987). After reviewing research on pictures, Fleming and Levie have concluded that, in general, "the public has come to expect color as a matter of course" (1993, p. 47). But Rudnick and Elsuydam (1973) argue that there is no difference in learning between color and black and white pictures, although color may accentuate details and elicit more emotion-based description. In pilot testing for this study, describers frequently mentioned color, but this was generally an initial observation preceding the structural description of an image or image component. In the current task, subjects' drawings were done with lead pencils, rendering actual color irrelevant in the communication of internal contextuality.

Five judges evaluated the set of images for equal distribution of complexity[48] and color within the three image domains. Following Jorgensen (1995) and Rogowitz et al. (1998), evaluation used a three-point scale (1 = minimal; 2 = average; 3 = dominant) for the visual criteria of form/pattern complexity, of variation in color and of the presence of identifiable (i.e., namable) objects or places[49]. The judges were recruited from faculty and professionals with training in art history, visual design, architecture and geographic information systems. Each judge was asked to evaluate a group of 280 images (105 ART

[48] Judgment of visual complexity is highly correlated with the number of angles in the stimulus (Hochberg & McAlister, 1953).

[49] Identifiable objects or places for SAT images were determined by ability to name the place/location of the image (e.g., Washington DC or San Francisco). Identification of objects such as river, street, or building was considered acceptable.

images, 92 SAT images and 83 MIC images) for variation in color and complexity using the three-point scale, eliminating any images with identifiable objects. Each judge was then asked to identify a set of exemplar images for each of the three image domains. The number of exemplar images selected was left to the discretion of each judge.[50]

Judges' ratings were evaluated for distribution both within and between domains based on the criteria of color-range, form/pattern complexity, and object recognition. Based on judges' ratings, images with the least amount of agreement were eliminated, as were any images judged to have recognizable objects. A subset of 235 images was selected based on judges' ratings for color variation and complexity and the need for equal distribution across domains. Two images with the lowest color-complexity rating were then selected from each domain to serve as training images. Two more images were then selected from each domain to serve as standard images[51] that would be described by all pairs of subjects. These six standard images were identified by overlap in the exemplars selected by each judge. There was exemplar agreement by three judges on one ART image and two SAT images. The remaining three exemplars were selected from those on which two judges had agreed. When there was more than one image on which two judges had agreed, the image with the lowest color-complexity rating was selected. Using randomly-generated numbers, 176 images (59 ART, 58 MIC and 59 SAT) were then selected from the remaining set of images to provide a total of 182 (including the six standard images).

Twenty-two sets of images were constructed consisting of fourteen images each: the six standard images and two or three images from each pool of ART, SAT and MIC images. The six standard images were presented to all pairs of subjects in three different combinations:

Table 4.1. *Combinations of standard images for distribution in image sets.*

Standard combination 1a	Standard combination 1b	Standard combination 2a	Standard combination 2b	Standard combination 3a	Standard combination 3b
ART1	ART2	ART1	ART2	ART1	ART2
MIC1	MIC2	MIC1	MIC2	MIC2	MIC1
SAT1	SAT2	SAT2	SAT1	SAT1	SAT2

Note: Domain numbers distinguish between two standard images from the same domain (e.g., SAT1 and SAT2).

50 A copy of the instructions for judges is available in Appendix I. Sample data from an evaluation by a single judge is available in Appendix J.

51 Standard images are reproduced in Appendix K.

Each combination of standard images was assigned a pair number (i.e., 1 to 22) and an a or b designation since there would be a set of materials for each subject in a pair (i.e., 44 subsets). The three sets of standard image combinations were assigned in sequential order.

The eight non-standard images for each material set were assigned based on random selection from the three domain groups. This was done by first randomly selecting a domain to begin the selection process, then randomly selecting an image from within the domain. The process was repeated until all 22 sets of materials contained eight non-standard images. Because there were three domains, this resulted in an uneven distribution of domain images in each set; however, each set had at least two images from each domain, assuring that each subset would include at least one image from each domain.

When combined with the three standard images, this provided a subset of seven images for each subject. There are twenty-six possible order-of-delivery permutations for three standard and four unique images in any subset of seven images. Eight of these order permutations were selected based on a relatively even distribution of standard images across each position:

Table 4.2. *Distribution of standard images across delivery positions in the image sets.*

	1st Position	2nd Position	3rd Position	4th Position	5th Position	6th Position	7th Position
Possible ordering set							
1	x	s	x	s	x	s	x
2	x	x	s	x	x	s	s
3	s	x	x	s	x	s	x
4	s	x	s	x	s	x	x
5	s	x	s	s	x	x	x
6	x	s	x	x	s	x	s
7	x	s	x	x	x	s	s
8	x	s	x	s	s	x	x
Number of occurrences of standard images by position	3	4	3	4	3	4	3

Note: A standard image is represented by s and a non-standard image is represented by x. Columns indicate the order in which images were presented to each subject.

This ensured that the standard images were as evenly distributed as possible in the order of presentation for description, from first to seventh, thereby accounting for learning effects due to shared communication. Images in each subset were randomly assigned to a position in the eight possible orderings.

4.1.3 Subjects

Forty-four volunteers were recruited for participation in the study. To solicit volunteers, advertisements[52] were posted on the campus of Indiana University (in the Herman B. Wells Library as well as other campus buildings) and in the Monroe County (Indiana) Public Library. An announcement was also posted to the discussion list for Indiana University's School of Library and Information Science (SLIS). Two eligibility requirements were clearly stated: volunteers must be native English speakers and volunteers could not have taken any college-level courses in art, architecture, photography, or geographic mapping. Of the 44 subjects, 40 were from the United States, one was from Australia and three were of British background. The prohibition against specific coursework was included because it was assumed that subjects who had studied in any of these areas might have acquired skills related to composition and spatial thinking, which could have influenced their descriptive vocabulary. The subjects ranged in age from 18 to 56, with an average age of 30. There were 14 men and 30 women.

4.1.4 Procedure

Subjects were randomly assigned to pairs. Experimental sessions with individual pairs were held in the SLIS Usability Lab or, in five cases where the Usability Lab was not available, in a quiet conference room. Before beginning the session, each subject was asked to read and sign a Human Subjects consent form[53] and then complete a short profile questionnaire.[54] Each pair of subjects was randomly assigned one of the image sets, and each subject was randomly assigned one subset of seven images. All sessions were recorded on audio-tape.

The same presentation format was followed for all pairs. Before beginning the describe-draw task, subjects read an instruction script describing the general purpose

---

[52] A copy of the advertisement is available in Appendix L.

[53] A copy of the Human Subjects Consent Form is available in Appendix M.

[54] A copy of the profile questionnaire is available in Appendix N.

of the study and the procedures that would be followed.[55] The researcher answered any questions asked by the subjects. In every case, this was a re-statement of what they had read in the instruction script. A small sign was made visible to both subjects that listed the only dialog allowed during the actual task: Please repeat, Please slow down, Wait-OK-ready.

To complete the describe-draw task, the subjects in a subject pair were asked to sit on either side of a visual barrier. After a training example that was identical across all pairs, Subject 1 was given an image in a book-like folder such that Subject 2 could not see the image. The first describer was predetermined by the random assignment of the image set. Subject 1 described the image for Subject 2; and Subject 2 used a lead pencil to draw the described image on a blank, A-size paper in a second book-like folder. Upon completion of the drawing, the researcher prompted the subjects to discuss the results of the task by asking leading questions as set forth in the instruction script. Subjects then reversed the describe/draw roles and Subject 2 described one picture for Subject 1 to draw (end of Round 1). During the task, the images to be described were uncovered one at a time by the researcher.

After each subject had completed an initial description, subjects reversed roles for two descriptions by Subject 1 (begin Round 2). They then reversed roles for two descriptions by Subject 2 (end of Round 2). This process of alternating roles for two descriptions by each subject was repeated for Rounds 3 and 4. Each subject completed four rounds of the describe-draw task, alternating between drawing and describing. Excluding the training example, each subject performed a total of seven descriptions and seven drawings. Following completion of Round 2, subjects were allowed to take a 10-minute break; they also had the option to take a second break between Rounds 3 and 4.

During each describe-draw task, the researcher made brief notes regarding subject fatigue or frustration, verbalizations that were difficult to hear or understand, and drawing actions that did not match verbal instructions. At the end of discussion of the last describe-draw task, subjects were allowed to ask questions about the study. Each subject was paid $10.

4.2 Data Processing

Processing the verbal description data involved three major steps: transcribing, faceting, and tabulating. Of the 22 pairs of subjects, one pair was eliminated due to

[55] A copy of the instruction script is available in Appendix O.

audiotape recorder malfunction, leaving 21 pairs or 42 individual subjects from whom data was collected.

In the following report and discussion, individual words and direct quotes from subjects are indicated by use of plain font; stop words are indicated by use of plain font in SMALL-CAPS; and conceptual antonyms and superordinate conceptual organizers are indicated by use of plain font *italicized*. When a word is determined to be a facet term, its initial letter is capitalized and the term is enclosed in angle brackets: <Plain> or *<Italicized>*. When a term is identified as a top-level facet, it is indicated by use of all capitals enclosed in angle brackets: <CAPITALIZED>. The following syntax is used when attributing examples from the data: (Pair number [1-22], decimal point, subject number [1 or 2], #xxx [picture], hyphen, 1st-14th [description order]). For example, (13.1#120-7th) indicates pair 13, subject 1 of 2, picture #120, 7th out of 14 descriptions. No attribution is made when words or phrases are quoted that were used across multiple subjects. Comments made by a subject during discussion following a describe-draw task are indicated as (13.1#120-sum); and comments made in the final summary period at the end of the session are reported as (13.1-sum).

4.2.1 Transcribing the oral descriptions

Transcription of the audiotapes involved listening to each tape three times[56]. The first pass was for direct typing; and the subsequent two passes were for error correction and checking code consistency. Transcribing the oral descriptions to written words was an iterative process for several reasons. First was the question of utterances or partial words and whether or not they should be included. An example of this would be transcribing uh on the left edge of the bor ... oh on the right edge of that border (20.2#248-14th). Here, uh and oh are utterances that serve as verbal placeholders while the subject is thinking. They were not transcribed. The utterance bor is subsequently proven to be the word border when the phrase is repeated, but to transcribe its first utterance as a word would be an assumption. Therefore, partial words were not transcribed. Repetitions were always transcribed whether or not they consisted of exactly the same words.

Secondly, the representation of spoken words in written notation is never straightforward, in particular with numeric amounts. For example, should the spoken word half be represented as a word or as the numeral ½? For expediency in transcription, all numeric values were indicated by their numeric figures. During later stages in data

[56] A sample transcription is available in Appendix P.

processing, fractions proved problematic due to varied sort techniques embedded in the software applications. Therefore, numeric representation of fractions was changed to #b# (e.g., 3b5) since some software could not read or appropriately sort the slash mark.

Another transcribing problem was the use of the hyphen and the apostrophe. These characters caused problems with the various software sorting functions: one ignored them and another treated them as a space, creating two words where there was only one. During this first transcription stage of the data processing, hyphens were retained whenever the transcriber deemed it appropriate to indicate the spoken words. However, at a later stage in data processing, every hyphen was examined for appropriateness of conceptual use and was retained as either a single string (e.g., off-white, bull's-eye, see-through) or separated into two (or more) words (e.g., color-scheme became color and scheme). In addition, contractions were later expanded into their full representation (e.g., can't became can not).

No other punctuation marks, such as commas or periods, were included in the transcription as there was no clear way to distinguish between them. In fact, few subjects spoke in complete sentences. However, it became evident that, in some instances during the transcribing process, there was a need to represent contextual meaning or nuance. This was done by appending a hyphen and a bracketed comment. For example, the word right could be used to mean *on the right*, *next to*, or *in the correct place*. These were transcribed as right, right-(just), and right-(correct). Similarly, the word over could be used in the context to mean over as in *above*, over as in *on top of*, or over as in *to the side*. These were transcribed as over, over-(2), over-(3), and a transcription key was maintained for later reference. In the final stages of data processing, when such variations were retained they were indicated as over, over2, and over3, and a scope note was added for clarification. When variant forms of these words were encountered, the numbering system was maintained, albeit awkwardly, to differentiate the various uses of the word: for example: over2ly, block2s, fat2ish.

After listening to several descriptions, the transcriber began to use shorthand codes for certain phrases that occurred regularly, e.g., kind-of (KO), sort-of (SO), kind-of-looks-like (KOLL), sort-of-looks-like (SOLL), and other permutations of those words. These codes were subsequently expanded during later data processing stages.

Transcriptions were also made of the subjects' summary evaluations at the end of each draw-describe task and at the end of the whole session. These were kept in separate coordinated files.

In summary, the following transcription rules, maintained in the transcriber key, were devised and applied through the iterative process of listening to the tapes several times:

1. No utterances (uh, oh).
2. No partial words.
3. Numbers are numerals.
4. Fractions are numerals and slash is b (3b5).
5. Hyphenate words at your discretion (off-white, bull's-eye).
6. Use contractions when they are verbatim usage.
7. No punctuation.
8. Capture context when appropriate (right-(just)), add new entries to transcriber key.
9. Use shorthand for repetitive phrases (SO, KOLL), add new entries to transcriber key.

4.2.2 Building the faceted controlled vocabulary

Words and phrases used in the oral descriptions from subjects[57] were extracted to identify a vocabulary for visual structure based on natural language. Transcripts for each subject were analyzed for term identification (Batty, 1989); and these terms were used for tabulation of term frequencies as established by research on category norms (Battig & Montague, 1969; Hunt & Hodge, 1971).

The process of building the controlled vocabulary and facet creation had three iterative phases: identification of stop words, syntactic normalization, and semantic normalization. Initially, a list was created of all unique words used by subjects as identified by a concordance-building software application. Definitions were then added based on contextual use in subject descriptions: Because this vocabulary is concerned with the physical aspects of images and is not interpretative, definitions of words used by subjects had to make sense within the context of physical structure.

For the purpose of identifying a concise and focused vocabulary for images from the verbal transcripts of subjects, the concept of stop words was applied but with a slight variation. Stop words are commonly used words that are deemed irrelevant for searching and indexing by search engines; because their frequency can interfere with search and index statistics, they were removed from consideration. The frequency of some words which carry little description of image physicality, such as pronouns, justified their elimination from the tabular analysis of vocabulary. However, stop word lists generally include some words that are descriptive of the physicality of images, such as the numbers

[57] The list of all natural language and phrases generated by subjects is available in Appendix A.

one through ten (which in the current context could indicate the number of objects appearing in an image) or the articles *a* and *an* (which would indicate that the number of objects appearing in the image is just one). Therefore, because some stop words can be meaningful in the representation of internal contextuality, traditional lists of stop words were used as a starting point in the current context. Stop words added to traditional lists included draw commands, describe-draw process dialog, indexical pointers, and verbal placeholders.[58] Differences in use of any given word had to be carefully examined during transcription, when identifying stop words, and during normalization. For example, paper is a stop word in the sense of go-all-the-way-down-the-paper but is not a stop word when it-looks-like-curled-paper.

Syntactic normalization involved basic concept identification and the grammatical analysis of variant forms of words with the same root in order to develop stemming guidelines. For example, rectangle, rectangular, and rectangularish were all reduced to the term rectangle by the guideline that noun forms take precedence. Another example was the initial reduction of all plural forms to singular (e.g., leaves to leaf, antennae to antenna, and feet to foot). The result of this process was a list of normalized words.[59]

Semantic normalization followed the guidelines for faceted thesaurus construction described by Batty (1989). The researcher identified synonym sets (synonyms and near-synonyms), determined an authorized value (isolate) to represent the set, grouped the values into concepts (facets), and created a hierarchical structure for each facet. For example, the synonyms beige, sandy, and tan have the value brown in the <PROPERTY> facet of <Color>, and brown appears after orange in the hierarchical structure. The faceting process had two results: an alphabetical list of all authorized terms with Use For (UF) references and scope notes (SN) (see Appendix D); and, a faceted controlled vocabulary of authorized terms arranged in a logical order and including scope notes (SN) (see Appendix E).

Although the process of vocabulary creation initially followed the order of stop word identification, syntactic normalization, and semantic normalization, the actual process was, in fact, iterative due to constant comparisons and referrals to the words in context. This led to refinements in the vocabulary as each subsequent word in the list was examined. For example, the first encounter with the word over was in the context of *to the side*, while subsequent encounters produced over2 meaning *above* and over3 meaning *in front of*.

---

[58] The stop word list is available in Appendix B.

[59] The list of words normalized by variant form is available in Appendix C.

The products of this process were a natural language word list that contains all unique words used by subjects; a faceted vocabulary; and a term list that includes all authorized terms with Used For (UF) references indicating unauthorized words.[60] The faceted vocabulary contains all authorized terms that resulted from the normalization process. Each term includes a scope note (SN) when needed for clarification. Definitions of words as they were used in the context of the image description task often formed the basis for scope notes that were added to clarify the controlled vocabulary.

4.2.3 Word List

Only the actual words that occurred during the description task are included in the word list which means that all possible inflected forms of a word or possible definitions are not included. Describer-invented words do appear, such as orientated, intestiney, amorphicky, and straightly. Concordance software (KWIC Concordance, version 4.7, 2004, Saturo Tsukamoto)[61] was used to generate alphabetical word lists from the transcriptions of the description tasks. The concordance output included every unique string of characters, so that snake, snakes, snaky and snaking all appear as individual items. These lists were copied to an MS Excel spreadsheet of unique natural language words, one word per row. Duplicate strings were skipped when they appeared in each subsequent description word list, until the spreadsheet contained a complete list of all the unique natural language strings used by subjects.

Definitions for each unique entry were then derived from a dictionary of current American English (Abate, 2002) and dictionaries of American slang (Ayto & Simpson, 2005; Chapman, 1987). Because the vocabulary was collected within the context of orally describing the visual structure of images, definition selection was not based on purely semantic relationships or details. If a controlled vocabulary could be identified, it should be one that identified visual similarities and relationships and would not be exhaustive regarding semantic relationships and details. For example, the word baby was used in the sense of small and not in reference to an infant; and the word clock was used in the context of position and not in reference to the actual object. A few words did not appear in the dictionaries, some clearly being verbal inventions by the subject: for example,

---

[60] At some point, all authorized terms should include a notational reference to the appropriate location in the faceted vocabulary, but a citation order and notation are yet to be determined as they should be based on the intended user of the vocabulary. At this stage in the vocabulary development, it is not clear whether the user would be a searcher, developer or a computer program that would use the vocabulary in order to aid the searcher.

[61] See http://www.chs.nihon-u.ac.jp/eng-dpt/twkamoto/kwic.e.html.

squibble (referring to a squiggly line), teench (meaning a small amount), or trippy (slang for psychedelic). Since language usage is not fixed but constantly evolving in its common use, there is a time lag in dictionary acceptance and publication of new words and phrases. In addition, there are many words and phrases, both accepted and colloquial or idiomatic, that are specific to specialized domains.[62] In these cases a definition was generated based on context, similar terminology, and domain specific dictionaries.

Homographs are different words with the same spelling. Semantic disambiguation was performed during the transcribing process, when a word with multiple meanings could be evaluated for its meaning in context. When homographs were identified through contextual evaluation, each concept was given its own entry in the natural language word list (see Appendix A) and was differentiated by an appended numeric code: for example, block (a group of buildings surrounded by four streets) and block2 (an obstruction).[63] When variant forms of a homograph occurred, the numeric code was retained at the root position to facilitate grouping during alphabetic sorting (e.g., fat2 and fat2est). Scope notes were added to differentiate the different meanings of the homographs.

Abbreviations were retained rather than expanded (e.g., DNA). Contractions were problematic because the concordance software which ignored the apostrophe and created two entries (e.g., can and t). The decision was made to expand contractions to their full verbal intent (e.g., can and not) rather than to maintain them as a single string (e.g., cant).

The use of hyphens in the transcripts presented another unique entry problem with the concordance software. Just as with the apostrophe, the concordance application treated the hyphen as a space. During the transcription process, artificially imposed hyphens were added by the transcriber to retain the grouping of some words for evaluation during the normalization process. This guaranteed the examination of the normal usage of the hyphen. Furthermore, multi-word concepts (phrases) would not be mechanistically separated into multiple concepts. For example, three-dimensional is a hyphenated concept in the dictionary and high-school appears as a single concept phrase referring to the entity and not the adjective *high*, as opposed to *low*, modifying *school*.

---

[62] For example, *Merriam-Webster's Medical Dictionary* (1995) includes the words castor-bean, xylose, yohimbe, zoonosis (a disease communicated from animals to humans), and the chemical elements thulium, praseodymium, protactinium. None of those words appear in the *The Oxford American dictionary of current English* (Abate, 2002).

[63] The sequential number was assigned by the order in which the various meanings were encountered in the process of building the word list. After tabulation, the numbers were re-assigned based on frequency of use.

During syntactic normalization, phrases and hyphenated words were retained as a single concept if they appeared in the dictionary (Abate, 2002) as a single entry. For interim tabulation purposes, spaces and hyphens were eliminated so these concepts became single strings (e.g., threedimensional and highschool), as required by the concordance application. When normalization and tabulation were completed, these terms were expanded to three-dimensional and high-school in both the term list and the faceted vocabulary.

During the transcription process, hyphens were liberally applied to create phrases that were not dictionary based but sometimes quite expansive. A simple example is the phrase the-yinyang-symbol (P11.2#133). Both yin and yang appear in the dictionary as separate concepts; but the contextual reference was to *the yin yang symbol* and not the individual philosophical principles. For this reason, the-yinyang-symbol was retained as a single concept. In further normalization iterations, the was eliminated as a stop word and symbol was identified as a facet concept, thus subsuming yinyang as an isolate. A more complex example is like-a-little-kid-would-draw-a-snake (15.1#70-8th) and like-a-little-kid-making-pictures-of-a-campfire (2.2#250-5th). Anticipating the need to retain context, the expanded phrases were created during transcription and then analyzed into concepts during iterations of the normalization processes. Eventually, the only words left that could not be easily accommodated as distinct concepts were kid-would-draw and kid-making-pictures. Both references are to a style of picture making, so the relevant concept in the phrases was transliterated to children's-art. This resolution was in line with the elimination of stop words but clarified the contextual use of the word kid. Notation of the transliteration was made in the scope note.

In summary, the guidelines for creating the word list were:

1. Abbreviations are retained.
2. Contractions are expanded.
3. Homographs are appended with sequential reference numbers which have scope notes to differentiate them.
4. Phrases and hyphenated words are retained as single concepts.
5. Transliterations are noted in scope notes.

4.2.4 Stop words

Stop words[64] are common words that are often used to modify other words but that carry no inherent meaning in and of themselves – words such as adverbs, conjunctions, articles, prepositions, or forms of the verb *to be*. In general, stop word lists

[64] The stop word list is available in Appendix B.

for Internet search engines are relatively short: for example, Google's stop word list only contains approximately 20 words.[65] These lists frequently include words that could be descriptive of the internal contextuality of images (e.g., *a, an, in, under*, and *within*). The stop word list used by van Rijsbergen (1979) is considerably longer, consisting of approximately 270 words, but it, too, includes words that could be descriptive of images (e.g., *around, below, behind,* and *fire*).

One of the first steps in the normalization process was to review existing lists of stop words in order to devise guidelines for decisions regarding stop words. Stop words were subsequently identified throughout the various iterations of the normalization process and ultimately included draw commands, describe-draw process dialog, indexical pointers, and verbal placeholders in addition to more conventional stop words such as pronouns.

Draw commands are those words and phrases used by subjects verbalized in the describe-draw task which were relevant only to the physical task of drawing an image. Draw commands included action verbs (CHANGE, MAKE, and DRAW), state of being words (BELONGS, CONSISTS, and HAS), and time related words (AWHILE, NOW, and STILL). For example, the following words and phrases produced by pair 16 were identified as stop words, but the context in which each occurred was carefully reviewed in the various iterations of normalization:

> PUT, COME, GO, GIVE, WHERE YOU STOPPED, THAT YOU JUST DREW, YOU STARTED WITH, WHERE YOU STARTED, YOU DREW, DRAW WHAT YOU IMAGINE WOULD BE, YOU JUST MADE, IF YOU'D LIKE TO ENVISION, SO YOU CAN SEE, WORKING YOUR WAY, YOU'VE GOT

The draw commands from pair 16 exemplify the fact that identification of stop words is necessarily context dependent for homographs and syntactically similar words. JUST is a stop word when it is used in the context of time and is represented as just to distinguish it from just meaning *exactly*, as in a-wave-just-like-the-Ocean-Spray-bottle (3.2#133-10th), or barely, as in just3-to-the-left. Similarly, imagine refers to an action on the part of the drawer and is a draw command in contrast to its syntactically similar use in the phrase an-imaginary-line (P7.1#115). The pair 16 draw commands also exemplify the use of a non-stop word in a stop word context, such as STOPPED. STOP is identified as a stop word in the draw-command context referring to the action of a drawer; but it is not a stop word when stop means *end*, as in end-of-the-line. However, since it is not a homograph, STOP was simply eliminated during the normalization process

[65] See <http://www.abcseo.com/seo-book/stop-words.htm>, accessed on July 4, 2006.

when it occurred in the draw command context.[66]

Describe-draw process dialog is related to draw commands in that it includes the verbal communication that was allowed as part of the describe-draw process (i.e., PLEASE REPEAT, PLEASE SLOW DOWN, WAIT, OK, READY). In addition, verbal placeholders, including a describer's dialog with herself, were eliminated as stop word phrases. These words and expressions are exemplified by the following excerpts from pair 16:

I'M SORRY
THIS IS GONNA TAKE MORE THAN 5 MINUTES
LET'S SEE THE BEST WAY I CAN DESCRIBE IT
OH-WOW

Indexical pointers were also categorized as stop words. Indexical pointers are words that refer to the object being described and include pronouns (e.g., IT, HE, and THIS) as well as indefinite references to objects (e.g., THING and CERTAIN), ordinals (e.g., FIRST, SECOND, and FIFTH), indefinite references to positional state (e.g., LAST and PREVIOUS), and references to the image being described (e.g., CANVAS, PICTURE, and SCENE).

Conjunctions, prepositions and adverbs were evaluated on a case-by-case basis. Conjunctions such as AND, BUT and OR were easily identified as stop words, but some conjunctions were deemed to be descriptive (e.g., as-if implying similar). Most prepositions, which are generally followed by a noun, a noun phrase or a pronoun, are descriptive (e.g., in, on, around, and between) but a few were generally used as stop words (e.g., FOR and OF). The use of TO was easily categorized as a stop word in its infinitive form, but its use indicating direction (e.g., to-the-left) was more problematic, as was the use of its counterpart FROM. Both were ultimately identified as draw commands and were, therefore, designated as stop words.

Adverbs were the most difficult to categorize as traditional controlled vocabularies are generally limited to nouns although they may sometimes include adjectival forms. Adverbs modify verbs and help to communicate how, when, where, and to what extent something is done. Most adverbs were held for evaluation until the syntactic normalization process when they could be converted to adjectival form. Adverbs that initially emerged as stop words because they bore little descriptive information or were related to drawing commands included ELSE, AS, HERE, ASIDE-FROM and OTHERWISE.

The guidelines for designating stop words were:

1. Allowable process dialog (please repeat, etc.).

[66] Stop words were not tabulated, but a gross number could be calculated by comparing the natural language words and the word list.

2. All words associated with verbal placeholders or self-dialog.
3. All words associated with the draw command context even if they are not normally stop words, including time references.
4. Indexical pointers.
5. Evaluate each conjunction, preposition, and adverb for its descriptive value.
6. Iteratively review all designations based on syntactic and semantic normalization.

4.2.5 Syntactic normalization

Because stop word identification and syntactic normalization are closely related activities, they were not executed in a sequential fashion after initial application of the first three stop word guidelines. The goal of syntactic normalization was to reduce the number of inflected forms of any given word without compromising its descriptive value. Each word in the controlled vocabulary should be a noun, adjective or preposition, with the noun being the form of choice. Each word and its variants were closely evaluated in context to determine stem relationships and to designate associated words in the word list. Associated words were often identified by creating a restatement of the descriptive verbalization using a noun, adjective or prepositional form. If a restatement did not change the meaning of a word in the context of its original verbal description, the form used in the restatement was deemed acceptable. For example, the words circle and circular represent an entity and an attribute respectively. Circular is defined as "having the form of a circle" (Abate, 2002) and could be restated by the phrase *like a circle*. For example, a tumbleweed-that's-circular (4.2#9-5th) could be restated as a tumbleweed that's *like a circle* without a change in the intended meaning. Similarly, the frequent use of statements such as kind-of-circular are easily re-stated as *like a circle*. Therefore, circular was syntactically normalized as the noun form circle.

Variants that were alphabetically consecutive or nearly consecutive were combined as a single entry in the word list with the preferred form appearing first and all other variants following in alphabetical order. Variant forms that did not occur in the descriptions were not included in the word list. However, describer created forms were included, such as amorphicky (P21.2#39-7th). Inflected forms that were routinely normalized included plurals of nouns, past tense, and the past and present participles of verbs as well as comparative and superlative forms of adjectives and adverbs. Words without inflected forms were simply retained.

The easiest words to normalize based upon their root words were the plural forms of nouns and noun phrases, which were always converted to the singular; For example, leaves was always normalized as leaf. Word variants referring to the same entity or concept were normalized to a single word as long as the intended meaning was retained. For example, gasoline became gas (i.e., the flammable liquid) and fries became French-fries. When there was more than one noun form from which to select, the simplest form took precedence. For example, given the variant forms smoke, smokey, smokiness, the normalized form selected was the noun smoke. Proper nouns were collected and set aside, as were any remaining pronouns, such as both, another, and everything. Proper nouns became isolates in the faceted structure, and pronouns were evaluated during semantic normalization.

Verbs were taken to their noun or adjectival form whenever possible. At this stage of normalization, the selection of the noun or adjective form of a verb was based solely on the variants that actually occurred in the descriptions. Many verbs occurred in the past participle form (e.g., deviated, distorted, shattered, dissected, and frayed) or, less frequently, in the present participle form (e.g., increasing). Generally, the past participle took precedence when there was a choice between verb forms. For example, from flare, flared, and flares, the normalized form became flared, just as tapered became the normalized form for taper, tapered, and tapers. The present participle was selected if the past participle did not occur in the description: meandering was selected as the normalized form for meanders and wandering for wanders. When the inflected forms of a verb included an adjective, the adjectival form took precedence over the participle forms: For example, for the related words blur, blurred, blurring, and blurry, the normalized form was designated as blurry. Verbs normalized as nouns included stack for stacked and angle for angling. Decisions were generally based on the restatement of the original description and evaluation of any potential change in meaning.

Some verbs that first appeared to be draw commands were subsequently identified as descriptive. For example, in the phrase where-to-draw-the-lines-separating-these-segments (5.2#89-5th), the words retained after stop word elimination were lines, separating, and segments, which were meaningful to the physical structure of the image. Since separate, separated, and separating all occurred in the descriptions, the normalized form became separated.

When draw command verbs occurred with prepositions, the verb was identified as a stop word but the preposition was generally retained as descriptive (e.g., come-down, come-in, come-into, and come-out). Adjectival suffixes indicating approximation, such as *-y* and *-ish*, were combined under their root word. For example, spikey became spike and squarish became square. Comparative and superlative adjectives were grouped together (e.g., good, better, and best; dark, darker, and darkest) and held for later semantic evaluation. Adverbs were generally left alone at this stage of normalization except in the rare case when an adjectival form existed in the description (e.g., slight and slightly).

The guidelines for normalizing variant forms were:

1. Noun forms take precedence, the simplest one if more than one is available.
2. Singular takes precedence over plural.
3. Verbs go to the noun or adjectival form, and noun takes precedence.
4. Adjectival forms take precedence over participles.
5. Past participle takes precedence over present participle.
6. Adjectives with approximation suffixes (-ish, -y, etc.) go to the root word.
7. Comparative and superlative adjectives are grouped.

4.2.6 Faceting and semantic normalization

The goal of the faceting process was a hierarchically structured controlled vocabulary of terms grounded in the words used by subjects to describe images. Faceting collects terms into groups with similar meaning, identifies or assigns a label for the concept group, and then creates an internal ordering (or order in array) for the terms within each facet. Semantic normalization refers to the identification of a single term to be used for a single concept, establishing Used For (UF) relationships between the authorized term and its synonyms and near-synonyms. In this process, a near-synonym is not necessarily a substitute for a word, but another way of saying what is connoted by a word. For example, a wave is an object but is defined as “undulating” or “alternate contrary curves” (Abate, 2002) and can be mapped to its near-synonym ripple, which is defined as having a “wavy appearance” (Abate, 2002).

Facet labels, whether terms or phrases, were not necessarily selected from the terms that resulted from the semantic normalization, but were sometimes introduced by the researcher in order to properly represent the concept and are referred to as superordinate conceptual organizers. For example, the word *vegetable* did not occur in the descriptions provided by subjects; rather, it is a superordinate concept that represents a category of similar entities (e.g., corn, pea, cauliflower, and onion, etc.) that did occur in the descriptions. The facet label *<Vegetable>* was therefore introduced as a conceptual organizer, which is indicated by the use of italics. Additionally, it is enclosed in angle brackets (<>) to indicate that it is a facet label in the controlled vocabulary. In some cases, a variant form of a term was used as the facet label (e.g., *<Sharpness>);* in other cases, a label was created (e.g., *<Movement-reflecting-growth>).* Where necessary, conceptual antonyms were required to ensure semantic integrity of a facet: For example, in the facet <*Dryness*>, the logic of *dry* is completed by its semantic antonym *wet*.

The last step in faceting was the determination of hierarchical structure within each facet. All of the terms in a given facet are arranged in an order that would potentially be of assistance to the user of the vocabulary. For example, in the facet <Blood-vessel>, terms are ordered using the logic of large to small: artery to vein to capillary.

In this dual process of semantic normalization and faceting, words were initially grouped according to similarity of meaning or closeness in concept. For example, all the color names were grouped together, as were all fruits, vegetables, animals, figures, numbers, places, etc. This process of grouping similar words continued until all individual words had been assigned to a group. A simple example is illustrated by the words church, museum, aquarium, and planetarium. These words were grouped because they were types of cultural buildings that were used in subject descriptions. The word list contained references to various types of buildings, including skyscraper, parking-garage, tepee, and trailer; but the grouping church, museum, aquarium, and planetarium demonstrated semantic coherence in that each of these words pointed to a kind of cultural building. All of these terms were identified as isolates nested under the facet <*Cultural-building*>. However, reference was made to a specific museum (e.g., Louvre); and, because proper nouns are organized as isolates nested under the appropriate concept, <Museum> was identified as a subfacet nested under <*Cultural-building*> with the proper noun Louvre as its isolate.

The faceting process was by no means straightforward. The evaluation of words for concept grouping often required returning to the raw transcription to discern, if not infer, the intent of a word's usage. For example, it would initially seem that the words stem and stemming could be semantically normalized to stem; however, on examination of descriptions, both terms were retained as separate concepts with stem indicating a part of a plant and stemming as a figural characteristic in the sense of branching or separating. Ultimately, stem was retained as a term in the <*Plant-body-part*> facet and stemming was identified as a synonym for branch in the <Angle> facet.

Sometimes redundancy in the definitions of words was enough to semantically normalize similar words. For example, because avenue and drive were defined as types of a street or road, both were semantically normalized to the term street. The selection of street as the authorized term rather than either avenue or drive was based on a guideline specifying selection of the least ambiguous or most common word; and, in this case, street was deemed the most common. An example of selecting the least ambiguous word occurs with tiny and baby, two words for a *very small thing*. Tiny was selected as the authorized term because it was less ambiguous than baby, which could also refer to an infant.

Some words required careful analysis due to subtleties of definition and usage. A good example is the word right used in contexts such as on-the-right-hand-side, on-the-right-side, and on-the-right. The word right can be used to reference direction, position, or specification; but the distinction was deemed trivial because there is too much overlap in the descriptions.. For example, when right is used as a direction, it can be considered a drawing command and therefore a stop word, yet it clearly indicates a relative position within an image. Similarly, when the word is used to specify a particular side or object, it can be considered an indexical pointer and therefore a stop word; yet, again, it indicates a relative position within the image. Therefore, the word right became an isolate under <Direction>. And, because the related phrase right-hand does not differentiate between spatial location and direction, it was normalized to right. Right-side, however, does indicate a more specific reference to the concept of spatial location, and the phrase was normalized as two terms (i.e., right and side). Right, meaning *exactly* or *precisely*, was normalized to exactly.

In some cases, semantic normalization changed what had been accomplished during syntactic normalization. For example, following syntactic guidelines, symmetry and symmetrical were normalized with the noun form taking precedence. However, as the <Attribute> facet evolved, symmetrical was determined to belong in the subfacet *<Regularity>*. Therefore, the authorized term became the adjective form symmetrical, following the guideline that semantic normalization would take precedence over syntactic normalization. This situation was also prevalent in the semantic normalization of certain verbs that had been taken to the past participle during syntactic normalization but were changed to the present participle during facet formation. For example, the <Action> facet proved to be both less awkward and more descriptive using present participles (e.g., exploding, gushing, and spraying) rather than past participles (e.g., exploded, gushed and sprayed).

Almost all adverbs were eliminated from the controlled vocabulary during semantic normalization because they could be grouped with closely associated, authorized adjectives. Examples include divided UF apart, similar UF somewhat, and main UF especially. A notable exception to this process was in the <Degree> facet which, by its very nature, remained predominantly adverbial with isolates such as approximately, barely, and very.

Verbal redundancies were originally identified during transcription by the creation of hyphenated phrases and were reviewed during the process of semantic normalization. An example of a verbal redundancy is a-helipad-a-place-where-a-helicopter-lands (22.1#133-3rd). In semantic normalization, where-a-helicopter-lands became an occurrence of helipad since it is the definition of a *helipad* and the term helipad had already been used. The word *helicopter* was not included as a term in the controlled vocabulary because the reference was not to the object *helicopter* but to the place helipad.

Comparative and superlative forms of adjectives were semantically normalized to their root form following the guideline that the comparative would be indicated by appending -x2 to the root and the superlative appending -x3. This holds true even if the root form is different than the other forms. For example, dark, darker, and darkest becomes dark, darkx2, and darkx3; and similarly, many, more, and most becomes many, manyx2, and manyx3.

In some cases during the faceting process, facets and isolates were added in order to establish the semantic coherence of the facet. For example, the <Emphasis> facet grouped the terms highlight and lowlight; but semantic coherence required that there be a state of *no highlights or lowlights*. Therefore, the term *no-light* was introduced and added as a conceptual antonym.

Finally, once the unauthorized words were identified and the authorized terms were grouped into a hierarchy of facets, subfacets and isolates, the isolates within each facet were arranged in a logical order. The primary guideline for ordering in array was to organize the set of facets or subfacets by what is most important in the domain. At the topmost level, this guideline produced the following order: <OBJECT>, <PLACE>, <PROPERTY>, <SPATIAL-LOCATION>. Within these main facets, sub-facets and isolates were ordered based on which of the following strategies was most appropriate to the concept group: big to little (mold to microbe); general to specific or whole to part (skeleton to tibia); top to bottom (thorn to root); and most common to least common (rooster to toucan). Alphabetical order was only used for certain groups of proper nouns, such as the isolates nested under *<Computer-game>* or *<Municipality>.*

Guidelines for the semantic normalization and faceting processes were:

1. Semantically normalize to the least ambiguous or most common word.
2. Semantic normalization takes precedence over syntactic normalization.
3. When possible, adverbs normalize to associated adjectives.

4. Semantically normalize verbal redundancies to a single term.
5. Comparative adjective forms are indicated by appending -x2, superlative by -x3.
6. Add superordinate conceptual organizers or conceptual antonyms when needed for continuity or clarification and indicate by use of italics.
7. Order in array is determined by what is most important in the domain, or by the most appropriate of the following organizing strategies: big to little, general to specific, top to bottom, most common to least common.

4.2.7 Finalizing the term list and concept hierarchy

The preliminary faceted vocabulary and the hierarchical structure of facets and isolates were then simplified and finalized. This was done by considering the implications of the visual structure of the terms. Examples of this process are the facets <Curve> and <Urban>.

<Curve> initially had 17 isolates: arc, arch, bump, curl, hook, loop, parabola, spiral, crescent, horseshoe, keyhole, teardrop, heart, circle, ring, oval, semicircle. Close examination identified a possible division between line curves (e.g., spiral) and closed curves (e.g., crescent). A possible exception to this organization was the isolate semicircle; but the visual structure of a *semi-circle* is similar to an arc, defined as "an arched line" (Abate, 2002), as are the visual structures of a *parabola* and an *arch*. Following this logic, <Arc> became a subfacet nested under the facet <Line-curve> and containing the isolates arch, semi-circle, and parabola. Bump was moved to <Protrusion> as a better visual match; and ring was identified as a near-synonym for circle. The subfacets under <Line-curve> were then ordered as <Spiral>, the fullest line curve, approaching circularity; <Loop>, a curl overlapping itself; <Curl>, an irregular curve; <Arc> indicating a symmetrical-shape; and <Hook>, a near-symmetrical shape having one shorter side. Closed curves were then ordered based on a visual structure from most circular to least circular: circle, oval, and crescent. The isolates horseshoe, teardrop, keyhole, and heart were identified as more appropriately belonging in the <Symbol> facet. The structure of the facet <Curve> was thus reduced to two semantically coherent subfacets: <Closed-curve> and <Line-curve>.

<Urban> initially had eight isolates: city, town, community, suburb, subdivision, trailer-park, neighborhood, and block. Semantically, it might seem to make sense to simplify this array by collapsing it to city, town, community (UF: suburb, subdivision, trailer-park), neighborhood, and city-block. But, when visual structure was considered, analysis identified a *suburb* as located next to a *city*;

a *subdivision* as demonstrating various street patterns; and a *trailer-park* as consisting of regular arrangements of rectangular shapes. The visual structure of a community was more ambiguous and so community was collapsed as a near-synonym for town. The isolates for <Urban> became: city, town, suburb, subdivision, trailer-park, neighborhood, and city-block.

In the end, however, more terms were retained in the faceted vocabulary than might be present in a standard controlled vocabulary. This was due to the visual nature of the vocabulary. For example, parking-garage and parking-lot are conceptually similar but visually distinct, so both terms were retained in the vocabulary.

The controlled vocabulary and faceted scheme for describing the internal contextuality of images could potentially include vocabulary representing the whole physical world, but that would provide for too much specificity to be useful. Unfortunately, the problem of excessive specificity is exacerbated by the nature of facet construction. For example, the facet <Seasonal-palette> contained only two terms generated in subject descriptions: summer-color and fall-color. A well-developed faceted vocabulary should account for all possible seasonal-palettes by adding isolates for autumn colors and winter colors. This is not a problem when there are only a total of four concepts, but it does become problematic when considering a facet such as <Sociopolitical-location>. Theoretically, every instance of a <Continent>, <Country>, <State>, and <Municipality> should be included in the controlled vocabulary as there is a finite number that exists, albeit far too large for this context.

The goal of this project was to identify kinds of concepts and not an exhaustive level of specificity. Therefore, because the words that were collected from the oral descriptions provided by subjects are representative of kinds of concepts, only the specific terms actually used by subjects are included in the vocabulary. Collapsing words to terms made possible a subject-generated level of specificity in the faceted hierarchy. Obviously, this is not a traditional controlled vocabulary where terms are limited to nouns and noun phrases. Adjectives, adverbs and prepositions were included because, in the context of image description, they were necessary for the comparative aspect (e.g., shallow, almost, and between) and could not be eliminated as stop words.

The semantic normalization process did not generate a faceted classification scheme because neither a notation system nor a citation order were created. These features were deemed to be beyond the scope of this research as they are dependent on the information needs of the end user. Creating a citation order involves understanding the use to which the controlled vocabulary will be put. In contrast, the objective of this research was to determine if a shared vocabulary of concepts could be identified for representation of the internal contextuality of images. As such, it was not relevant to determine whether the end users of this vocabulary would be CBIR developers, vocabulary developers, or actual searchers.

CHAPTER 5
Results

The existence of a shared vocabulary for visual structure was investigated by presenting 22 pairs of subjects with images selected from three domains (i.e., ART, MIC, and SAT) and by collecting the natural language words subjects used to describe those images. The words used to describe images were identified by subject pair and, within each subject pair, by image number (1 to 14) and by the domain from which each image was taken. After stop words (see Appendix B) had been eliminated from the list of natural language words generated by subjects (see Appendix A), frequency of occurrence was tallied for each unique natural language word across all the descriptions generated by subject pairs. As words were evaluated, first by normalizing variant forms (Appendix C), then by normalizing semantic referents and situating the resulting terms within the hierarchical structure of the faceted vocabulary (Appendix E), frequency of occurrence was tallied for each word, for each normalized form, for each grouping of synonymous and near-synonymous terms, and, finally, for each concept represented by an isolate or facet label. The units of analysis, therefore, are the terms and, more importantly, the superordinate concepts under which terms are grouped in the faceted vocabulary.

Each term or concept had the potential to be used by 22 pairs of subjects in each of the three image domains. However, due to malfunctioning of the audiotape recorder, one subject pair was eliminated from the analysis, leaving data for 21 subject pairs.

Transcriptions of the subject descriptions generated a total of approximately 107,581 natural language words (see Appendix A for a list of all unique natural language words). An exact count of the natural language words was not maintained because interpretations of meaning based on context were appended to some words, hyphenated phrases were retained until they could be explicated, and the development of processing guidelines was iterative during normalization, (e.g., the rule that utterances and partial words were not to be counted). Stop words accounted for approximately 50% of the total words generated by subjects. Stop words that were identified during transcription are included in the list of natural language words in Appendix A. Additionally, a list of the complete set of 221 unique stop words is provided in Appendix B.

There was a total count of 51,629 natural language word occurrences after stop words were eliminated. Once these natural language words had been normalized and variant forms eliminated, 2,075 unique words remained. After these 2,075 words had been semantically normalized by collapsing synonyms and near-synonyms, 1,319 unique terms remained. This final list of terms included 225 superordinate conceptual organizers and conceptual antonyms that had been introduced during construction of the faceted vocabulary.

5.1 Term frequencies

The results of tabulating all occurrences for each of the 1,319 terms across all pairs of subjects are included in Appendix D. Totals for frequency of occurrence of these terms ranged from zero to 3,695 per term, with zero frequency of occurrence indicating superordinates and antonyms introduced during the process of constructing the faceted vocabulary. Descriptions by individual subjects ranged from a total of 12 to 826 words per description;[67] and the number of words per description by image domain ranged from 20 to 826 for ART; from 12 to 353 for MIC; and from 13 to 702 for SAT. The median number of words per description was 155. The range for frequency of occurrence for any individual term within a description generated by a single subject pair was 1 to 28.

5.1.1 Terms with highest frequency and pair count

Pair count was used as a criterion for selecting terms with the highest frequency counts. There are 60 terms that had high frequency counts and occurred in the descriptions of 20 to 21 (20/21) subject pairs. These terms are reported in Table 5.1. In limiting Table 5.1 to terms with a pair count of 20/21, the assumption was made that any difference between usage by 20 versus 21 pairs of subjects is due to the specificity of the image, subject individuality, or simple chance. The high frequency terms listed in Table 5.1 account for all the terms that occurred in 20/21 pairs, but they do not necessarily represent the highest frequency counts.[68] However, it can be assumed that the 60 terms used by 20/21 subject pairs indicate the potential for a shared vocabulary.

---

[67] Full transcription of the description with only 12 words (P18.2#3,-12th) is: a dark square, a dark box, 4 seeds in it, 4 seeds.

[68] Although the term inch had a high frequency count (544), it only occurred in the descriptions of 17 subject pairs. For this reason it was not included in Table 5.1.

Table 5.1. *All terms with high frequency counts that were used by 20/21 subject pairs.*

Freq. Rank	Term	Pair Total	Freq. Count	Freq. Rank	Term	Pair Total	Freq. Count
1	similar	21	2670	31	<Part>	21	354
2	a	21	2174	32	<Joined>	21	347
3	<Inside-of>	21	1763	33	<Triangle>	20	345
4	left	21	1524	34	almost	20	325
5	right	21	1458	35	more	21	309
6	top	21	1152	36	end	21	307
7	bottom	21	1078	37	straight	21	305
8	side	21	941	38	<Exact>	21	285
9	center	21	937	39	<Cross>	21	285
10	<On>	21	913	40	1	20	278
11	small	21	877	41	<Off>	21	271
12	<Rectangle>	21	860	42	long	20	242
13	down	21	853	43	same	20	224
14	negation	21	701	44	<Surrounding>	21	218
15	corner	21	690	45	3	21	216
16	<Shape>	21	686	46	many	21	214
17	circle	21	682	47	some	21	198
18	edge	21	637	48	<Beside>	20	188
19	2	21	631	49	squiggle	20	185
20	very	21	581	50	below	20	163
21	all	21	521	51	<Between>	20	151
22	large	21	462	52	different	21	144
23	<Dot>	21	442	53	<Color>	20	140
24	start	21	374	54	<Whole>	21	132
25	square	21	367	55	<Area>	20	130
26	<Line>	20	1286	56	above	20	129
27	<Approximate>	20	1066	57	4	20	115
28	up	20	825	58	tiny	20	104
29	1/2	20	537	59	<Land>	20	100
30	<Outside-of>	21	355	60	aerial-view	20	89

Note. Rankings by frequency of occurrence across all terms have been assigned to facilitate comparisons.

Frequency of occurrence for the set of terms used by 20/21 subject pairs ranged from a high of 2670 for similar to a low of 89 for aerial-view. There are other terms with higher frequency counts than aerial-view that do not appear in Table 5.1 because they did not occur in 20/21 pairs. High frequency terms with pair value less than 20 are excluded from this listing as they are not indicative of a shared vocabulary of terms under the strict 20/21 pair criteria.

5.1.2 Image Domain Term Counts

The high frequency terms with 20/21 pair occurrences listed in Table 5.1 were investigated for distribution across the three image domains. Table 5.2 indicates that there is general distribution of term occurrence across the three domains, with the ART domain having a slightly higher percent of total terms used. When considering the number of unique terms used in each domain, the distribution is fairly even. However, it must be remembered that the totals reported for unique terms by domain represent a simple count of terms used out of the total 1,319 unique terms possible and do not imply use of the same terms across domains.

Table 5.2. *Distribution of term frequencies across image domains.*

	ART	MIC	SAT	Total terms
Total term frequency	20,618	13,694	17,317	51,629
% of Total term frequency	40%	27%	34%	
Total unique terms	684	624	622	1,319
% of Total unique terms	63%	57%	57%	

Note. Total counts for unique terms do not equal the sum of unique terms for the three domains because the same term may be used in multiple domains.

Table 5.3 presents a different view of term distribution across domains in that it identifies the terms with the highest frequency counts within each domain. A frequency count of 100 or more was used as the criterion for inclusion in Table 5.3 to demonstrate the overlap in term use across domains. Although the majority of these terms occurred in 19 to 21 subject pairs, the actual range for pair occurrence was 8 to 21. Table 5.4 lists all terms used by 20/21 subject pairs regardless of frequency of occurrence.

Table 5.3. *Domain occurrence of terms with frequency count of 100 or more.*

ART				**ART**			
Freq. Rank	**Terms**	**Pair Total**	**Freq. Count**	**Freq. Rank**	**Terms**	**Pair Total**	**Freq. Count**
1	similar	21	838	25	square	21	193
2	a	21	836	26	very	21	185
3	<Inside-of>	21	691	27	<Triangle>	20	182
4	<Line>	20	676	28	large	21	181
5	left	21	638	29	all	21	176
6	right	21	605	30	<Joined>	21	172
7	<Approximate>	20	537	31	circle	21	164
8	<Rectangle>	21	501	32	straight	21	154
9	top	21	455	33	long	20	149
10	side	21	448	34	<Outside-of>	21	144
11	bottom	21	431	35	1	20	142
12	down	21	379	36	<Part>	21	139
13	<On>	21	357	37	<Cross>	21	138
14	<Shape>	21	333	38	1/3	15	135
15	up	20	318	39	start	21	129
16	center	21	316	40	<Horizontal>	17	127
17	negation	21	294	41	centimeter	8	125
18	2	21	290	42	<Section>	17	123
19	1/2	20	276	43	end	21	119
20	edge	20	276	44	3	21	114
21	small	20	270	45	more	21	111
22	<Dot>	21	244	46	almost	20	105
23	inch	17	225	47	<Curve>	19	104
24	corner	21	196				

Note. The ART domain had a total of 684 unique terms out of 1319 possible unique terms. Table 5.3 continued on next page.

Table 5.3 continued.

MIC

Freq. Rank	Terms	Pair Total	Freq. Count
1	similar	21	858
2	<Inside-of>	21	550
3	a	21	485
4	right	21	382
5	left	21	349
6	center	21	315
7	circle	21	314
8	top	21	313
9	bottom	21	285
10	small	20	284
11	<On>	21	264
12	<Line>	20	263
13	corner	21	215
14	side	21	215
15	<Shape>	21	203

MIC

Freq. Rank	Terms	Pair Total	Freq. Count
16	<Rectangle>	21	201
17	very	21	200
18	2	21	191
19	up	20	189
20	negation	21	184
21	all	21	181
22	<Approximate>	20	176
23	down	21	171
24	edge	21	167
25	large	21	147
26	<Cell>	15	114
27	squiggle	20	113
28	inch	17	107
29	more	21	101

SAT

Freq. Rank	Terms	Pair Total	Freq. Count
1	similar	21	974
2	a	21	803
3	left	21	537
4	<Inside-of>	21	522
5	right	21	471
6	top	21	384
7	bottom	21	362
8	<Approximate>	20	353
9	<Line>	20	347
10	small	20	323
11	up	20	318
12	center	21	306
13	down	21	303
14	<On>	21	292
15	corner	21	279
16	side	21	278
17	<Road>	19	260
18	negation	21	223
19	inch	17	212
20	circle	21	204

SAT

Freq. Rank	Terms	Pair Total	Freq. Count
21	very	21	196
22	edge	21	194
23	1/2	20	166
24	all	21	164
25	<Rectangle>	21	158
26	<Curve>	19	155
27	<Shape>	21	150
28	2	21	150
29	start	21	148
30	<Part>	21	137
31	large	21	134
32	<Dot>	21	131
33	almost	20	129
34	end	21	126
35	<Exact>	21	125
36	<Outside-of>	21	122
37	square	21	116
38	straight	21	105
39	<Surrounding>	21	105
40	<Land>	20	100

Note. The MIC domain had a total of 624 unique terms out of 1319 possible unique terms; the SAT domain had a total of 622 unique terms out of 1319.

Table 5.4. *Terms with 20/21 subject pair counts ordered first by pair count and then by frequency count.*

ART

Pair Rank	Freq. Rank	Terms	Pair Total	Freq. Count
1	1	similar	21	838
2	2	a	21	836
3	3	<Inside-of>	21	691
5	5	left	21	638
6	6	right	21	605
8	8	<Rectangle>	21	501
9	9	top	21	455
10	10	side	21	448
11	11	bottom	21	431
12	12	down	21	379
13	13	<On>	21	357
14	14	<Shape>	21	333
16	16	center	21	316
17	17	negation	21	294
18	18	2	21	290
19	20	edge	21	276
21	22	<Dot>	21	244
22	24	corner	21	196
23		square	21	193
24		very	21	185
25		large	21	181
26		all	21	176
27		<Joined>	21	172
28		circle	21	164
29		straight	21	154
30		<Outside-of>	21	144
31		1	21	142
32		<Part>	21	139
33		<Cross>	21	138
34		start	21	129
35		end	21	119
36		3	21	114
37		more	21	111
38		<Off>	21	95
39		<Exact>	21	78
40		<Surrounding>	21	76
41		some	21	66
42		different	21	60
43		many	21	56
44		<Whole>	21	42
4	4	<Line>	20	676
7	7	<Approximate>	20	537
15	15	up	20	318
45	19	1/2	20	276
20	21	small	20	270
46		<Triangle>	20	182
47		long	20	149
48		almost	20	105
49		below	20	87
50		same	20	86
51		<Beside>	20	67
52		<Between>	20	61
53		4	20	59
54		<Color>	20	46
55		squiggle	20	35
56		above	20	34
57		<Area>	20	33
58		tiny	20	16
59		<Land>	20	1
60		aerial-view	20	1

Note. Terms are the same ones listed in Table 5.1. Frequency ranks come from Table 5.3. Table 5.4 continued on next two pages.

Table 5.4 continued.

MIC

Pair Rank	Freq. Rank	Terms	Pair Total	Freq. Count
1	1	MIC Terms	21	858
2	2	<Inside-of>	21	550
3	3	a	21	485
4	4	right	21	382
5	5	left	21	349
6	6	center	21	315
7	7	circle	21	314
8	8	top	21	313
9	9	bottom	21	285
10	11	<On>	21	264
11	13	corner	21	215
12	14	side	21	215
13	15	<Shape>	21	203
14	16	<Rectangle>	21	201
15	17	very	21	200
16	18	2	21	191
17	20	negation	21	184
18	21	all	21	181
19	23	down	21	171
20	24	edge	21	167
21	25	large	21	147
22	29	more	21	101
23		start	21	97
24		<Dot>	21	91
25		<Outside-of>	21	89
26		many	21	88
27		<Exact>	21	82
28		<Joined>	21	81
29		<Off>	21	79
30		<Part>	21	78

MIC

Pair Rank	Freq. Rank	Terms	Pair Total	Freq. Count
31		some	21	73
32		1	21	69
33		end	21	62
34		square	21	58
35		3	21	56
36		<Cross>	21	50
37		straight	21	46
38		<Whole>	21	38
39		different	21	38
40		<Surrounding>	21	37
41	10	small	20	284
42	12	<Line>	20	263
43	19	up	20	189
44	22	<Approximate>	20	176
45	27	squiggle	20	113
46		1/2	20	95
47		almost	20	91
48		same	20	83
49		<Color>	20	66
50		<Triangle>	20	66
51		long	20	65
52		<Between>	20	56
53		<Beside>	20	47
54		below	20	46
55		above	20	39
56		tiny	20	29
57		4	20	28
58		<Area>	20	19
59		<Land>	20	1
60		aerial-view	20	0

Note. Table 5.4 continued on next page.

Table 5.4 continued.

SAT					SAT				
Pair Rank	Freq. Rank	Terms	Pair Total	Freq. Count	Pair Rank	Freq. Rank	Terms	Pair Total	Freq. Count
1	1	similar	21	974	31		more	21	97
2	2	a	21	803	32		<Off>	21	97
3	3	left	21	537	33		<Cross>	21	97
4	4	<Inside-of>	21	522	34		<Joined>	21	94
5	5	right	21	471	35		1	21	74
6	6	top	21	384	36		many	21	70
7	7	bottom	21	362	37		some	21	59
8	12	center	21	306	38		<Whole>	21	52
9	13	down	21	303	39		3	21	46
10	14	<On>	21	292	40		different	21	46
11	15	corner	21	279	41	8	<Approximate>	20	353
12	16	side	21	278	42	9	<Line>	20	347
13	18	negation	21	223	43	10	small	20	323
14	20	circle	21	204	44	11	up	20	318
15	21	very	21	196	45	23	1/2	20	166
16	22	edge	21	194	46	33	almost	20	129
17	24	all	21	164	47	40	<Land>	20	99
18	25	<Rectangle>	21	158	48		<Triangle>	20	97
19	27	<Shape>	21	150	49		aerial-view	20	88
20	28	2	21	150	50		<Area>	20	78
21	29	start	21	148	51		<Beside>	20	74
22	30	<Part>	21	137	52		tiny	20	59
23	31	large	21	134	53		above	20	56
24	32	<Dot>	21	131	54		same	20	55
25	34	end	21	126	55		squiggle	20	37
26	35	<Exact>	21	125	56		<Between>	20	34
27	36	<Outside-of>	21	122	57		below	20	30
28	37	square	21	116	58		<Color>	20	28
29	38	straight	21	105	59		long	20	28
30	39	<Surrounding>	21	105	60		4	20	28

There is considerable overlap between terms with 100 or more occurrences in a given domain and the 60 high frequency terms listed in Table 5.1. This is demonstrated in Table 5.5 where the top 60 terms are listed on the left in the same order used in Table 5.1, allowing for the comparison of term occurrence across domains. Table 5.6 combines the 60 high frequency terms from Table 5.1 with terms that had a frequency count of 100 or more within one or more of the three domains. Each of the domains has high frequency terms that are not included in the list of 60 high frequency terms: cell is a high frequency term in MIC; road, and curve are high frequency terms shared by ART and SAT; and curve, section, 1/3, centimeter, and horizontal are high frequency terms in ART.

Table 5.5: *Domain distribution of terms from Table 5.1 ordered by total frequency count.*

Pair Total	Total Freq. Count	Terms	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	ART Freq. Rank	MIC Freq. Rank	SAT Freq. Rank
21	2670	similar	**838**	**858**	**974**	1	1	1
21	2174	a	**836**	**485**	**803**	2	3	2
21	1763	<Inside-of>	**691**	**550**	**522**	3	2	4
21	1524	left	**638**	**349**	**537**	5	5	3
21	1458	right	**605**	**382**	**471**	6	4	5
20	1286	<Line>	**676**	**263**	**347**	4	12	9
21	1152	top	**455**	**313**	**384**	9	8	6
21	1078	bottom	**431**	**285**	**362**	11	9	7
20	1066	<Approximate>	**537**	**176**	**353**	7	22	8
21	941	side	**448**	**215**	**278**	10	13	16
21	937	center	**316**	**315**	**306**	16	6	12
21	913	<On>	**357**	**264**	**292**	13	11	14
21	877	small	**270**	**284**	**323**	21	10	10
21	860	<Rectangle>	**501**	**201**	**158**	8	16	23
21	853	down	**379**	**171**	**303**	12	23	13
20	825	up	**318**	**189**	**318**	15	19	11
21	701	negation	**294**	**184**	**223**	17	20	17
21	690	corner	**196**	**215**	**279**	23	14	15
21	686	<Shape>	**333**	**203**	**150**	14	15	24
21	682	circle	**164**	**314**	**204**	30	7	18
21	637	edge	**276**	**167**	**194**	19	24	20
21	631	2	**290**	**191**	**150**	18	18	25
21	581	very	**185**	**200**	**196**	25	17	19
20	537	1/2	**276**	95	**166**	20	29	21
21	521	all	**176**	**181**	**164**	28	21	22
21	462	large	**181**	**147**	**134**	27	25	28
21	442	<Dot>	**244**	91	**131**	22	30	29
21	374	start	**129**	97	**148**	37	28	26
21	367	square	**193**	58	**116**	24	45	34
21	355	<Outside-of>	**144**	89	**122**	33	32	33
21	354	<Part>	**139**	78	**137**	35	38	27
21	347	<Joined>	**172**	81	94	29	36	42

Note. Terms from Table 5.1 have 20/21 subject pair counts and high frequency. Domain distributions with total frequency counts of 100 or more are in **bold**. See also Table 5.6 for the same terms ordered according to position in the faceted structure. Table 5.5 continued on next page.

Table 5.5 continued.

Pair Total	Total Freq. Count	Terms	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	ART Freq. Rank	MIC Freq. Rank	SAT Freq. Rank
20	345	<Triangle>	**182**	66	97	26	41	39
20	325	almost	**105**	91	**129**	41	31	30
21	309	more	**111**	**101**	97	40	27	37
21	307	end	**119**	62	**126**	38	44	31
21	305	straight	**154**	46	**105**	31	50	35
21	285	<Exact>	78	82	**125**	45	35	32
21	285	<Cross>	**138**	50	97	36	48	40
21	278	1	**142**	69	74	34	40	45
21	271	<Off>	95	79	97	42	37	38
20	242	long	**149**	65	28	32	43	59
20	224	same	86	83	55	44	34	51
21	218	<Surrounding>	76	37	**105**	46	55	36
21	216	3	**114**	56	46	39	46	53
21	214	many	56	88	70	52	33	47
21	198	some	66	73	59	48	39	48
20	188	<Beside>	67	47	74	47	49	46
20	185	squiggle	35	**113**	37	55	26	55
20	163	below	87	46	30	43	51	57
20	151	<Between>	61	56	34	49	47	56
21	144	different	60	38	46	50	53	54
20	140	<Color>	46	66	28	53	42	58
21	132	<Whole>	42	38	52	54	54	52
20	130	<Area>	33	19	78	57	58	44
20	129	above	34	39	56	56	52	50
20	115	4	59	28	28	51	57	60
20	104	tiny	16	29	59	58	56	49
20	100	<Land>	1	1	98	59	59	41
20	89	aerial-view	1	0	88	60	60	43

Table 5.6. *Terms with 20/21 subject pair counts or frequency count of 100 or more ordered according to position in the faceted structure.*

Pair Total	Total Freq. Count	Total Freq. Rank	Term	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	ART Freq. Rank	MIC Freq. Rank	SAT Freq. Rank
			<OBJECT>						
21	686	19	<Shape>	**333**	**203**	**150**	14	15	27
21	442	28	<Dot>	**244**	91	**131**	22	32	32
20	1286	6	<Line>	**676**	**263**	**347**	4	12	9
20	185	56	squiggle	35	**113**	37	61	27	62
21	690	18	corner	**196**	**215**	**279**	24	14	15
19	278	41	<Curve>	**104**	19	**155**	47	63	26
21	682	20	circle	**164**	**314**	**204**	31	7	20
20	345	34	<Triangle>	**182**	66	97	27	43	42
21	860	14	<Rectangle>	**501**	**201**	**158**	8	16	25
21	367	30	square	**193**	58	**116**	25	47	37
15	114	65	<Cell>	0	**114**	0	68	26	68
19	266	44	<Road>	4	2	**260**	65	66	17
			<PLACE>						
20	100	67	<Land>	1	1	98	66	67	44
20	130	62	<Area>	33	19	78	63	64	47

Note. A total of 68 terms are included: 60 terms from Table 5.1 (i.e., high frequency and occurring in 20/21 subject pairs) and an additional eight terms from Table 5.3 (i.e., frequency of 100 or more). Table 5.6 continued on next two pages.

Table 5.6 continued.

Pair Total	Total Freq. Count	Total Freq. Rank	Term	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	ART Freq. Rank	MIC Freq. Rank	SAT Freq. Rank
			<PROPERTY>						
21	701	17	negation	**294**	**184**	**223**	17	20	18
21	132	61	<Whole>	42	38	52	60	59	57
21	354	32	<Part>	**139**	78	**137**	36	40	30
17	228	47	<Section>	**123**	55	50	42	50	59
15	254	45	1/3	**135**	42	77	38	55	48
20	537	25	1/2	**276**	95	**166**	20	31	23
20	224	48	same	86	83	55	50	36	56
20	2670	1	similar	**838**	**858**	**974**	1	1	1
21	144	59	different	60	38	46	56	58	61
21	2124	2	a	**836**	**485**	**803**	2	3	2
21	521	26	all	**176**	**181**	**164**	29	21	24
21	214	52	many	56	88	70	58	35	51
21	198	54	some	66	73	59	54	41	53
21	278	42	1	**142**	69	74	35	42	49
21	631	22	2	**290**	**191**	**150**	18	18	28
21	216	51	3	**114**	56	46	44	48	60
20	115	64	4	59	28	28	57	62	67
8	206	53	centimeter	**125**	16	65	41	65	52
17	544	24	inch	**225**	**107**	**212**	23	28	19
20	104	66	tiny	16	29	59	64	61	54
20	877	13	small	**270**	**284**	**323**	21	10	10
21	462	27	large	**181**	**147**	**134**	28	25	31
20	242	46	long	**149**	65	28	33	45	66
20	325	35	almost	**105**	91	**129**	46	33	33
21	581	23	very	**185**	**200**	**196**	26	17	21
20	1066	9	<Approximate>	**537**	**176**	**353**	7	22	8
21	285	39	<Exact>	78	82	**125**	51	37	35
21	309	36	more	**111**	**101**	97	45	29	40
21	305	38	straight	**154**	46	**105**	32	53	38
20	89	68	aerial-view	1	0	88	67	68	46
20	140	60	<Color>	46	66	28	59	44	65

Note. Table 5.6 continued on next page.

Table 5.6 continued.

Pair Total	Total Freq. Count	Total Freq. Rank	Term	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	ART Freq. Rank	MIC Freq. Rank	SAT Freq. Rank
			<SPATIAL-LOCATION>						
21	941	10	side	**448**	**215**	**278**	10	13	16
21	1152	7	top	**455**	**313**	**384**	9	8	6
21	937	11	center	**316**	**315**	**306**	16	6	12
21	1078	8	bottom	**431**	**285**	**362**	11	9	7
21	374	29	start	**129**	97	**148**	39	30	29
21	307	37	end	**119**	62	**126**	43	46	34
21	637	21	edge	**276**	**167**	**194**	19	24	22
21	285	40	<Cross>	**138**	50	97	37	51	43
21	347	33	<Joined>	**172**	81	94	30	38	45
21	1763	3	<Inside-of>	**691**	**550**	**522**	3	2	4
21	218	49	<Surrounding>	76	37	**105**	52	60	39
21	355	31	<Outside-of>	**144**	89	**122**	34	34	36
21	271	43	<Off>	95	79	97	48	39	41
21	913	12	<On>	**357**	**264**	**292**	13	11	14
20	188	55	<Beside>	67	47	74	53	52	50
20	151	58	<Between>	61	56	34	55	49	63
20	129	63	above	34	39	56	62	57	55
20	163	57	below	87	46	30	49	54	64
20	825	16	up	**318**	**189**	**318**	15	19	11
21	853	15	down	**379**	**171**	**303**	12	23	13
17	218	50	<Horizontal>	**127**	40	51	40	56	58
21	1524	4	left	**638**	**349**	**537**	5	5	3
21	1458	5	right	**605**	**382**	**471**	6	4	5

Analysis of term frequency counts in Table 5.4 (i.e., terms with 20/21 subject pair counts in each domain) indicates that 15 (25%) of the 60 high frequency terms used by 20/21 subject pairs (see Table 5.1) are not among the high frequency terms for any of the three domains: land, area, whole, same, different, many, some, 4, tiny, aerial-view, color, off, beside, between, above. This indicates that, because these terms have relatively high frequency counts, their use is not specific to a particular domain, but, in general, that the use of these terms is more evenly distributed across all three domains (with the exception of aerial-view and <Land>). For example, the term some has domain distributions of 66 (ART), 73 (MIC), and 32 (SAT), as compared to the term squiggle, which has a domain distribution indicating a higher frequency of use in MIC images (i.e., domain distributions of 35 [ART], 113 [MIC], and 32 [SAT]).

Although the 60 terms used by 20/21 pairs (see Table 5.1) appear to be evenly distributed across the three domains, if the actual distribution of individual term frequency counts is taken into consideration, this may not be the case. For example, the terms land and aerial-view aptly illustrate the disparity in distribution across the three domains (see Table 5.7). Although these two terms occur in the descriptions of 20/21 subject pairs, with the exception of three occurrences, they are used exclusively in the SAT domain. This indicates that some terms in the shared vocabulary are highly domain specific.

Table 5.7. *Disparity in domain distribution of frequency counts for* land *and* aerial-view.

Term	Total Frequency	ART Frequency	MIC Frequency	SAT Frequency
Land	100	1	1	98
Aerial-view	89	1	0	88

5.1.3 Lowest occurrence terms

Low term frequency counts and pair occurrence values are also of interest because they are assumed to be of less value to a shared vocabulary. There were 447 terms with a pair count of 1 and 225 terms with a pair count of 0, the latter representing the superordinates and conceptual antonyms introduced during construction of the faceted vocabulary. Frequency counts for the 447 terms with pair count of 1 ranged from 1 to 28. Interestingly, the high frequency count of 28 was for the term pin, which was used by a single subject as a reference point for describing all aspects of a single image (P14.1 #234-1st). Other examples of terms used repeatedly by a single pair are: matte, C, H, Q, panel, stitching, checkmark, zebra, surfboard, and olive. Of the 447 terms with a pair count of 1, 73% are isolates and 25% are facet labels. In contrast, of the 225 terms with a pair count of 0, 7% are isolates and 97% are facet labels.

For the 165 terms with a pair count of 2, term frequency ranges from 2 to 24. Following the same rationale used in the combination of 20 and 21 subject pair occurrences as indicative of shared vocabulary, terms with pair counts of 1 and 2 are assumed to indicate vocabulary that is not shared. This is particularly true of the 46 terms with a pair count 2 and a frequency count of 2, demonstrating that each of these terms was only used once by each of two pairs. Applying pair frequency of 1 or 2 pair occurrences as the criteria for low pair count, there are 612 terms (46%) of the 1,319 total unique terms that appear not to be shared across subject pairs.

After removing from consideration the 612 low count terms as well as the 225 terms with pair count of 0, there are 492 terms that remain as potentially shared.

However, as demonstrated by the distribution of term occurrence across domains (see Table 5.2), frequency counts may indicate common usage, but they do not demonstrate common usage across domains.

5.1.4 Parts of speech

As described in Chapter 4, the controlled vocabulary developed in this study did not always adhere to traditional conventions regarding parts of speech. Thus, for example, adverbs were included as information bearing terms in the context of internal contextuality. Table 5.8 provides an overview of frequency counts by parts of speech for the set of 1,094 unique terms that were actually used in subject descriptions 9i.e., 1319 total unique terms minus 225 terms with pair count of 0). These frequency counts for parts of speech are only approximations because the determination of a term's part of speech is dependent on its usage in the research task and may have been changed during the normalization process.

Table 5.8. *Percent of total terms by part of speech.*

Part of speech	**% of total unique terms (1094)**	**% of total frequency count (51,629)**
Noun	63%	40%
Adjective	32%	44%
Adverb	3%	5%
Preposition	2%	10%
Other: Conjunction, phrase, etc.	<1%	1%

Note. 1,094 represents the total number of unique terms in the faceted vocabulary (1,319) minus the superordinates terms and conceptual antonyms that were introduced during construction of the faceted vocabulary (225). Participles and numeric values are included in the adjective count.

5.1.5 Comparatives and superlatives

Comparative and superlative forms of terms were used by all 21 pairs of subjects. There is the potential for adding comparative endings to many base words, whether or not that creates an accepted word. Therefore, occurrence counts for adjectives and adverbs are necessarily dependent on the relationship between the comparative or superlative and its base word. In the tabulation process, comparative and superlative forms were treated as synonyms and were grouped with the base adjective or adverb. Thus, the occurrence count reported for any individual adjective or adverb includes all comparative or superlative forms that were used by subject pairs. For example, the term small has a frequency count of 877 which includes all comparative and superlative forms of small used by subject pairs.

As indicated in Table 5.9, there are 19 terms for which a comparative or superlative form was used. There were 3,078 occurrences of comparatives and superlatives, which accounted for 6% of total word occurrences. Terms that took the comparative or superlative form generally fell in the <PROPERTY> sub-facets of <Comparative-measure>, <Condition>, <Judgment>, and <Color-quality>, although there were exceptions: For example, few is in the facet <Linguistic-quantity>, and far and near are in the facet <Distance>.

Table 5.9. *Comparative and superlative forms of adjectives and adverbs.*

Superordinate	Base word	Synonyms	Pair Count	Freq. Count
<Linguistic-quantity>	Few	(fewer)	8	23
<Size>	Small	(smaller, smallest)	21	877
	Large	(larger, largest; big, bigger, biggest; great, greater)	21	462
<Width>	Thin	(thinner; narrow, narrower; skinny, skinnier)	18	99
	Wide	(wider; broad, broadest)	19	108
	Thick	(thicker, thickest; fat, fattest)	19	99
<Length>	Short	(shorter, shortest)	15	44
	Long	(longer, longest)	20	242
<Height>	Low	(lower, lowest)	19	169
	High	(higher; tall, taller, tallest)	15	63
<Concentration>	Dense	(heavy, heaviest)	10	28
<Sharpness>	Sharp	(sharper)	12	63
<Goodness>	Good	(better)	7	11
<Abnormal>	Strange	(weird, weirdest)	12	35
<Tone>	Pale	(light2, light2er, light2est)	18	73
	Dark	(darker, darkest)	19	126
<Brightness>	Bright	(brighter, brightest)	9	23
<Distance>	Far	(farther, farthest)	16	83
	Near	(nearest; close, closer, closest)	18	95

5.2 Collapsing to concepts

In order to determine the potential of a shared vocabulary of concepts, each isolate was collapsed into its superordinate facet. Collapsing was based on the assumption that isolates were recognizable as instances under the superordinate concept category. For example, blue is readily identifiable as an instance of the facet <Hue> and asterisk as an instance of the facet <Punctuation>.

Isolates at the lowest levels of the faceted hierarchy were collapsed to the superordinate facet, with the frequency count for the superordinate term (i.e., facet) accruing the frequency and pair counts of all subordinate isolates. Thus, the frequency count for a superordinate facet is the sum of the frequency counts for all isolate terms nested under the superordinate and the frequency count associated with the superordinate itself. Pair count for a superordinate facet is the union of the pair counts for all isolate terms and the pair count of the superordinate. In this process, superordinate organizers and conceptual antonyms were always included even though they might have values of 0 for pair count and frequency of use. Thus all terms are accounted for, but at the more general level of concept rather than the specific level of the isolate.

The process of collapsing isolates resulted in 545 concept categories. The frequency counts and subject pair values for concept categories is provided in Appendix F. There were 80 superordinate facets that had no subordinate isolates due to introduction during the construction of the faceted vocabulary. They retained the original pair value and frequency count of 0. Evaluation of all 545 concepts would skew calculation because it would include these 80 concepts that were never used by subjects. Therefore, the 80 concepts with pair and frequency counts of 0 were removed from calculations, leaving a set of 465 concepts for evaluation.

5.2.1 Concept frequencies and counts

There were 48 concepts that occurred in 20/21 pairs. However, because three concepts with relatively high frequency counts (i.e., <Section>, <Rotated> and <Curve>) occurred in 19 pairs, the reporting range was increased to include all terms with pair occurrence of 19 or greater. Table 5.10 lists the 59 concepts used by 19 or more subject pairs ordered by frequency of occurrence. Table 5.11 lists these same concepts ordered by position in the hierarchical structure of the faceted vocabulary.

Table 5.10. *Concepts used by 19 to 21 subject pairs ordered by total frequency count.*

Freq. Rank	Concept	Pair Total	Freq. Count
1	<Horizontal>	21	3200
2	*<General-part>*	21	3167
3	<Similarity>	21	3059
4	*<Linguistic-quantity>*	21	2875
5	<Vertical>	21	1823
6	<Inside-of>	21	1773
7	<Line>	21	1643
8	<Size>	21	1558
9	*<Certitude>*	21	1361
10	*<Extremity>*	21	1345
11	<Rectangle>	21	1227
12	<Degree>	21	1157
13	<Fraction>	20	1053
14	<Number>	21	1021
15	*<Presentation>*	21	980
16	<On>	21	913
17	<Angle>	20	861
18	*<Closed-curve>*	21	776
19	*<Hue>*	21	706
20	*<Validation>*	21	701
21	<Shape>	21	686
22	<Dot>	21	466
23	<Road>	21	359
24	<Joined>	21	359
25	<Outside-of>	21	355
26	<Part>	21	354
27	<Triangle>	20	345
28	<Unequal>	21	336
29	<Length>	21	335
30	*<Extension>*	21	317
31	<Piece>	21	315
32	*<Width>*	21	306
33	<Section>	19	294
34	*<Vertical-perspective>*	21	292
35	<Cross>	21	285
36	*<Rotated>*	19	283
37	<Curve>	19	278
38	<Beside>	20	276
39	<Off>	21	271
40	<Height>	21	259
41	<Concentration>	20	235
42	<Distance>	19	232
43	<Free-form>	21	222
44	<Tone>	21	222
45	<Surrounding>	21	218
46	<Urban>	19	162
47	<Between>	20	151
48	<Area>	20	149
49	<Emphasis>	20	146
50	<Color>	20	140
51	<Division>	19	140
52	<Whole>	21	132
53	<Symbol>	19	110
54	*<By-position>*	20	107
55	<Land>	20	100
56	<Building>	19	86
57	<Cut>	19	85
58	*<Existence>*	19	79
59	<Photograph>	19	55

Note. Some concepts appeared in preceding tables as terms with lower pair count occurrences or frequency counts. Counts have increased for these concepts due to accrual of counts from isolates that were collapsed into the superordinate concept.

Table 5.11. *Concepts used by 19 to 21 subject pairs according to position in faceted structure.*

Freq. Rank	Concept	Pair Total	Freq. Count
	<OBJECT>		
59	<Photograph>	19	55
53	<Symbol>	19	110
21	<Shape>	21	686
22	<Dot>	21	466
8	<Line>	21	1643
17	<Angle>	20	861
37	<Curve>	19	278
18	*<Closed-curve>*	21	776
27	<Triangle>	20	345
12	<Rectangle>	21	1227
43	<Free-form>	21	222
56	<Building>	19	86
23	<Road>	21	359
46	<Urban>	19	162
	<PLACE>		
55	<Land>	20	100
48	<Area>	20	149
	<PROPERTY>		
58	*<Existence>*	19	79
20	*<Validation>*	21	701
52	<Whole>	21	132
26	<Part>	21	354
50	<Division>	19	140
33	<Section>	19	294
31	<Piece>	21	315
14	<Fraction>	20	1053
4	<Similarity>	21	3059
1	*<Linguistic-quantity>*	21	2875
5	<Number>	21	1021
9	<Size>	21	1558
32	*<Width>*	21	306

Freq. Rank	Concept	Pair Total	Freq. Count
29	<Length>	21	335
40	<Height>	21	259
13	<Degree>	21	1157
10	*<Certitude>*	21	1361
28	<Unequal>	21	336
49	<Emphasis>	20	146
41	<Concentration>	20	235
30	*<Extension>*	21	317
57	<Cut>	19	85
54	*<By-position>*	20	107
51	<Color>	20	140
19	*<Hue>*	21	706
44	<Tone>	21	222
	<SPATIAL-LOCATION>		
15	*<Presentation>*	21	980
36	*<Rotated>*	19	283
3	*<General-part>*	21	3167
11	*<Extremity>*	21	1345
42	<Distance>	19	232
35	<Cross>	21	285
24	<Joined>	21	359
7	<Inside-of>	21	1773
45	<Surrounding>	21	218
25	<Outside-of>	21	355
39	<Off>	21	271
16	<On>	21	913
38	<Beside>	20	276
47	<Between>	20	151
34	*<Vertical-perspective>*	21	292
6	<Vertical>	21	1823
2	<Horizontal>	21	3200

Note. Concepts are the same as listed in Table 5.10.

Concepts with a frequency count of 100 or more have a subject pair occurrence ranging from 15 to 21. This includes all of the concepts with subject pair occurrence of 19 to 21 except the concepts <Building>, <Cut> and <Existence>, which had frequency counts of 86, 85, and 79 respectively. These concepts are listed in Table 5.12. The ten high frequency concepts used by less than 19 subject pairs account for 15% of concepts with a frequency count of 100 or more (i.e., <Measurement-unit, <Letter>, <Diagonal>, <Increase>, <Regularity>, <Cell>, <Probability>, <Opening>, <Liquid>, and <Cover>). This indicates that, although these concepts have relatively high frequency counts, their use is generally not shared across subject pairs according to a strict 19 to 21 subject pair criteria.

Table 5.12. *Concepts with frequency count of 100 or more.*

Freq. Rank	Concept	Pair Total	Freq. Count	Freq. Rank	Concept	Pair Total	Freq. Count
1	<Horizontal>	21	3200	34	<Section>	19	294
2	*<General-part>*	21	3167	35	*<Vertical-perspective>*	21	292
3	<Similarity>	21	3059	36	<Cross>	21	285
4	*<Linguistic-quantity>*	21	2875	37	*<Rotated>*	19	283
5	<Vertical>	21	1823	38	<Curve>	19	278
6	<Inside-of>	21	1773	39	<Beside>	20	276
7	<Line>	21	1643	40	<Off>	21	271
8	<Size>	21	1558	41	<Height>	21	259
9	*<Certitude>*	21	1361	42	<Concentration>	20	235
10	<Extremity>	21	1345	43	<Letter>	15	233
11	<Rectangle>	21	1227	44	<Distance>	19	232
12	<Degree>	21	1157	45	<Diagonal>	17	222
13	<Fraction>	20	1053	46	<Free-form>	21	222
14	<Number>	21	1021	47	<Tone>	21	222
15	*<Presentation>*	21	980	48	<Surrounding>	21	218
16	<On>	21	913	49	<Increase>	16	172
17	<Angle>	20	861	50	<Urban>	19	162
18	*<Measurement-unit>*	17	784	51	*<Regularity>*	18	152
19	*<Closed-curve>*	21	776	52	<Between>	20	151
20	*<Hue>*	21	706	53	<Area>	20	149
21	*<Validation>*	21	701	54	<Cell>	16	147
22	<Shape>	21	686	55	<Emphasis>	20	146
23	<Dot>	21	466	56	*<Probability>*	17	143
25	<Joined>	21	359	57	<Color>	20	140
24	<Road>	21	359	58	<Division>	19	140
26	<Outside-of>	21	355	59	*<Opening>*	18	139
27	<Part>	21	354	60	<Whole>	21	132
28	<Triangle>	20	345	61	<Liquid>	15	120
29	<Unequal>	21	336	62	<Symbol>	19	110
30	<Length>	21	335	63	*<By-position>*	20	107
31	*<Extension>*	21	317	64	<Cover>	18	100
32	<Piece>	21	315	65	<Land>	20	100
33	*<Width>*	21	306				

5.2.2 Image domain concept counts

Collapsing isolates into superordinate concepts does not change frequency distributions across domains. Thus the frequency counts for concepts across domains will be the same as the frequency counts for terms presented in Table 5.2

5.2.3 Concepts in the faceted hierarchy

The hierarchical structure of the faceted vocabulary provides a framework for evaluating the concept categories of terms generated by subject pairs. The top level concepts are the four facets <OBJECT>, <PLACE>, <PROPERTY>, and <SPATIAL-LOCATION>. Table 5.13 shows the number of concepts used in each of the top level facets and their distribution by domain. Table 5.14 shows the frequency distributions of concepts across top level facets within image domains.

Table 5.13. *Concepts used in the four top level facets and their distribution across image domains.*

Concepts	**Total concepts (465)**	**ART Freq.**	**MIC Freq.**	**SAT Freq.**
<OBJECT>	227	165	152	110
<PLACE>	90	40	38	79
<PROPERTY>	108	100	93	90
<SPATIAL-LOCATION>	40	40	39	39

Note. Total counts for unique concepts in each of the four top level facets do not equal the sum of the three domains because the same concept may be used in multiple domains.

Table 5.14. *Frequency distribution of concepts across top level facets and their distribution across image domains.*

	Total frequency	**ART**	**MIC**	**SAT**
<OBJECT>	9826	4328	2850	2648
% Total Freq	19 %	8.5 %	5.5 %	5 %
<PLACE>	2092	266	287	1539
% Total Freq	4 %	0.5 %	0.5 %	3 %
<PROPERTY>	22832	9090	6159	7583
% Total Freq	44 %	17.5 %	12 %	14.5 %
<SPATIAL-LOCATION>	16879	6933	4399	5547
% Total Freq	33 %	13.5 %	8.5 %	11 %
Total freq. for all concepts	51629	20617	13695	17317
% Total Freq	100%	40 %	26.5 %	33.5 %

It is possible to compute concept frequencies at progressively subordinate levels of the faceted hierarchy but discerning valuable information from such computations is problematic. At the second level in the faceted hierarchy, there are 14 subordinate concepts (see Table 5.15). However, these concepts are unevenly distributed across the top level facets and, at this point, are only of interest because they indicate the variety by volume of unique terms used in each concept category. At the third subordinate level, there are 47 concepts, again unevenly distributed. Counts between concepts cannot be computationally compared between concepts at this level because of their uneven distribution; but they do provide insight into the distribution of concepts across subordinate levels in the hierarchy. More detailed data could be computed for each subsequent subordinate level, but such computations would only be of use in analyzing individual top level facets and could not be used for comparison across facets because of the hierarchical structure within each top level facet. Such analyses would be of primary interest for determining where the emphasis should be placed in the development of individual retrieval vocabularies for particular domains or user groups.

Table 5.15: *Concepts in the top three levels of the hierarchy of the faceted structure.*

Concepts: Hierarchy Level 1	**Concepts: Hierarchy Level 2**	**Concepts: Hierarchy Level 3**
<OBJECT> (601)		
	<*Image*> (17)	
		<*Kind-of-image*> (10)
		<*Image-foundation*> (7)
	<*Non-living-thing*> (353)	
		<Figure> (105)
		<*Artifact*> (178)
		<*Mechanical-part*> (5)
		<Substance> (8)
		<*Naturally-occurring-phenomena*> (57)
	<Living-organism> (233)	
		<Animal-life> (55)
		<Plant> (33)
		<Body> (117)
		<*Aspects-of-living-thing*> (27)
<PLACE> (224)		
	<*Constructed-environment*> (130)	
		<*Water-based-environment*> (9)
		<*Land-based-environment*> (109)
		<*Locale*> (11)
	<*Natural-place*> (62)	
		<Sky> (2)
		<*Body-of-water*> (17)
		<Shore> (4)
		<*Land-water-formation*> (6)
		<Terrain> (20)
		<*Ecosystem*> (4)
	<*Sociopolitical-location*> (26)	
		<Continent> (2)
		<Country> (9)
		<State> (4)
		<*Municipality*> (9)
	<*Generic-place*> (14)	
		<Area> (5)
		<*Opening*> (4)
		<Joint> (4)

Note. Counts for individual concepts are indicated in parentheses following each concept and do not include counts for any subordinate concepts. Table 5.15 continued on next page.

Table 5.15 continued.

Concepts: Hierarchy Level 1	Concepts: Hierarchy Level 2	Concepts: Hierarchy Level 3
<PROPERTY> (410)		
	<*General-concept*> (16)	
		<*Existence*> (3)
		<Domain> (9)
		<*Validation*> (3)
	<*Attribute*> (391)	
		<*Visibility*> (5)
		<*Gestalt*> (67)
		<Quantity> (79)
		<Comparison> (32)
		<*Condition*> (45)
		<*Judgment*> (30)
		<*Change-in-condition*> (14)
		<Action> (22)
		<*Art-and-craft-process*> (66)
		<Color> (38)
<SPATIAL-LOCATION> (85)		
	<*Format*> (3)	
	<Position> (65)	
		<*Indexical*> (23)
		<*Relational*> (42)
	<Direction> (10)	
		<Pointing-to> (1)
		<Vertical> (3)
		<Horizontal> (3)
		<Diagonal> (1)
		<Perpendicular-to> (1)
	<Compass-orientation> (9)	
	<Clock-orientation> (1)	

In the subset of unique concept counts in Table 5.16, frequency distribution by domain is provided for a subset of concepts from Table 5.15. Frequency numbers and domain distribution for the second level facet <Constructed environment> indicate both that concept use emphasized <Land-based-environment> and that, within <Land-based-environment>, frequency of use was dominated by subject descriptions of images from the SAT domain. Furthermore, the nine instances of <Water-based-environment> concepts occurred solely within the SAT domain.

Looking at other concepts listed in Table 5.16, the information to be gleaned is not as clear cut. Each concept and its distributions must be analyzed individually and in relation to its superordinate and coordinate concepts: For example, use of <Body-

part> concepts is skewed toward MIC, while occurrences of <Aspects-of-living-thing> concepts are skewed toward ART. The high counts for unique concepts and frequencies of occurrence for the subfacets <Figure> and <Artifact> indicate that any meaningful analysis would have to occur at lower levels in the hierarchy.

Table 5.16. *A subset of concepts in the top three levels of the faceted vocabulary and their frequency distribution across domains.*

Concepts: Hierarchy Levels 1 and 2	Concepts: Hierarchy Level 3	ART Freq. Count	MIC Freq. Count	SAT Freq. Count	Total Freq. Count
<OBJECT> (601)					
<Image> (17)					
	<Kinds-of-image> (10)	23	16	52	91
	<Image-foundation> (7)	2	6	5	13
<Non-living-thing> (353)					
	<Figure> (105)	3460	1869	2010	7339
	<Artifact> (178)	285	179	128	592
	<Mechanical-part> (5)	10	2	9	21
	<Substance> (8)	17	49	103	169
	<Naturally-occurring-phenomena> (57)	64	169	202	435
<Living-organism> (233)					
	<Animal-life> (55)	109	98	29	236
	<Plant> (33)	25	56	48	129
	<Body> (117)	300	416	57	773
	<Aspects-of-living-thing> (27)	28	11	4	43
<PLACE> (224)					
<Constructed-environment> (130)					
	<Water-based-environment> (9)	0	0	22	22
	<Land-based-environment> (109)	86	119	770	975
	<Locale> (11)	10	12	147	169
<Natural-place> (62)					
	<Sky> (2)	8	3	5	16
	<Body-of-water> (17)	1	19	147	167
	<Shore> (4)	7	1	51	59
	<Land-water-formation> (6)	0	3	39	42
	<Terrain> (20)	25	55	99	179
	<Ecosystem> (4)	5	1	60	66
<Sociopolitical-location> (26)					
	<Continent> (2)	0	0	3	3
	<Country> (9)	4	1	12	17
	<State> (4)	0	1	21	22
	<Municipality> (9)	1	0	21	22

Note. Counts for individual concepts are indicated in parentheses following each concept and do not include counts for any subordinate concepts.

By way of comparison, Table 5.17 lists the 65 concepts with a frequency count of 100 or more. The minimum of 100 counts is an arbitrary cut off point but one that provides a reasonable number of concepts to examine concept use within the top level facets. The range of pair occurrence is from 15 to 21. In this list there are 16 concepts (24%) from the <OBJECT> facet, 29 concepts (44%) from the <PROPERTY> facet, and 19 concepts (29%) from the <SPATIAL-LOCATION> facet; but there are only three concepts (3%) from the <PLACE> facet (e.g., <Area>, <Opening> and <Land>). Interestingly, 44% of these high frequency concepts are concentrated in four subfacets: There are 11 concepts from the <Figure> facet; eight from the <Gestalt> facet; four from the <Comparative-Measure> facet; and six from the <Comparison> facet. While <Figure> is nested under the top level facet <OBJECT>, the subfacets <Gestalt>, <Comparative-Measure> and <Comparison> are all nested under the top level facet <PROPERTY>. Table 5.17 also indicates extremes in frequency of use across domain distributions: For example, <Cell> is predominantly used in the domain MIC; <Land>, <Road>, and <Urban> in the domain SAT; and <Letter> in the domain ART.

Table 5.17. *Concepts with frequency counts of 100 or more ordered according to position in the hierarchical structure of the faceted vocabulary.*

Hierarchical Structure	Concept	Pair Total	Freq. Rank	Total Freq. Count	ART Freq. Count	MIC Freq. Count	SAT Freq. Count
<OBJECT>							
<Non-living-thing>							
<Figure>	<Letter>	15	43	233	153	22	58
<Figure>	<Symbol>	19	63	110	50	25	35
<Figure>	<Shape>	21	22	686	333	203	150
<Figure>	<Dot>	21	23	426	241	67	118
<Figure>	<Line>	21	7	1647	830	413	404
<Figure>	<Angle>	21	17	861	282	238	341
<Figure>	<Curve>	19	38	278	104	19	155
<Figure>	*<Closed-curve>*	21	19	776	218	343	215
<Figure>	<Triangle>	20	28	345	182	66	97
<Figure>	<Rectangle>	21	11	1227	694	259	274
<Figure>	<Free-form>	21	45	222	87	72	63
<Substance>	<Liquid>	15	62	120	5	20	95
<Living-organism>							
<Body>	<Cell>	16	54	147	9	138	0
<PLACE>							
<Constructed-environment>							
<Land-based-environment>	<Road>	21	24	359	6	4	349
<Locale>	<Urban>	19	50	162	10	10	142
<Natural-pace>							
<Terrain>	<Land>	20	65	100	1	1	98
<Generic-place>							
<Area>		20	53	149	38	21	90
<Opening>		18	59	139	58	40	41
<PROPERTY>							
<General-concept>							
<Validation>		21	21	701	294	184	223
<Attribute>							
<Gestalt>	<Whole>	21	60	132	42	38	52
<Gestalt>	<Part>	21	27	354	139	78	137
<Gestalt>	<Division>	19	57	140	76	37	27
<Gestalt>	<Section>	19	34	294	151	78	65
<Gestalt>	<Piece>	21	32	315	122	104	89
<Gestalt>	*<Fraction>*	20	13	1053	558	173	322
<Gestalt>	*<Regularity>*	18	51	152	72	38	42
<Gestalt>	<Similarity>	21	3	3059	993	983	1083

Note. The same concepts are listed in Table 5.12. Table 5.17 continued on next page.

Table 5.17 continued.

Hierarchical Structure	Concept	Pair Total	Freq. Rank	Total Freq. Count	ART Freq. Count	MIC Freq. Count	SAT Freq. Count
<Quantity>	*<Linguistic-quantity>*	21	4	2875	1279	607	989
<Quantity>	<Number>	21	14	1021	144	452	425
<Quantity>	*<Measurement-unit>*	17	18	784	357	129	298
<Quantity>	<Size>	21	8	1558	508	495	555
<Quantity>	*<Width>*	21	33	306	122	106	78
<Quantity>	<Length>	21	30	335	205	81	49
<Quantity>	<Height>	21	41	259	97	89	73
<Comparison>	<Increase>	16	49	172	77	48	47
<Comparison>	<Degree>	21	12	1157	415	348	394
<Comparison>	*<Certitude>*	21	9	1361	615	268	478
<Comparison>	*<Probability>*	17	56	143	60	21	62
<Comparison>	*<Unequal>*	21	29	336	118	106	112
<Comparison>	<Emphasis>	20	55	146	42	19	85
<Condition>	<Concentration>	20	42	235	75	102	58
<Condition>	*<Extension>*	21	31	317	159	53	105
<Art-and-craft-process>	<Pattern>	17	61	124	34	40	50
<Art-and-craft-process>	*<By-position>*	20	64	107	1	17	89
<Art-and-craft-process>	<Color>	20	58	140	46	66	28
<Art-and-craft-process>	*<Hue>*	21	20	706	326	175	205
<Art-and-craft-process>	*<Tone>*	21	46	222	67	72	83
<SPATIAL-LOCATION>							
<Position>							
<Indexical>	*<Presentation>*	21	15	980	478	224	278
<Indexical>	*<Rotated>*	19	37	283	141	61	81
<Indexical>	*<General-part>*	21	2	3167	1202	913	1052
<Indexical>	*<Extremity>*	21	10	1345	536	339	470
<Relational>	<Distance>	19	44	232	93	56	83
<Relational>	<Cross>	21	36	285	138	50	97
<Relational>	<Joined>	21	25	359	175	85	99
<Relational>	<Inside-of>	21	6	1773	696	550	527
<Relational>	<Surrounding>	21	48	218	76	37	105
<Relational>	<Outside-of>	21	26	355	144	89	122
<Relational>	<Off>	21	40	271	95	79	97
<Relational>	<On>	21	16	913	357	264	292
<Relational>	<Beside>	20	39	276	94	66	116
<Relational>	<Between>	20	52	151	61	56	34
<Relational>	*<Vertical-perspective>*	21	35	292	121	85	86
<Relational>	<Cover>	18	66	100	30	36	34
<Relational>	<Vertical>	21	5	1823	790	387	646
<Relational>	<Horizontal>	21	1	3200	1370	771	1059
<Relational>	<Diagonal>	17	47	222	61	89	72

Although comparisons of concept occurrence across the faceted hierarchy are interesting and of potential value to developers of CBIR systems, such analyses are beyond the scope of the current research and are therefore left for future investigation. Instead, the question of a shared vocabulary of concepts must be addressed by the broader consideration of whether a concept was used by a preponderance of subjects across multiple domains.

5.3 Shared concepts: Shared-ness rating

The purpose of the research was to assess the extent to which subjects used the same concepts to describe images from more than one domain. Evaluation of the domain and frequency distributions of individual concepts used by subjects may indicate that a concept was used in multiple domains, but it does not indicate that subjects were actually using a shared vocabulary of concepts. In order to determine if the concepts generated by subject pairs were used across domains, a rating scale was devised based on the actual use of each concept by each subject pair. This scale is referred to as the *shared-ness rating*. The shared-ness rating scale is a continuum from 0 to 3: 0 indicates that a given concept was not used at all by a subject pair; 1 indicates the concept was used in only one domain by a subject pair; 2 indicates usage in two domains; and 3 indicates usage in all three domains. This shared-ness rating captures the breadth of each concept's use by each subject pair (i.e., whether a subject pair used a particular concept in all three domains, in two domains, in one domain, or not at all). Appendix F provides a summary of each concept's use by the 21 subject pairs.

Shared-ness ratings provide a general means for measuring subject agreement on concept use. For each concept, a shared-ness measure was computed by summing shared-ness ratings across all 21 subject pairs. For example, the concept <Triangle> was used by five subject pairs across all three domains, by ten subject pairs in two domains, by five subject pairs in one domain, and by one subject pair in no domain (i.e., the concept was not used by any pair). The resulting shared-ness measure for <Triangle> is 41 (i.e., 15 + 20 + 5 + 0 = 40). Concepts with different degrees of shared-ness can be arranged in a rank order according to shared-ness measure (i.e., from 0 to 63). Table 5.18 lists the 150 concepts that were used in more than one domain by more than one subject pair ordered by shared-ness measure. In Table 5.19 these same concepts are ordered by their position in the hierarchical structure.

Table 5.18. *Concepts used in more than one domain and more than one subject pair, ordered by the sum of the domain ratings.*

Concept number	Concept	Pair Count	Freq. Total	Shared-ness rating 0	1	2	3	Sum
338	<Similarity	21	**3059**	0	0	0	21	63
340	*<Linguistic-quantity>*	21	**2174**	0	0	0	21	63
460	<Vertical>	21	**1823**	0	0	0	21	63
442	<Inside-of>	21	**1773**	0	0	0	21	63
346	<Size>	21	**1558**	0	0	0	21	63
355	<Degree>	21	**1157**	0	0	0	21	63
461	<Horizontal>	21	**3200**	0	0	1	20	62
433	*<General-part>*	21	**3167**	0	0	1	20	62
22	<Line>	21	**1647**	0	0	1	20	62
434	*<Extremity>*	21	**1345**	0	0	1	20	62
341	<Number>	21	**1021**	0	0	1	20	62
429	*<Presentation>*	21	**980**	0	0	1	20	62
448	<On>	21	**913**	0	0	1	20	62
23	<Angle>	21	**861**	0	0	1	20	62
34	<Rectangle>	21	**1227**	0	0	2	19	61
356	*<Certitude>*	21	**1361**	0	0	3	18	60
320	*<Validation>*	21	**701**	0	1	2	18	59
20	<Shape>	21	**686**	0	0	6	15	57
328	*<Fraction>*	20	**1053**	1	1	2	17	56
445	<Outside-of>	21	**355**	0	0	7	14	56
347	*<Width>*	21	**306**	0	0	7	14	56
453	*<Vertical-perspective>*	21	**292**	0	1	5	15	56
327	<Piece>	21	**315**	0	0	8	13	55
359	*<Unequal>*	21	**336**	0	2	5	14	54
370	*<Extension>*	21	**317**	0	3	3	15	54
450	<Beside>	20	**276**	1	1	5	14	53
21	<Dot>	21	**426**	0	2	6	13	53
324	<Part>	21	**354**	0	1	8	12	53
348	<Length>	21	**335**	0	3	4	14	53
417	*<Hue>*	21	**706**	0	1	9	11	52
440	<Joined>	21	**359**	0	2	7	12	52
435	<Distance>	19	**232**	2	0	5	14	52
431	*<Rotated>*	19	**283**	2	2	3	14	50

Note. Concept number indicates concept position in the faceted hierarchy and has been introduced to facilitate comparison with Table 5.19. Concepts with frequency counts of more than 100 are indicated by bold. For each concept a shared-ness measure (Sum) has been computed by summing shared-ness ratings across all subject pairs. See complete list of shared-ness measures for all concepts in Appendix F. Table 5.18 continued on following three pages.

Table 5.18 continued.

Concept number	Concept	Pair Count	Freq. Total	Shared-ness Ratings				
				0	1	2	3	Sum
362	<Concentration>	20	**235**	1	3	4	13	50
447	<Off>	21	**271**	0	4	6	11	49
444	<Surrounding>	21	**218**	0	3	8	10	49
462	<Diagonal>	17	**222**	3	1	4	13	48
25	*<Closed-curve>*	21	**776**	1	1	10	9	48
422	*<Tone>*	21	**222**	0	3	9	9	48
349	<Height>	21	**259**	0	6	4	11	47
323	<Whole>	21	**132**	0	4	8	9	47
37	<Free-form>	21	**222**	0	4	10	7	45
332	*<Regularity>*	18	**152**	3	2	5	11	45
451	<Between>	20	**151**	2	2	8	9	45
326	<Section>	19	**294**	2	3	7	9	44
345	*<Measurement-unit>*	17	**784**	3	5	1	12	43
314	*<Opening>*	18	**139**	2	4	8	7	41
32	<Triangle>	20	**345**	1	5	10	5	40
360	<Emphasis>	20	**146**	1	8	4	8	40
357	*<Probability>*	17	**143**	3	2	10	6	40
415	<Color>	20	**140**	1	5	11	4	39
439	<Cross>	21	**285**	0	4	16	1	39
24	<Curve>	19	**278**	2	5	9	5	38
325	<Division>	19	**140**	2	7	6	6	37
318	*<Existence>*	19	79	2	7	8	4	35
313	<Area>	20	**149**	1	11	4	5	34
18	<Symbol>	19	**110**	2	6	12	1	33
455	<Cover>	18	**100**	3	6	10	2	32
454	<Layer>	16	88	4	5	9	3	32
404	<Cut>	19	85	2	9	7	3	32
438	<Through>	17	63	2	9	7	3	32
352	<Increase>	16	**172**	4	6	8	3	31
403	<Pattern>	17	**124**	4	6	9	2	30
412	*<By-position>*	20	**107**	1	12	7	1	29
372	*<Protrusion>*	16	75	5	6	7	3	29
307	*<Land-formation>*	16	92	6	6	5	4	28
365	*<Clarity>*	14	55	6	4	10	1	27
358	<Equal>	16	51	5	9	3	4	27
459	*<Pointing-to>*	16	54	4	9	6	2	27
411	<View>	17	66	3	11	5	2	27
368	*<Sharpness>*	15	89	6	8	3	4	26
336	*<Alignment>*	16	65	5	9	4	3	26
458	<Direction>	15	43	6	5	9	1	26

Note. Table 5.18 continued on next page.

Table 5.18 continued.

Concept number	Concept	Pair Count	Freq. Total	Shared-ness Ratings				
				0	1	2	3	Sum
14	<Letter>	15	**233**	6	7	6	2	25
456	*<Whole-coverage>*	15	62	8	5	4	4	25
269	<Road>	21	**359**	0	17	4	0	25
408	<Contour>	15	59	6	8	5	2	24
397	<Modern-art>	17	72	4	11	5	1	24
333	*<Ordered>*	17	44	4	12	3	2	24
232	*<Residential-building>*	18	59	2	14	5	0	24
6	<Photograph>	19	55	2	14	5	0	24
119	<Liquid>	15	**120**	5	9	7	0	23
452	*<Horizontal-perspective>*	16	46	4	11	6	0	23
287	<Urban>	19	**162**	2	15	4	0	23
425	<Shaded>	13	71	8	6	5	2	22
321	<Visible>	14	28	6	9	5	1	22
231	<Building>	19	86	2	16	3	0	22
319	*<Domain>*	13	61	8	7	4	2	21
391	*<Fluid-movement>*	15	49	7	8	5	1	21
390	*<Violent-movement--non-fluid>*	16	39	5	12	3	1	21
169	<Head>	17	80	4	14	2	1	21
388	*<Movement-reflecting-growth>*	11	61	10	4	5	2	20
369	*<Smoothness>*	9	24	10	5	3	3	20
449	<Facing>	13	37	8	7	5	1	20
191	*<Non-human-animal-appendage>*	14	40	7	9	4	1	20
55	*<Viewing>*	16	73	5	12	4	0	20
126	*<Atmosphere>*	11	31	10	5	5	1	18
139	<Wave>	11	76	9	7	4	1	18
381	*<Abnormal>*	12	52	7	11	2	1	18
308	*<Eco-system>*	16	66	6	13	1	1	18
334	*<Grouped>*	11	33	9	8	3	1	17
305	<Mountain>	14	47	8	10	2	1	17
337	*<Frequency>*	11	30	8	9	4	0	17
361	<Consistency>	13	29	8	9	4	0	17
376	*<Perfection>*	12	28	8	10	2	1	17
384	*<Damaged>*	11	32	7	11	3	0	17
430	<Orientation>	8	25	13	2	4	2	16
29	<Arc>	8	54	12	5	1	3	16
401	*<Production-media>*	12	70	9	9	2	1	16
227	<PLACE>	12	27	9	9	2	1	16
457	*<Partial-coverage>*	9	23	12	6	0	3	15
117	<Substance>	9	27	11	6	3	1	15

Note. Table 5.18 continued on next page.

Table 5.18 continued.

Concept		Pair	Freq.	Shared-ness Ratings				
number	Concept	Count	Total	0	1	2	3	Sum
437	<Intersecting>	11	41	10	8	2	1	15
343	*<Specific-dimension>*	9	24	11	6	4	0	14
339	*<Quantity>*	9	22	11	6	4	0	14
329	*<Closure>*	11	20	10	8	3	0	14
93	<Fiber>	9	20	10	9	1	1	14
12	<Sign>	8	59	13	4	3	1	13
187	<Finger>	9	37	13	4	3	1	13
1	<OBJECT>	8	20	13	4	3	1	13
354	*<Rate-of-change>*	8	23	12	5	4	0	13
39	<Block>	2	2	12	6	2	1	13
389	*<Random-movement>*	9	27	11	7	3	0	13
28	<Curl>	9	27	11	8	1	1	13
331	*<Simplicity>*	9	20	11	7	3	0	13
258	<Wall>	7	43	14	4	1	2	12
17	*<Punctuation>*	8	35	13	4	4	0	12
335	<Random>	8	18	13	4	4	0	12
294	<River>	9	77	12	6	3	0	12
182	<Arm>	9	55	12	7	1	1	12
165	*<Vegetable>*	9	45	12	6	3	0	12
423	*<Brightness>*	9	34	12	6	3	0	12
40	<Sphere>	7	55	14	3	4	0	11
373	<Indentation>	8	28	13	6	1	1	11
432	<Reversed>	8	22	12	7	2	0	11
310	*<Country>*	8	17	12	7	2	0	11
289	<Ocean>	9	14	12	7	2	0	11
402	<Design>	6	16	15	3	2	1	10
464	*<Compass-orientation>*	9	62	13	6	2	0	10
205	<Blood-vessel>	8	28	13	6	2	0	10
375	*<Neatness>*	7	11	13	6	2	0	10
194	<Hair>	6	28	15	4	1	1	9
118	<Solid>	5	13	15	4	1	1	9
225	<Artist>	7	19	14	5	2	0	9
9	*<Film>*	6	10	14	5	2	0	9
374	*<Goodness>*	7	11	15	4	2	0	8
342	<Dimension>	5	21	17	2	1	1	7
322	<Invisible>	8	16	16	3	2	0	7
303	<Valley>	3	6	18	1	2	0	5
428	<Position>	3	5	18	1	2	0	5

Table 5.19. *Concepts used in more than one domain and by more than one subject pair ordered by position in the hierarchical structure of the faceted vocabulary.*

Concept number	Concept	Pair Count	Freq. Total	Shared-ness rating 0	1	2	3	Sum
1	<OBJECT>	8	20	13	4	3	1	13
6	<Photograph>	19	55	2	14	5	0	24
9	*<Film>*	6	10	14	5	2	0	9
12	<Sign>	8	59	13	4	3	1	13
14	<Letter>	15	**233**	6	7	6	2	25
17	*<Punctuation>*	8	35	13	4	4	0	12
18	<Symbol>	19	**110**	2	6	12	1	33
20	<Shape>	21	**686**	0	0	6	15	57
21	<Dot>	21	**426**	0	2	6	13	53
22	<Line>	21	**1647**	0	0	1	20	62
23	<Angle>	21	**861**	0	0	1	20	62
24	<Curve>	19	**278**	2	5	9	5	38
25	*<Closed-curve>*	21	**776**	1	1	10	9	48
28	<Curl>	9	27	11	8	1	1	13
29	<Arc>	8	54	12	5	1	3	16
32	<Triangle>	20	**345**	1	5	10	5	40
34	<Rectangle>	21	**1227**	0	0	2	19	61
37	<Free-form>	21	**222**	0	4	10	7	45
39	<Block>	2	2	12	6	2	1	13
40	<Sphere>	7	55	14	3	4	0	11
55	*<Viewing>*	16	73	5	12	4	0	20
93	<Fiber>	9	20	10	9	1	1	14
117	<Substance>	9	27	11	6	3	1	15
118	<Solid>	5	13	15	4	1	1	9
119	<Liquid>	15	**120**	5	9	7	0	23
126	*<Atmosphere>*	11	31	10	5	5	1	18
139	<Wave>	11	76	9	7	4	1	18
165	*<Vegetable>*	9	45	12	6	3	0	12
169	<Head>	17	80	4	14	2	1	21
182	<Arm>	9	55	12	7	1	1	12
187	<Finger>	9	37	13	4	3	1	13
191	*<Non-human-animal-appendage>*	14	40	7	9	4	1	20
194	<Hair>	6	28	15	4	1	1	9
205	<Blood-vessel>	8	28	13	6	2	0	10
225	<Artist>	7	19	14	5	2	0	9

Note. Concepts are the same as Table 5.18. Concept number indicates the position of a concept in the hierarchical structure of the faceted vocabulary. Table 5.19 continued on next three pages.

Table 5.19 continued.

Concept number	Concept	Pair Count	Freq. Total	Shared-ness Ratings 0	1	2	3	Sum
227	<PLACE>	12	27	9	9	2	1	16
231	<Building>	19	86	2	16	3	0	22
232	*<Residential-building>*	18	59	2	14	5	0	24
258	<Wall>	7	43	14	4	1	2	12
269	<Road>	21	**359**	0	17	4	0	25
287	<Urban>	19	**162**	2	15	4	0	23
289	<Ocean>	9	14	12	7	2	0	11
294	<River>	9	77	12	6	3	0	12
303	<Valley>	3	6	18	1	2	0	5
305	<Mountain>	14	47	8	10	2	1	17
307	*<Land-formation>*	16	92	6	6	5	4	28
308	*<Eco-system>*	16	66	6	13	1	1	18
310	*<Country>*	8	17	12	7	2	0	11
313	<Area>	20	**149**	1	11	4	5	34
314	*<Opening>*	18	**139**	2	4	8	7	41
	<PROPERTY>							
318	*<Existence>*	19	79	2	7	8	4	35
319	*<Domain>*	13	61	8	7	4	2	21
320	*<Validation>*	21	**701**	0	1	2	18	59
321	<Visible>	14	28	6	9	5	1	22
322	<Invisible>	8	16	16	3	2	0	7
323	<Whole>	21	**132**	0	4	8	9	47
324	<Part>	21	**354**	0	1	8	12	53
325	<Division>	19	**140**	2	7	6	6	37
326	<Section>	19	**294**	2	3	7	9	44
327	<Piece>	21	**315**	0	0	8	13	55
328	*<Fraction>*	20	**1053**	1	1	2	17	56
329	*<Closure>*	11	20	10	8	3	0	14
331	*<Simplicity>*	9	20	11	7	3	0	13
332	*<Regularity>*	18	**152**	3	2	5	11	45
333	*<Ordered>*	17	44	4	12	3	2	24
334	*<Grouped>*	11	33	9	8	3	1	17
335	<Random>	8	18	13	4	4	0	12
336	*<Alignment>*	16	65	5	9	4	3	26
337	*<Frequency>*	11	30	8	9	4	0	17
338	<Similarity	21	**3059**	0	0	0	21	63
339	*<Quantity>*	9	22	11	6	4	0	14
340	*<Linguistic-quantity>*	21	**2174**	0	0	0	21	63
341	<Number>	21	**1021**	0	0	1	20	62
342	<Dimension>	5	21	17	2	1	1	7

Note. Table 5.19 continued on next page.

Table 5.19 continued.

Concept number	Concept	Pair Count	Freq. Total	Shared-ness Ratings				
				0	1	2	3	Sum
343	*<Specific-dimension>*	9	24	11	6	4	0	14
345	*<Measurement-unit>*	17	**784**	3	5	1	12	43
346	<Size>	21	**1558**	0	0	0	21	63
347	*<Width>*	21	**306**	0	0	7	14	56
348	<Length>	21	**335**	0	3	4	14	53
349	<Height>	21	**259**	0	6	4	11	47
352	<Increase>	16	**172**	4	6	8	3	31
354	*<Rate-of-change>*	8	23	12	5	4	0	13
355	<Degree>	21	**1157**	0	0	0	21	63
356	*<Certitude>*	21	**1361**	0	0	3	18	60
357	*<Probability>*	17	**143**	3	2	10	6	40
358	<Equal>	16	51	5	9	3	4	27
359	*<Unequal>*	21	**336**	0	2	5	14	54
360	<Emphasis>	20	**146**	1	8	4	8	40
361	<Consistency>	13	29	8	9	4	0	17
362	<Concentration>	20	**235**	1	3	4	13	50
365	*<Clarity>*	14	55	6	4	10	1	27
368	*<Sharpness>*	15	89	6	8	3	4	26
369	*<Smoothness>*	9	24	10	5	3	3	20
370	*<Extension>*	21	**317**	0	3	3	15	54
372	*<Protrusion>*	16	75	5	6	7	3	29
373	<Indentation>	8	28	13	6	1	1	11
374	*<Goodness>*	7	11	15	4	2	0	8
375	*<Neatness>*	7	11	13	6	2	0	10
376	*<Perfection>*	12	28	8	10	2	1	17
381	*<Abnormal>*	12	52	7	11	2	1	18
384	*<Damaged>*	11	32	7	11	3	0	17
388	*<Movement-reflecting-growth>*	11	61	10	4	5	2	20
389	*<Random-movement>*	9	27	11	7	3	0	13
390	*<Violent-movement--non-fluid>*	16	39	5	12	3	1	21
391	*<Fluid-movement>*	15	49	7	8	5	1	21
397	<Modern-art>	17	72	4	11	5	1	24
401	*<Production-media>*	12	70	9	9	2	1	16
402	<Design>	6	16	15	3	2	1	10
403	<Pattern>	17	**124**	4	6	9	2	30
404	<Cut>	19	85	2	9	7	3	32
408	<Contour>	15	59	6	8	5	2	24
411	<View>	17	66	3	11	5	2	27
412	*<By-position>*	20	**107**	1	12	7	1	29

Note. Table 5.19 continued on next page.

Table 5.19 continued.

Concept number	Concept	Pair Count	Freq. Total	Shared-ness Ratings				
				0	1	2	3	Sum
415	<Color>	20	**140**	1	5	11	4	39
417	*<Hue>*	21	**706**	0	1	9	11	52
422	*<Tone>*	21	**222**	0	3	9	9	48
423	*<Brightness>*	9	34	12	6	3	0	12
425	<Shaded>	13	71	8	6	5	2	22
	<SPATIAL-LOCATION>							
428	<Position>	3	5	18	1	2	0	5
429	*<Presentation>*	21	**980**	0	0	1	20	62
430	<Orientation>	8	25	13	2	4	2	16
431	*<Rotated>*	19	**283**	2	2	3	14	50
432	<Reversed>	8	22	12	7	2	0	11
433	*<General-part>*	21	**3167**	0	0	1	20	62
434	*<Extremity>*	21	**1345**	0	0	1	20	62
435	<Distance>	19	**232**	2	0	5	14	52
437	<Intersecting>	11	41	10	8	2	1	15
438	<Through>	17	63	2	9	7	3	32
439	<Cross>	21	**285**	0	4	16	1	39
440	<Joined>	21	**359**	0	2	7	12	52
442	<Inside-of>	21	**1773**	0	0	0	21	63
444	<Surrounding>	21	**218**	0	3	8	10	49
445	<Outside-of>	21	**355**	0	0	7	14	56
447	<Off>	21	**271**	0	4	6	11	49
448	<On>	21	**913**	0	0	1	20	62
449	<Facing>	13	37	8	7	5	1	20
450	<Beside>	20	**276**	1	1	5	14	53
451	<Between>	20	**151**	2	2	8	9	45
452	*<Horizontal-perspective>*	16	46	4	11	6	0	23
453	*<Vertical-perspective>*	21	**292**	0	1	5	15	56
454	<Layer>	16	88	4	5	9	3	32
455	<Cover>	18	**100**	3	6	10	2	32
456	*<Whole-coverage>*	15	62	8	5	4	4	25
457	*<Partial-coverage>*	9	23	12	6	0	3	15
458	<Direction>	15	43	6	5	9	1	26
459	*<Pointing-to>*	16	54	4	9	6	2	27
460	<Vertical>	21	**1823**	0	0	0	21	63
461	<Horizontal>	21	**3200**	0	0	1	20	62
462	<Diagonal>	17	**222**	3	1	4	13	48
464	*<Compass-orientation>*	9	62	13	6	2	0	10

The rank ordering of concepts by shared-ness measure is important because it demonstrates which concepts were more broadly shared across subject pairs. The higher the shared-ness measure of a concept, the more likely it is that a concept is part of a shared vocabulary for describing the internal contextuality of images. For example, the concepts <Similarity>, <Size> and <Degree> all have a shared-ness measure of 63, which demonstrates their importance in a vocabulary for the description of images. In contrast, <Valley> and <Film> both have a shared-ness measure of 23, indicating that these two concepts do not have particular relevance for a cross-domain vocabulary.

Determining if subject pairs agree on concepts used in all three domains is a problem of interrater reliability. The scale used to assign shared-ness ratings for concepts (i.e., 0 to 3) represents the distribution of concept use across image domains during the entire describe-draw session for a single subject pair and thus provides a scoring rubric similar to that used when judges are knowingly performing a rating task.

According to Stemler (2004), there are three types of methods used for estimating interrater reliability: consensus, consistency and measurement. Consensus is most useful with nominal data and commonly measures agreement on application of a scoring rubric as a percentage. Consistency is a within judge measure and reflects how consistently a judge applies the scoring rubric. Measurement considers data from all judges to compute an estimate of reliability for a single judge. Because measurement approaches emphasize the reliability of individual judges and require that there be more raters than items rated[69], this approach was not appropriate for determining the existence of a shared vocabulary. Methods for measuring consensus were also ruled out because agreement is computed as a simple percentage and thus cannot establish statistical significance. Furthermore, as Stemler (2004) observes, "if the behavior of interest has a low incidence of occurrence in the population, then it is possible to get artificially inflated percent agreement figures simply because most of the values fall under one category of the rating scale" (Popular Methods of computing Consensus Estimates, paragraph 2).

Stemler (2004) contends that "consistency estimates of interrater reliability are based upon the assumption that it is not really necessary for two judges to share a common meaning of the rating scale, so long as each judge is consistent in classifying the phenomena according to his or her own definition of the scale" (Consistency Estimates, General Description, paragraph 1). In this study, the question was whether subjects (judges), who are not necessarily aware of an image's domain, are assigning concepts to the image (or classifying the phenomena) based on a common vocabulary of concepts.

[69] In this study, there were 465 items but only 21 subject pairs (raters).

Agreement across subject pairs would indicate the presence of a vocabulary of concepts that was applicable across image domains.

The most appropriate indicator of interrater reliability that can be computed for this data would be the Spearman rank correlation coefficient. Although the Pearson correlation coefficient is commonly used, it requires that data be normally distributed, which is not the case here since many concepts were used by only one subject pair. According to Stemler (2004), the Spearman rank correlation coefficient offers "an approximation of the Pearson correlation coefficient, but may be used in circumstances where the data under investigation are not normally distributed" (Popular Methods for Computing Consistency Estimates, paragraph 3).

Spearman rank correlation coefficients were computed for the set of 465 concepts that remained after collapsing of isolates and removal of the 80 concepts introduced during vocabulary construction and used by subject pairs. For each subject pair, sharedness ratings (i.e., 0, 1, 2, or 3) were assigned for each of the 465 concepts (see Appendix F and Table 5.20). A Spearman rank correlation coefficient was then computed for each combination of two subject pairs (e.g., P1 and P2). This produced a matrix of 210 correlations, one for each possible combination of the 21 subject pairs (see Table 5.21, located at the end of this chapter). Each correlation indicates the extent to which two subject pairs agreed on the overall use of concepts across domains.

Table 5.20. *Occurrences of each shared-ness rating by each subject pair.*

		Subject pair										
		21	**20**	**19**	**18**	**17**	**16**	**15**	**14**	**13**	**12**	**11**
Rating	**0**	264	266	310	306	372	341	297	296	270	251	184
	1	111	93	76	86	56	77	101	97	111	132	148
	2	42	46	28	18	19	24	30	49	36	36	58
	3	48	60	51	55	18	23	37	23	48	46	75

		Subject pair									
		10	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**
Rating	**0**	341	314	329	377	316	304	323	299	283	311
	1	65	74	80	51	72	90	67	86	105	81
	2	35	29	26	15	40	28	32	37	38	31
	3	24	48	30	22	37	43	43	43	39	42

Note. Ratings are for 465 concepts.

According to Cohen (1988), Spearman correlation values of .30 to .49 are of medium magnitude and values of .50 to 1.0 are of large magnitude. The correlations computed for subject pairs using the shared-ness rating for 465 unique concepts ranged from a low of .535 to a high of .821, making them all large in magnitude. The majority of correlations below .612 are associated with subject pair 17, and the highest correlation (.821) is associated with P1 and P2. A mean correlation was also computed for all 210 subject pair combinations to measure overall agreement in regard to shared-ness. The mean correlation computed for the 465 concepts was .70 and thus of large magnitude.

Statistical significance means that the degree of correlation observed is unlikely to have occurred on the basis of chance alone. As part of the correlation analysis, SPSS performed a series of 2-tailed *t* tests to assess statistical significance of the correlations and reported the probability value for each correlation coefficient. SPSS automatically reports any *p* value below .001 as .000. Therefore, based on the SPSS output in Table 5.21, all correlations achieved statistical significance at the .001 level.[70] Thus, the results of the correlation analyses indicate that subject pairs do tend to agree in the extent to which they use certain concepts across multiple domains. Concepts with high shared-ness measures, as reported in Table 5.18, demonstrate some confidence that most of the subject pairs used these concepts to describe multiple domains, thus pointing toward a shared vocabulary. The 14 concepts with shared-ness sums of 62 or 63 would form the heart of this shared vocabulary.

[70] It should be noted, however, that the statistical significance of these correlations may be, in part, a function of the large sample size of 465 concepts.

Table 5.21. *Spearman rank order correlations computed for subject pairs across the set of 465 concepts*.

	Pair Number	21	20	19	18	17	16	15
Correlation	**21**	1	.711(**)	.735(**)	.696(**)	.615(**)	.678(**)	.677(**)
p =		.	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**20**	.711(**)	1	.716(**)	.708(**)	.585(**)	.642(**)	.685(**)
p =		0.000	.	0.000	0.000	0.000	0.000	0.000
Correlation	**19**	.735(**)	.716(**)	1	.789(**)	.625(**)	.689(**)	.723(**)
p =		0.000	0.000	.	0.000	0.000	0.000	0.000
Correlation	**18**	.696(**)	.708(**)	.789(**)	1	.638(**)	.708(**)	.729(**)
p =		0.000	0.000	0.000	.	0.000	0.000	0.000
Correlation	**17**	.615(**)	.585(**)	.625(**)	.638(**)	1	.730(**)	.629(**)
p =		0.000	0.000	0.000	0.000	.	0.000	0.000
Correlation	**16**	.678(**)	.642(**)	.689(**)	.708(**)	.730(**)	1	.706(**)
p =		0.000	0.000	0.000	0.000	0.000	.	0.000
Correlation	**15**	.677(**)	.685(**)	.723(**)	.729(**)	.629(**)	.706(**)	1
p =		0.000	0.000	0.000	0.000	0.000	0.000	.
Correlation	**14**	.644(**)	.661(**)	.709(**)	.720(**)	.571(**)	.661(**)	.690(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**13**	.685(**)	.691(**)	.780(**)	.781(**)	.613(**)	.670(**)	.691(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**12**	.698(**)	.675(**)	.728(**)	.740(**)	.596(**)	.683(**)	.709(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**11**	.624(**)	.652(**)	.663(**)	.667(**)	.535(**)	.579(**)	.623(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**10**	.686(**)	.607(**)	.723(**)	.691(**)	.678(**)	.659(**)	.649(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**9**	.698(**)	.678(**)	.765(**)	.764(**)	.628(**)	.672(**)	.732(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**8**	.692(**)	.702(**)	.737(**)	.745(**)	.670(**)	.699(**)	.721(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**7**	.616(**)	.578(**)	.687(**)	.702(**)	.712(**)	.759(**)	.662(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**6**	.692(**)	.642(**)	.737(**)	.765(**)	.603(**)	.664(**)	.703(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**5**	.700(**)	.675(**)	.756(**)	.720(**)	.584(**)	.645(**)	.714(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**4**	.701(**)	.646(**)	.737(**)	.787(**)	.613(**)	.671(**)	.696(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**3**	.719(**)	.664(**)	.712(**)	.689(**)	.633(**)	.695(**)	.709(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**2**	.732(**)	.685(**)	.742(**)	.739(**)	.583(**)	.696(**)	.734(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**1**	.698(**)	.697(**)	.746(**)	.754(**)	.592(**)	.690(**)	.710(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note. N = 465. Mean correlation = .70. All correlations are significant at the 0.001 level (2-tailed). SPSS reports a *p* value of .000 if the significance of a correlation drops below .001. Table 5.21 is continued on next two pages.

Table 5.21 continued.

	Pair Number	14	13	12	11	10	9	8
Correlation	**21**	.644(**)	.685(**)	.698(**)	.624(**)	.686(**)	.698(**)	.692(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**20**	.661(**)	.691(**)	.675(**)	.652(**)	.607(**)	.678(**)	.702(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**19**	.709(**)	.780(**)	.728(**)	.663(**)	.723(**)	.765(**)	.737(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**18**	.720(**)	.781(**)	.740(**)	.667(**)	.691(**)	.764(**)	.745(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**17**	.571(**)	.613(**)	.596(**)	.535(**)	.678(**)	.628(**)	.670(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**16**	.661(**)	.670(**)	.683(**)	.579(**)	.659(**)	.672(**)	.699(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**15**	.690(**)	.691(**)	.709(**)	.623(**)	.649(**)	.732(**)	.721(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**14**	1	.728(**)	.638(**)	.612(**)	.641(**)	.670(**)	.696(**)
p =		.	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**13**	.728(**)	1	.707(**)	.682(**)	.677(**)	.721(**)	.697(**)
p =		0.000	.	0.000	0.000	0.000	0.000	0.000
Correlation	**12**	.638(**)	.707(**)	1	.621(**)	.702(**)	.718(**)	.711(**)
p =		0.000	0.000	.	0.000	0.000	0.000	0.000
Correlation	**11**	.612(**)	.682(**)	.621(**)	1	.620(**)	.658(**)	.642(**)
p =		0.000	0.000	0.000	.	0.000	0.000	0.000
Correlation	**10**	.641(**)	.677(**)	.702(**)	.620(**)	1	.675(**)	.680(**)
p =		0.000	0.000	0.000	0.000	.	0.000	0.000
Correlation	**9**	.670(**)	.721(**)	.718(**)	.658(**)	.675(**)	1	.732(**)
p =		0.000	0.000	0.000	0.000	0.000	.	0.000
Correlation	**8**	.696(**)	.697(**)	.711(**)	.642(**)	.680(**)	.732(**)	1
p =		0.000	0.000	0.000	0.000	0.000	0.000	.
Correlation	**7**	.592(**)	.677(**)	.648(**)	.550(**)	.666(**)	.665(**)	.694(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**6**	.662(**)	.718(**)	.727(**)	.699(**)	.721(**)	.737(**)	.718(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**5**	.670(**)	.722(**)	.682(**)	.604(**)	.680(**)	.717(**)	.705(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**4**	.694(**)	.714(**)	.683(**)	.608(**)	.676(**)	.724(**)	.703(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**3**	.692(**)	.714(**)	.665(**)	.629(**)	.678(**)	.726(**)	.689(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**2**	.696(**)	.697(**)	.734(**)	.639(**)	.706(**)	.728(**)	.703(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**1**	.660(**)	.692(**)	.721(**)	.618(**)	.662(**)	.724(**)	.712(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000

Note. Table 5.21 continued on next page.

Table 5.21 continued.

	Pair Number	7	6	5	4	3	2	1
Correlation	**21**	.616(**)	.692(**)	.700(**)	.701(**)	.719(**)	.732(**)	.698(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**20**	.578(**)	.642(**)	.675(**)	.646(**)	.664(**)	.685(**)	.697(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**19**	.687(**)	.737(**)	.756(**)	.737(**)	.712(**)	.742(**)	.746(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**18**	.702(**)	.765(**)	.720(**)	.787(**)	.689(**)	.739(**)	.754(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**17**	.712(**)	.603(**)	.584(**)	.613(**)	.633(**)	.583(**)	.592(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**16**	.759(**)	.664(**)	.645(**)	.671(**)	.695(**)	.696(**)	.690(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**15**	.662(**)	.703(**)	.714(**)	.696(**)	.709(**)	.734(**)	.710(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**14**	.592(**)	.662(**)	.670(**)	.694(**)	.692(**)	.696(**)	.660(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**13**	.677(**)	.718(**)	.722(**)	.714(**)	.714(**)	.697(**)	.692(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**12**	.648(**)	.727(**)	.682(**)	.683(**)	.665(**)	.734(**)	.721(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**11**	.550(**)	.699(**)	.604(**)	.608(**)	.629(**)	.639(**)	.618(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**10**	.666(**)	.721(**)	.680(**)	.676(**)	.678(**)	.706(**)	.662(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**9**	.665(**)	.737(**)	.717(**)	.724(**)	.726(**)	.728(**)	.724(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**8**	.694(**)	.718(**)	.705(**)	.703(**)	.689(**)	.703(**)	.712(**)
p =		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**7**	1	.650(**)	.640(**)	.670(**)	.680(**)	.637(**)	.672(**)
p =		.	0.000	0.000	0.000	0.000	0.000	0.000
Correlation	**6**	.650(**)	1	.716(**)	.760(**)	.683(**)	.755(**)	.716(**)
p =		0.000	.	0.000	0.000	0.000	0.000	0.000
Correlation	**5**	.640(**)	.716(**)	1	.691(**)	.696(**)	.718(**)	.719(**)
p =		0.000	0.000	.	0.000	0.000	0.000	0.000
Correlation	**4**	.670(**)	.760(**)	.691(**)	1	.712(**)	.744(**)	.735(**)
p =		0.000	0.000	0.000	.	0.000	0.000	0.000
Correlation	**3**	.680(**)	.683(**)	.696(**)	.712(**)	1	.759(**)	.672(**)
p =		0.000	0.000	0.000	0.000	.	0.000	0.000
Correlation	**2**	.637(**)	.755(**)	.718(**)	.744(**)	.759(**)	1	.821(**)
p =		0.000	0.000	0.000	0.000	0.000	.	0.000
Correlation	**1**	.672(**)	.716(**)	.719(**)	.735(**)	.672(**)	.821(**)	1
p =		0.000	0.000	0.000	0.000	0.000	0.000	.

CHAPTER 6
Discussion

6.1 The descriptive process

In the experimental situation, a pair of subjects took turns orally describing images to each without benefit of communicative feedback during the description process. Results of the shared communication studies conducted by Krauss and Fussell (Krauss & Fussell, 1991; see also Krauss & Weinheimer, 1964, 1967) suggest that the need to communicate in the experimental situation requires subjects to use more general words that would be readily understood. The images used in Krauss and Fussell's experiments were selected on the basis of their lack of identifiable objects under the assumption that they would be described in terms of familiar geometric elements that are part of a shared communicative environment. Following this argument that a communicative environment would be indicated by a shared vocabulary of terms and concepts, it was anticipated that the natural language vocabulary used by subject pairs in the current research, if shared, could provide the basis for a controlled vocabulary to describe the internal contextuality of images. More recent studies in referential communication indicate that shared knowledge is developed over the course of an interaction (Galantucci, 2005; Lau & Chui, 2001; Nohara-LeClair, 2001). This points to the possibility that, over the course of 14 descriptions, a subject pair would develop a shared vocabulary, in particular because communication was allowed between subjects after each description task had been completed. Although dialog between subjects was initiated by the researcher's questions about the ease or difficulty of the describe-draw task (see Appendix O), subjects explored a wider range of issues. Examples of this dialog include:

> I wasn't prepared for how hard this would be, can we assume if the person says repeat, repeat, repeat that they don't understand and try a different approach? (P9.2-sum)

> Good description, what mattered was the concept not the details. (P13.1)

> Wanted you to draw a big upsilon but didn't know if you'd know the Greek letter. (P3.2-sum)

> So many lines and squiggles, I was trying to pick the most prominent ones. (P15.1-sum)
>
> Scallop is a better word than concavity. (P14.2-sum)
>
> I should have described these crystals as being diagonal in a parallel fashion. Should have told you how many were diagonal and how many were standing up. (P20.1-sum)
>
> I should have said stylized. I should have said the pickets dominate. (P12.2-sum)
>
> Don't know what a sea-urchin is like so I drew a blob. (P4.2-sum)
>
> Got it right but not the size. I had a whale, she had a minnow. (P7.1-sum)

There was some terminological confusion evident during the experimental task which may have affected frequency counts. For example, a subject might interchange the word rectangle for triangle or say left when she meant right. Other observations during the describe-draw task, as well as subject's comments on descriptive approaches and terminological choices, indicate that terms for geometric elements may not be assumed to be part of a shared descriptive vocabulary. For example, when a subject says It's hard talking about things in geometric terms and distances, there's no 'impression' I can give you, I just had to do it geometrically (P21.1-sum), it indicates that, while geometric words may be assumed to be part of a general vocabulary, there is some discomfort in using them.

Frequently, it appeared to be difficult for subject pairs to estimate or communicate size. If the describer used quarter ($.25) as a referent for size, the drawer might make a dime-size shape (e.g., P9 and P13); or the describer would say to go two inches, but the drawer would only go one inch (P7). Other describers used referents for size that were more successful, such as pea or peanut (P2) or the index-finger (P7). Many size references were in terms of fractions of the page, but that also created confusion because the frame of reference used by the drawer was sometimes a rectangle and sometimes the whole page. In addition, all subjects tended to use about or approximate to describe size and position references (e.g., about-dead-center). Did the subject mean to communicate exactly, or was it a communication hedge because she could not actually measure and was insecure about her description? Studying the descriptive strategies subject's use with a particular focus on how size is referenced could extend current CBIR interface design to assist searchers by including referential models, whether terminological, visual or both.

The oral nature of the describe-draw task introduces the problem of implied words that could not be captured by the transcription. For example, transcription of One straight line and one kind-of curved (P5.1#106) fails to capture the implied repetition of the word line, and the second close in Don't close it … the oval shape does (P7.1#200) is similarly lost. Thus actual frequency counts for these and other concepts may be misleading as to their relative importance or shared-ness. The communicative aspect of the describe-draw task suggests that selection of natural language words was more strongly influenced by the need to communicate than by descriptive accuracy. This raises a question as to whether an orally generated vocabulary in a shared communicative environment would be differ from a written vocabulary in a single subject situation; and that question is worthy of future research.

Frequency counts as a basis for determining shared-ness of vocabulary may be skewed by the loss of implied words and also by the fact that some subjects were more verbose than others. Some subjects went into great detail, and some selected only the most dominant features to describe. For example, in the description of a satellite photograph, the describer said draw what you would see flying over Indianapolis (P16.2#7), a total of eight words containing five stop words (i.e., DRAW, WHAT, YOU, WOULD, SEE). In contrast, another subject said this is an aerial view of a city (P11.2#10), a total of eight words but containing only three stop words (i.e., THIS, IS, OF). And then there was the subject who dispensed with full sentence structure and described the same concept with no stop words: another aerial view (P10.2#12). Referential communication studies suggest that lack of feedback generally increases the number of words used by speakers when encoding referents (Lau & Chui, 2001). In this research situation, there was no feedback during the actual describing process, yet the result was a range of 12 to 826 words per description. Individual differences may account for the wide discrepancy in numbers of words used to describe a single image, as well as descriptive strategies, but this is a question to be explored in future research. However, it is worth noting that raw word-for-word counts may be affected by varied communication styles as much as descriptive variation.

6.2 Terms

Term frequency counts and pair occurrences were used to evaluate the possibility of a shared vocabulary. Following a strict interpretation of shared vocabulary, the focus was on terms that were used by 20 or 21 subject pairs and had frequency counts of 100 or more. This provided a manageable set of terms used by a majority of subjects, although it is understood that they may not be representative of terms shared across the three image domains.

For parts of speech, the greatest number of unique terms generated was nouns (Table 5.8), as might be expected. Yet total frequency counts were similar for both nouns and adjectives, indicating that there were more noun terms, but that they were used with less frequency. This was not unexpected, as nouns were generally used analogically (i.e., looks-like) since, by design, the images had no "of-ness." Nouns were frequently specific to a given image, thus exhibiting greater variation and representing more categories, whereas an adjective could be associated with nouns of any category and images of any domain.

Adverbs are notable for their absence in most controlled vocabularies. Although adverbs account for only 3% of the total terms and 5% of the total frequency, these numbers are misleading due to normalization. For example, subjects typically used the word about whenever they felt that a shape or color was not an exact match to the term they were using; and frequently used adverbial phrases were not-too-much and pretty-much, meaning *not exactly*, *close to,* or *almost*. Many variant forms of adverbs were collapsed to an adjectival form. For example, <Approximate> subsumed the words about, generally, loosely, and roughly; and <Exact> subsumed definitely, specifically, and technically. Both <Approximate> and <Exact> are among the high frequency terms listed in Table 5.1. The repeated use of approximations indicates the efforts of subjects to describe aspects of images in terms of <Comparison>. It also supports the inclusion in the vocabulary of those adverbial forms which do not readily lend themselves to mapping to an adjectival form.

Prepositions are noteworthy because of their high frequency counts relative to the low number of unique terms. This appears to support Landau and Jackendoff's (1993) hypothesis that prepositions are devoted to expressing spatial relations.[71] Landau and Jackendoff also contend that languages contain relatively few prepositions (i.e., 80 to 100) because the "class of spatial relations available to be expressed in language … is

[71] The notable exception is the word without which was subsumed by the term negation.

extremely limited" (1993, p. 224). Hence, prepositions represent a finite set of terms that would be used more frequently and should be explored as potential CBIR image search operators (see, for example, Cobb & Petry, 1998).

Some terms could be domain specific, but frequency counts do not indicate domain distribution. Thus, domain distributions for term frequencies in Table 5.2 are relatively equal, with ART being at the high end (40%) and MIC at the low end (27%). This disparity may be accounted for by the possibility that art, as embodied intention, inspires creativity and imagination and thus fosters more detailed descriptions. In contrast, the MIC and SAT images were photographs that required a minimal amount of mediation effort on the part of the describer because of their strong representational nature. This is indicated by use of such domain specific terms as land and aerial-view (Table 5.7) and other terms referencing recognizable objects in the SAT images (e.g., road, house, and stadium) as well as proper nouns (e.g., Shedd-Aquarium, White-House and Jefferson-Monument). Despite the ratings of judges involved in the image selection process, these examples suggest that the choice of satellite images did not lack recognizable objects as was intended.

Frequency counts for unique terms for each of the three domains are similar (see Table 5.2), indicating that domain overlap must be studied in order to understand if individual terms are shared. Table 5.5 provides a view of the overlap across high frequency, high pair value terms, indicating how many are shared and their use in each domain. Table 5.3 suggests that term ranking relationships might be another approach to evaluation of cross-domain vocabulary. For example, circle ranked seventh among the high frequency terms in MIC, but only twentieth in SAT and thirty-first in ART. Table 5.3 highlights the facets that have high domain-specific frequency (e.g., <Cell> in MIC, <Road> in SAT, and <Horizontal> in ART), while Table 5.4 indicates terms used by all subject pairs, but without high frequency within a single domain (e.g., whole, different, beside, and tiny). These terms are of interest because, in order to have high frequency of occurrence, they would need relatively high frequencies in several domains, indicating that they were shared terms, though not necessarily meeting the strict criteria of 20/21 pair count.

The results represented in Table 5.1 show an extreme range of frequencies (from 100 to 2770), yet there are few surprises in the set of terms with high frequencies and high pair counts. The one possible exception is the term squiggle, a specific type of line that was used with high frequency, particularly in MIC. This was an unexpected term and yet visually descriptive of many of the shapes seen in the MIC images. In general, however, the terms listed in Table 5.1 could reasonably be expected given the descriptive task since they emphasize shapes, numbers, relationships, and degree.[72]

6.3 Concepts

The potential of a shared vocabulary is more readily evaluated at the concept level where synonymous and near-synonymous terms have been clustered and the specificity of a term has less import. The result of collapsing isolates into concepts had two advantages: generalization and manageability. Generalization alleviated the specificity of many isolates. This included collapsing individual color values under <Hue> and proper-nouns under their appropriate superordinate concept (e.g., Louvre to <Museum> and Texas to <State>). Antonyms that had been added to the vocabulary for semantic coherence were also subsumed by their superordinates in the process of collapsing isolates to facets. The collapsing of each isolate into its immediate superordinate facet provided a manageable number of 545 concepts that could be analyzed to consider shared-ness. Of these 545 concepts, 80 were never used by subjects and were removed from further analysis. The result was 465 concepts that were actually used by subjects. This total includes 139 terms that might otherwise have been eliminated from consideration due to low occurrence criteria.

One disadvantage of the process of collapsing isolates was that those isolates with high pair occurrences and high frequency were effectively lost when they were subsumed. For example, in Table 5.1, left and right are subsumed by the superordinate <Horizontal> and square is subsumed by <Rectangle>. While it is of interest that left, right, and square are high frequency shared terms, when identifying a controlled vocabulary, it is more useful to know which superordinate concepts are shared across subjects so that specific terms can be accommodated in the future as they arise. This focus on facets rather than isolates avoids the problem of having to enumerate all possible instances of a concept and identifies categories of terms that can be used in CBIR research as potential image search operators or descriptive mechanisms.

6.3.1 Faceted vocabulary

The faceted vocabulary and the conceptual structure that emerged from the faceting process provided a set of mutually exclusive categories, organized within a

[72] Negation stands out in that it could be considered an outlier, an anti-concept that never stands alone but always in contradiction to a concept. Its antonym affirmation was added to complete the logic of <Validation> but was never used because it represents the state of expressing a concept.

nested hierarchy, that standardized the way concepts were referenced. Because of the intense, time-consuming nature of the development process, a faceted vocabulary is typically constructed for a specific application. In this study, no effort was made to simplify the vocabulary beyond collapsing of synonyms and near-synonyms because of the potential visual information provided by specific terms within any given category (e.g., squiggle and wisp in the facet <Line>). This faceted vocabulary was constructed for a specific application: namely, description of the internal contextuality of images. As such, it had two intended audiences: broadly, any image searcher on the Web and, more narrowly, CBIR developers. Since terms were collected from actual natural language descriptions, it would be possible to construct a vocabulary based upon term frequency or subject pair counts. However, neither of these approaches would generate a vocabulary that was conceptually organized.

A faceted thesaurus would provide an alphabetical listing of all terms in the vocabulary and would include a well-developed syndetic structure of relationships including *See, See Also, Used For, Broader Term*, and *Narrower Term.* This faceted vocabulary detailed in Appendix E was not developed as a thesaurus, and, while it does include an alphabetical list of all terms in the vocabulary, only the *Used For* (UF) relationship is designated.

The sets of objects and places that were identified during the faceting process represent paradigmatic relationships available within a structural syntax. These are the low-level elements (isolates) that can be combined to form higher-order elements (facets). The preponderance of syntagmatic concepts that can be used to set boundaries constraining image searches are represented in the <SPATIAL-LOCATION> facet. The <PROPERTY> facet provides paradigmatic descriptors that can be used to limit concepts in the <PLACE> and <OBJECT> facets.

The <Gestalt> sub-facet contributes principles of perceptual organization (completeness, closure, equilibrium, continuation and similarity) as attributes that can be used to describe a specific image. In addition to the general part-whole concepts in the <Gestalt> sub-facet of <Completeness>, partonomies are built into the faceted vocabulary with such concepts as <Mechanical-part>, <Appendage>, <Tissue>, <Construction-material>, and <Building-component>. Partonomies serve to separate entities into structural components. The faceted vocabulary also incorporates the gestalt principle of figure-ground: Figure, in its most rudimentary form, is represented by the <Figure> facet and also by the subfacets of the top level facets <OBJECT> and <PLACE>. Ground is accommodated most specifically by the <Layer> facet, but it is also supported by other properties that differentiate figure from ground, including <Position> and <Color>.

Just as the perceptual principles of gestalt inform the viewer during visual analysis, they can be applied as a geometric basis for image analysis. These principles demonstrate that similarity and adjacency of forms result in visual grouping. When paradigmatic forms change, a new relationship metric need not be required every time change occurs if gestalt principles are operationalized for application across images. The terms and concepts nested within the <Gestalt> facet could facilitate an understanding of how this operationalization might be accomplished.

The property <Hue> is worth special mention as it forms the foundation of CBIR research and development. Berlin and Kay (1969) noted that color perception and color identification do not coincide and that "category boundaries proved to be so unreliable, even for an individual informant, that they have been accorded a relatively minor place in the analysis, consequently … we refer to the foci of categories" (p.13). Despite the fact that the draw task only permitted the use of a lead pencil, all 21 subject pairs used color and hue vocabulary. There were 702 instances of color words which were collapsed into the 11 color terms used by Berlin and Kay and nested under the facet <Hue>. Yet, in any given image, if the color palette were to change, everything else would remain the same. Color, therefore, should be considered a secondary attribute of the internal contextuality of images.

The faceted vocabulary and color frequency occurrence (see Appendix F) indicate that a simplified list of 11 basic colors provides an adequate interface (van den Broek, Kisters, & Vuurpijl, 2004). These basic colors could be modified by combining color-quality facets, as was attempted by Mojsilovic et al. (2004), although without reliance on a structured vocabulary. Under the <Color> facet, both <Hue> and <Tone> had frequency counts over 100 and similar levels of domain distribution (Table 5.17). In retrospect, it seems evident that <Shaded> should have been normalized to <Tone>, which would have given <Tone> more than double its current pair occurrence value.

Looking at placement within the hierarchical structure of concepts, the high frequency terms from Table 5.1 occur in the following facets (listed without <> for readability):

Object: Non-living-thing: Figure: Shape: 2-dimensional-shape
Place: Generic-place: Area
Place: Natural-place: Terrain: Land
Property: Negation

Property: Attribute: Gestalt: Completeness: Part, Whole
Property: Attribute: Gestalt: Similarity
Property: Attribute: Quantity: Linguistic-quantity, Number, Size
Property: Attribute: Quantity: Comparative-measure: Length
Property: Comparison: Degree, Certitude, Equality
Property: Condition: Extension
Property: Attribute: Art-and-craft-process: View
Property: Attribute: Color
Spatial-location: Position: Indexical: Presentation, General-part, Extremity
Spatial-location: Position: Relational: Interconnection, Containment, Placement
Spatial-location: Direction

These high frequency terms are not evenly distributed across the top level facets. As reported in Table 5.13, the top level facet <OBJECT> has the highest number of unique concepts (265 or 49% of 545) and <SPATIAL-LOCATION> the lowest (41 or 8% of 545). But, when considering frequency counts for words, <OBJECT> has a relatively low frequency count (9826 or 19% of 51,629) compared to that of <SPATIAL-LOCATION> (16,879 or 33% of 51,629). This may be explained in part by Landau & Jackendoff's (1993) hypothesis regarding prepositions: because concepts available to describe spatial relations are limited, these concepts were used more frequently by subjects. This has implications for CBIR development. If there is a small set of spatial operators or relationships that can be identified, it could be used to focus development of metrics for the evaluation of images. In many text-based information systems, both the syntagmatic operators *and* and *or* and the paradigmatic relationships *broader term* and *narrower term* constrain the results of searches. A close evaluation of the <SPATIAL-LOCATION> concepts could potentially identify a set of operators for image retrieval systems that would specify object positions and relationships, thereby helping to constrain the results.

In analyzing the top level facet <OBJECT> and its high concept count yet relatively low frequency of occurrence, the rationale for separating <PLACE> from <OBJECT> should be reconsidered. This separation may have been an instance of researcher bias: Knowing that the images came from three domains, the <PLACE> facet may have been artificially highlighted rather than being placed under <OBJECT>. Alternatively, it is possible that the broader concept of <*ENTITY*> could subsume <OBJECT> and <PLACE> The frequency counts in Table 5.14 appear to support this notion of combining <PLACE> with <OBJECT> since it would result in a small increase for <OBJECT> from 9826 to 11,918, or 23% of 51629. In the logical organization, <PLACE> and <OBJECT> would occur as coordinates under <*ENTITY*>, which would simplify the vocabulary from the perspective of CBIR development.

It should be considered, however, that, in the current context of image description, objects were generally used in analogies, whereas references to places may not always be analogies. Since images were selected for their lack of identifiable objects, the use of object terminology and concepts cannot be representative of the actual entity, but rather is used in the sense of "looks-like." Analogies are devices for performing categorization (Holland, Holyoak, Nisbett, & Thagard, 1986). As a device, the value of analogy is in its ability to simplify. Holland *et al*. discuss analogy as a method for problem solving in which attention is given to information other than that derived from the problem at hand. In this case, the problem is image description, and it can be surmised that the gestalt of perception is used to find an object that can serve as a relevant source analog. Analogies depend on the overlap of features between the problem and the selected analogical entity. Since analogies can have different degrees of similarity, they need to be explored. Why does an entity look like a ball? In this context, looking like an object indicates an analogical relationship which could be interpreted as a shape or a shape indicator.

Place, however, is not always an analogy, but sometimes names the place in a particular SAT image, again pointing to the problem that some of the SAT images may have been recognizable. By its very nature, however, satellite imagery is going to be a picture *of* something or somewhere. Sometimes the place term really is a picture of the specific place (e.g., Jefferson-Memorial) and at other times it is an analogy: For example, several subjects observed that ART image #9 looked like the pyramid structure of the Louvre-Museum.

6.3.2 Shared-ness: Concept frequencies and counts

Evaluating the set of concepts listed in Table 5.18, the following concepts stand out because they had shared-ness ratings of 3 for either 20 or 21 subject pairs:

> <Line>, <Angle>, <Similarity>, <Linguistic-quantity>, <Number>, <Size>, <Degree>, <Presentation>, <General-part>, <Extremity>, <Inside-of>, <On>, <Vertical>, <Horizontal>

These fourteen concepts represent the heart of a shared vocabulary: Given the shared-ness ratings for these concepts, it is not unreasonable to assume that subjects would use them to describe images from all domains. The number of concepts with shared-ness ratings of 3 then drops off rather quickly: one concept for 19 pairs; two concepts for 18 pairs; one concept for 17 pairs; and three concepts for 15 pairs. These particular concepts should probably be the starting point for CBIR research, but not to the exclusion of the relationships these concepts have with other concepts that are superordinate or subordinate. In addition, investigation of those concepts with strong domain specificity would be of value in more focused image collections.

The set of 150 concepts with shared-ness ratings of 2 or 3 listed in Table 5.18 represents 32% of the total concepts used by subjects (i.e., 465). These 150 concepts encompass all but two of the concepts with high-frequency counts listed in Table 5.17. Interestingly, comparing frequency distribution for those 2 concepts indicates that they are domain specific: <Land> was used in SAT; and, <Cell> was used in MIC.

Examining the concepts in Table 5.19 that are nested under the <OBJECT> facet, it is evident that <Figure> is a superordinate concept with 17 subordinate concepts (i.e., concepts numbered between 18 and 40) that could be considered shared because of both their shared-ness ratings and their frequency counts. This may be because they are easily recognizable shapes -- visual building blocks that can be interpreted as examples of Focillon's (1948) form without meaning.

Except for the concept <Liquid>, the next 15 concepts nested under <OBJECT> in Table 5.19 have frequency counts of less than 100. Because these concepts are included in the list of concepts that were used in at least two domains, the fact that they were not used across all three domains may indicate that they are not appropriate for describing the internal contextuality of images from the third domain. With the exception of <Substance> and <Vegetable>, the rest of the concepts under <OBJECT> are nested under the superordinate facets *<Image>, <Artifact>, <Animal-body-part> <Naturally-occurring-phenomena>,* and *<Aspects-of-living-thing>*, which are of utility only as a means for organizing subordinate concepts. Thus the superordinate concepts are not as important as the subordinate concepts that may actually represent kinds of shapes: Should <Wave> be nested under <Shape> or <Water-event>? Ultimately, it seems that all objects are shapes in the visual context and used only as analogies.

Similar to the <OBJECT> facet, most of the 14 concepts listed under the <PLACE> facet are nested under the organizing facets *<Constructed-environment>* and *<Natural-place>.* The labels of these organizing facets suggest that their subordinate concepts be further examined for the potential of SAT domain specificity. The remaining concepts listed under <PLACE> are <Country>, which consists solely of proper nouns, and the concepts <Opening> and <Area> nested under *<Generic-place>.* These two concepts nested under *<Generic-place>* also have frequencies of more than 100 and subject pair use of 18 and 20 respectively and are worthy of further examination to determine their distribution across the three domains.

Under the top level facet <PROPERTY>, there are seven more <Gestalt> concepts included in Table 5.19 than were in Table 5.17 (i.e., *<Closure>, <Simplicity>*, <Ordered>, <Grouped>, <Random>, *<Alignment>* and *<Frequency>)* and they represent 75% of the <Gestalt> facets. This emphasizes the importance of gestalt principles as visual organizers. Other concepts from the top level <PROPERTY> facet are subordinates of <Condition> (consistency, concentration, clarity, *sharpness, smoothness*, *extension*, protrusion, indentation), *<Judgment>* (*goodness, neatness*, *perfection*, *abnormal*) and <Action> (*movement-reflecting-growth, random-movement, violent-movement—non-fluid, fluid-movement*). These subordinate concepts are of interest because they are included based on shared-ness ratings rather than high frequency counts. As demonstrated by their subordinates, *<Condition>* and <Action> both serve as organizers for concepts that are references for shapes. But what is the shape of sharpness or of fluid-movement? Even *<Judgment>* may be shape related in that, to be abnormal, there would have to be an identifiable normal shape from which to vary. These examples indicate that a visual shape catalogue could be of value for comparative purposes.

Within *<Art-and-craft-process>,* which is also nested under the top level facet <PROPERTY>, there are several concepts that require further examination for domain specificity. The concepts nested under the facets <Art-style>, *<Production-media>,* and <Technique> (<Design>, <Pattern>, <Cut>, <Contour>, <View>, *<By-position>*) bring to mind the work of Kirsch and Kirsch (1985) on defining the grammars of Modern Art as well as the early Morelli system evaluating painterly qualities (Stenvert, 1992). However, the concept <View> and its subordinate *<By-position>* appear to be a reaction to image domain when considering the isolates nested under them (aerial-view, cross-section, and magnification).

Finally, there is the set of concepts that are nested under <SPATIAL-LOCATION>. This top level facet was well represented in the list of concepts with high frequency counts; and additional concepts appear when considering shared-ness. Each of the six *<Indexical>* concepts is included. Under the facet *<Relational>,* 77% of its subordinates appear in Table 5.19, with several subordinate concepts represented by all of their subfacets: <Interconnection>, <Placement> and <Cover>. These subfacets are three of the seven concepts in <Relational> that are not included in the list of concepts used in two or three domains and demonstrate that superordinate organizers bring focus to conceptual groupings worthy of future research.

6.4 Limitations

This study has been designed to explore the potential for a structured, controlled vocabulary for description of the internal contextuality of images. In retrospect there are several issues which should be discussed.

First and foremost, the terms and concepts that resulted from the analysis of word frequencies must be understood as dependent upon the corpora from which they were drawn. Obviously, the descriptions provided by subjects were affected by the selection of images. In addition, the transcription process itself could not account for nuances in the descriptions or capture relationships between words and phrases even though the original transcript or tape recording was referred to for contextual information during the normalization processes. Furthermore, researcher bias may have affected both the syntactic and semantic structure of the vocabulary through the initial selection of isolates, indicating that both the vocabulary and its organization will require validation.

Although terms used by only one pair were included in the result summaries, they may be of little value in the determination of a shared vocabulary due to their singular use. In a study of category members, Tversky and Hemenway (1984) eliminated every term listed by less than one-third of the subjects on the assumption that these terms were idiosyncratic. In the current research, terms used by only one subject pair were included following the assumption that, while those terms would be of minimal use to a shared vocabulary, they may become significant when agglomerated to their superordinate level and thus contribute to the significance of a concept. In addition, even though a term may have been elicited only once, it may represent a concept or structure that was not generally represented in this set of images.

One particular problem for analysis was deciding at what level of pair agreement and frequency of use is a term to be considered shared? For example, the top 100 frequency scores (Table 5.3) have a subject pair range of 8 to 21, indicating that some high frequency scores were a result of high usage by a relatively small number of pairs. This could be an artifact of individual variation in the communication styles of subjects, a reaction to the describe-draw task inhibiting verbalization, or the reaction of a subject to specific images. For example, the term pin occurred 28 times in a single description [P14.1#234-1st] because the subject used it as the point of reference for describing the image's internal contextuality.

Finally, there is limited use for this controlled vocabulary in its current stage of development. The vocabulary requires validation at both the isolate and facet levels. As it stands, it should not be adapted to image search interfaces; but it could be used to inform future research and to provide a logical structure for coordinating research efforts. In fact, one reason for constructing the faceted vocabulary was as a starting point for CBIR developers to explore conceptual structure in order to resolve current terminological confusions and to develop controlled vocabularies in order to close the semantic gap.

CHAPTER 7
Future Work

The goal of this research was to determine if a shared vocabulary exists for the description of the internal contextuality of images. The vocabulary elicited from subjects was organized into a hierarchical structure of concepts (facets). The Spearman rank correlation coefficient for interrater reliability was used to evaluate consistency of concept use across subject pairs for a set of 465 concepts used by subjects (see Appendix F). Correlations computed using shared-ness ratings for the 465 concepts ranged from a low of .535 to a high of .821 and were statistically significant at the .001 level. Because the degree of correlation observed was unlikely to have occurred by chance, the results of the correlation analyses indicate that subject pairs did agree on a shared vocabulary of concepts.

Having established the viability of a shared vocabulary of concepts for describing the internal contextuality of images, this vocabulary can be used to inform future research in the areas of image vocabulary development, identification of operators for image searching, construction of CBIR metrics for similarity judgments, and the design of interfaces for image retrieval systems.

7.1 Focus on descriptions and subject characteristics

There are refinements of the current study that could be implemented to focus acquisition of natural language vocabulary in order to address specific concerns for CBIR similarity measures. Focusing subject pair description on specific concepts would help to clarify those concepts and gain greater detail about structural components relevant to the concepts. For example, the investigation of elements contributing to pattern by focusing on the descriptions of pattern (texture), as begun by Caivano (1990) and others, would benefit from use of structured descriptive vocabulary so that results could be shared across studies and terminological confusion could be minimized. Assuming that descriptive vocabulary will become more complex as the images described become more complex, a more rigorous categorization of test images into different complexities (e.g., simple design, medium design, and complex design) could help to elicit a broader range of terms (isolates) and, perhaps, concepts that could be applied across image domains. An alternative approach would be to include in the introduction to the describe-draw task instructions focusing the descriptive vocabulary of subjects on a particular aspect or dimension of interest. For example, subjects might be instructed to focus on the shape of dominant features in an image. This would facilitate identifying terms and concepts used to describe particular configurations. Additionally, setting a time limit for descriptions might encourage subjects to focus on the most salient features of an image.

The effects of individual subject characteristics were not explored in the current study. It is possible that these characteristics might influence the shared-ness of vocabulary. For example, do women have a broader color vocabulary than men? Does field dependence or field independence affect the selection and application of concepts? Or does education in a particular field (e.g., fine art, geography, or the biological sciences) affect the choice of terminology for different domains? Another issue to explore is what prompts some subjects to provide lengthy descriptions and others to be more succinct in their use of vocabulary. Is this a subject characteristic, a communication style, or a reaction to the images? And would the vocabulary be affected if the task were written rather than oral?

In addition, a textual analysis of *how* people describe images might be explored and might influence the application of vocabulary in the design of user interfaces by addressing strategies used by subjects to describe images. Descriptive strategies used by the subjects in this study included establishing an initial orientation (e.g., a horizontal rectangle) and identifying shared experiences (e.g., imagine-a-small-rural-home-where-they-took-wood-panels-and-stuck-them-together-as-a-fence [P13.2#109]). In this study, many subjects had trouble both describing size and using comparative measurements within images. One strategy used to address this problem was identification of a spatial measurement tool as a reference point (e.g., dime-size) or referencing index-finger for length. Investigation of descriptive strategies could inform CBIR metrics as well as approaches to search interfaces.

7.2 Develop data processing techniques

One of the biggest stumbling blocks for this study was the lack of a sophisticated database to handle the development and tracking of the vocabulary from its initial collection in the transcriptions, extraction of stop words, and iterative normalization processes to the development of the faceted vocabulary. Identifying or developing a database that could manage the iterative nature of the vocabulary building process would facilitate the process of terminological organization and analysis as well as the ability to share data with other researchers. In particular, analysis of three-dimensional data (domain, term, shared-ness) would be enhanced by use of a relational database. Incorporating the original transcripts in such a database would facilitate the difficult task of textual analysis during word collocations. A standardized approach to vocabulary management would also allow multiple studies to merge into a single organized structure for evaluation and application. And, because validation of the vocabulary is an essential issue for future research, a shared database would help to facilitate the validation process as vocabularies evolve.

The controlled vocabulary that emerged from this study needs validation at each step in the construction process. Transcriptions should be submitted to evaluators who would apply the sets of rules that were used to extract stop words and normalize the vocabulary (see Chapter 4). The resulting vocabularies could then be compared for similarity. Validation is also important for the determination of synonyms and near synonyms and the choice of an authorized term for synonym groups. Another strategy for validating the vocabulary would involve submitting the faceted structure to a review group, as suggested by the hybrid approach to constructing faceted vocabularies (Yang et al., 2004). And, finally, a textual analysis of subjects' post-task discussions would offer an untapped resource for evaluating the validity of the vocabulary as well as providing information about the strategies subjects used in the descriptive process.

7.3 Deeper data analysis

Analysis of the standard images described by all subject pairs was not undertaken in the current study. Similarities and differences between these descriptions could provide insight regarding the shared vocabulary and might also help to focus attention on domain-specific vocabulary. Analysis of standard image descriptions might also contribute to an understanding of the extent to which the occurrence of a concept in one domain predicts its use in another domain.

During the transcription process, long phrases were initially retained so as to preserve the context of the words. These phrases were evaluated and normalized to single terms whenever possible in order to build the controlled vocabulary, even though the contextual relationship between terms was sometimes lost. Examining the collocation of terms and concepts for patterns of combination is an important next step in the analysis of the vocabulary.

The existing data still needs a more detailed analysis of the concepts that did not fall at the extremes of frequency counts and subject pair usage. If a concept were not in the high range of frequency use (i.e., a frequency count of 100 or more) or shared-ness rating (i.e., shared-ness ratings of 2 or 3), then what was its importance in the natural language descriptions that generated the concept vocabulary? More sophisticated statistical methods need to be discovered for this type of analysis, perhaps taking advantage of frequency rankings or domain usage variation. Many of the terms with mid-range frequencies may have been used because of the nature of the image being described: for example, crystal did not emerge as a shared term but was used consistently to describe a single standard MIC image. This points to additional questions: Do individual pictures evoke the same terms from subjects? Does a particular image domain produce more descriptive words? And what level of frequency indicates that a term is domain specific?

The controlled vocabulary developed here could also be used as a basis to coordinate existing categories in CBIR research or for studying the utility of categories already used across extant picture collections and vocabularies. For example, the AAT should accommodate the controlled vocabulary that resulted from this study; but, since it does not exhibit a true faceted structure, initial perusal indicates that it is inadequate for facilitating description of the internal contextuality of images such as those produced in this study. Only a few shapes are included in the AAT terminology (e.g., curve and arch, but not square), which targets expert description of art styles and objects. As has been argued, the problem of terminological overlap in CBIR research could be eased by analyzing how CBIR researchers have defined terms and categories and then mapping them to a consistent and structured vocabulary such as the one produced in this research.

Finally, the purpose of a controlled vocabulary and faceted hierarchical structure needs to be clarified for CBIR developers, interface designers, and actual end-users. The vocabulary could then be represented in thesaural form by identifying terminological relationships (i.e., BT, NT, and USE) and a classification scheme could be developed by adding notation to the existing faceted structure and determining a standard or default citation order. This should not be attempted, however, until the vocabulary has been validated.

7.4 Applications of the controlled vocabulary

This research has the potential to inform CBIR developers and interface designers about user-generated vocabulary at both the levels of term and concept. Adopting a controlled vocabulary will lessen the semantic gap through use of standardized vocabulary to inform threshold setting for interface choices, CBIR similarity metrics, and relevance feedback. Identification of those concepts with high shared-ness ratings could inform CBIR researchers regarding image attributes that are prominent from the user's perspective and could be used to develop new image differentiation metrics. The faceted vocabulary itself offers an organizational structure that could facilitate the combination of CBIR research agendas through the coordination or differentiation of various attribute concepts. Furthermore, it suggests attributes that could be pursued for metric evaluation such as the distinction between approximate and exact or operationalization of <SPATIAL-LOCATION> concepts.

7.4.1 Overlap in CBIR properties

There are many different properties that are addressed in CBIR and image research. But such research is generally based on assumptions about which properties to study, and too often researchers create their own idiosyncratic category names. Categories are rarely identified by the research process or organized in a logical structure, as was attempted in this study. The result is category overlap and terminological confusion.

There is much confusion regarding categories for the description of an image's internal contextuality. As discussed in Chapter 2, Bertin (1980) refers to height as elevation and only addresses length and height in his research, not attending to size; Caivano (1990) addresses size, but only in relation to texture; and Landau (1993) refers to all of these properties as dimensions. Design can refer to pattern, composition and style (Stenvert, 1992), and color can mean hue, chroma, and saturation (Rorvig, 1999). Cobb & Petry (1998) define "during" to mean one object occurring within the x-y coordinate extensions of another object without overlap (see Figure 2.12), whereas Schwering & Raubal (2005) call this a distance function. Lin et al. (2001) refer to texture in terms of rough, coarse and smooth, whereas Caivano (1990) refers to the density of texture; and Belkhatir et al. (2005) attempt to categorize all texture as either bumpy, cracked, disordered, interlaces, lines, marbled, netlike, smeared, spotted, uniform, or whirly.

The problem with failure to employ standard terminology is most evident in the last list of texture types offered by Belkhatir et al. (2005). These terms can refer to more than just the texture of an image. An image could have a composition that is netlike or disordered; two objects in an image could be interlaced or uniform; and lines can refer to the edges of objects in an image as well as the components of a pattern. When terms are not carefully defined and organized, the result for CBIR research is the constant re-creation of definitions and metrics when new image categories and objects emerge.

The faceted vocabulary could facilitate a narrowing of the semantic gap and alleviate terminological confusion by providing CBIR researchers with a structured approach to the terminology that they are attempting to link to image evaluation metrics. Consideration of frequency counts and shared-ness ratings in the faceted vocabulary could also provide focus for CBIR research regarding similarity metrics and for research exploring relationships among image parts. Because a faceted approach allows vocabulary to evolve and new concepts to be added, coordinating the various efforts of CBIR researchers would make an important contribution to bridging the semantic gap.

7.4.2 CBIR metrics and operators

New directions in CBIR operators and metrics continue to affect the development of image search interfaces. An expanded shared vocabulary would provide an opportunity for dialog between the searcher and the system that could help to clarify and narrow user searches. For example, in the Hermitage Museum interface, the user-system dialog could be extended and enhanced by asking the searcher to use a controlled vocabulary to indicate which features of the images in a result set are most critical. The interface might also take free-form text input which could be analyzed by comparison to the controlled vocabulary or it might present options to the searcher based on concepts in the faceted scheme.

Using the full range of specificity to generality in such a vocabulary, current thresholds for analyzing various image features in CBIR search engines could be re-evaluated. For example, when using the color layout option in the Hermitage Museum interface, the user is presented with a color wheel of millions of colors. Yet, when describing images in this research, 24 different color names were used, which were subsequently normalized to only 13 different colors. This supports research indicating that color selection for search interfaces may be more successful with fewer searcher choices (van den Broek et al., 2004).

Both of the examples from the Hermitage Museum search interface are relevant to current search techniques supporting QBPE or find-me-more-like-this strategies. The enhancements offered by application of the controlled vocabulary would provide a mechanism for the searcher to communicate her information need more effectively.

For example, if a searcher selects a stylized picture of a woman holding a chicken as the sample image, she may not want more pictures of chickens and women, but something that *looks like* the sample image. In this context, what characteristics might *looks like* imply? Using low-level properties identified in the faceted vocabulary, CBIR metrics could be developed that would allow the searcher to communicate with the system by selecting simple shape and bright hue as her *looks like* criteria.

The shared concept vocabulary could support the development of additional tools for incorporation into the image search interface. Currently, image retrieval systems evince a paradigmatic focus on pattern matching and the naming of objects in an image (i.e., ofness). Syntagmatic relationships, whether applied to text-based or graphic language resources, demand methods of analysis that go beyond pattern matching into the identification of relationships and roles. A faceted controlled vocabulary for visual structure with an associated set of paradigmatic visual elements could provide structure and insight for the discovery of syntagmatic relationships and roles similar to Kirsch and Kirsch's (1985) codification of the composition of Diebenkorn's painting.

CBIR researchers should explore the potential for syntagmatic operators to be used to constrain searches in CBIR systems in the same way that Boolean operators function in the text-based environment. Analysis of term frequency counts suggests that visual search operators should be explored for their potential to express relationships such as those represented by the terms and concepts nested within the facets <Gestalt> and <SPATIAL-LOCATION>. Using the high frequency shared concepts subsumed by <Gestalt> and <SPATIAL-LOCATION>, operators could be developed that would allow searchers to define relationships between elements when describing the internal contextuality of the desired image. A size and proportion tool could provide samples from which to select, an exact-approximate attribute could be offered, and a shape catalogue or visual vocabulary could be made available to provide both geometric and analogical shapes.

7.4.2 Shape catalog

Developing a visual vocabulary of shapes and relationships could be an important application of the controlled vocabulary that emerged from this study. Although one objective of this study was to identify a vocabulary for representing the internal contextuality of an image, language does not always work well for describing images.

However, the controlled vocabulary could serve as the basis for a visual vocabulary of shapes and relationships that would capture the visual implications of searchers' analogical use of terms.

One basic problem with the vocabulary developed in this study is the inclusion of many isolates. The rationale for retention of so many specific low-level terms was the wide visual variation among synonyms and near-synonyms. If the faceted vocabulary seems to have too much detail, it is because descriptions of images are specific while words are general. Collapsing words generated in the context of describing visual structure risks losing visual subtleties that may be important in image retrieval. For example, the differences among arch, parabola, and hook can be clearly represented with images, but verbally they are near-synonyms and can be subsumed under the single facet <Arc>.

The facet construction process resulted in the retention of many more terms and concepts than might be expected in a more traditional vocabulary due to the wide visual variation of terms in synonym groupings. An emphasis on visual variation would indicate the need to enumerate every possible entity within a conceptual grouping, such as <Animal> or <Artifact>, thus defeating the purpose of identifying synonymous relationships among terms. Given that the images being described were selected, in part, for the absence of recognizable objects, the presence of <OBJECT> and <PLACE> facets in the scheme implies subject reliance on analogical relationships (i.e., looks-like) to describe the internal contextuality of images. The identification of generalizable shapes represented by these terms could provide the basis for the visual expression of synonymous or near-synonymous relationships to describe the internal contextuality of the image, both in its parts and as a whole.

Reliance of subjects on the use of analogy to describe images suggests that the faceted vocabulary of terms and concepts could be used to provide both the user and the CBIR system with a link to the visual shape represented by a verbal construct. Because linguistic mechanisms can not always distinguish between terms at the most appropriate level of visual distinction, a visual vocabulary could be constructed as a catalog of shapes. For example, the concept <Chimney> subsumes the term smokestack; but, while a smokestack is a kind of <Chimney>, the concept <Chimney> and the term smokestack represent shapes that are visually distinct. Furthermore, the basic conceptual structure of a group of synonyms may benefit from visual support. For example, under <Urban>, the seven isolates are visually distinctive: city (large), town (small), suburb (on the edge of a city), sub-division (with a specific street pattern), trailer-park (rows of rectangles), neighborhood (subset of a town or city), and block (generally a rectangle within a neighborhood). The notion of a shape catalog supports current CBIR directions grounded in Biederman's (1987) Recognition-By-Component theory using geons to construct objects (Di Sciascio, Donini, & Mongiello, 2002; Xing, Liu, & Yuan, 2005). Identifying visual prototypes could collocate terms which are conceptually distinct but have similar visual structure, such as sphere, sun, and soccer-ball. A shape catalogue would allow for description using terms and concepts based on visual structure and thus provide a more effective means for describing the internal contextuality of an image.

To create a visual vocabulary or shape catalog, the referent of a term used by subjects could be traced or outlined to produce a figure that would recreate the visual structure represented by the term. For example, <Sun>, <Sphere> and soccer-ball could be represented by the same figure. Such a shape catalog based on simple perceptual organization would build on Focillon's (1948) notion of form devoid of meaning. It would also approximate Biederman's (1987) geon theory regarding the construction of images from components. Verbal access to concepts would lead to visual references for terms via the shape catalog. Thus concepts would serve as entry points leading into the visual vocabulary where shapes would be paired with specific low-level terms.

The terms in the controlled vocabulary are coordinated based on conceptual similarity. Many terms retained in this vocabulary need to be analyzed with respect to visual structure so that potential visual relationships can be constructed. While subtle differences between terms within a concept may develop into significant visual-verbal relationships, this is a problem for future research.

7.5 Conclusion

This research started out to discover if there is a shared vocabulary that could be used to describe images from three domains. The results of normalizing and organizing the transcribed image descriptions of 21 subject pairs indicate that there is a set of terms and concepts for describing visual structure as it occurred in the tested set of images. It was possible to construct a faceted vocabulary from the natural language generated by subject pairs. This vocabulary was organized as a hierarchical structure that established relationships among its terms and concepts. Through an interrater reliability test, it was established that a subset of the concepts in the vocabulary were shared by subjects and that the use of those concepts to describe images in the three test domains was relatively consistent across subject pairs.

In describing images, subjects focused on terms and concepts nested within the facets <Gestalt> and <SPATIAL-LOCATION>. This appears to support the applicability of the pyramid model (Jaimes & Chang, 2000) for representation of visual properties because it highlights terminology representing the syntactic characteristics of local structure and global composition. High frequencies of occurrence for terms and concepts in these facets also support Jorgensen's (2001) contention that, when describing images for retrieval purposes, searchers will use properties that occur at the lower levels of perception. The recommendation for a visual vocabulary in the form of a shape catalogue is supported by the overlap between the conceptual structure of the faceted vocabulary and the descriptions of internal contextuality – the internal physical structure of the image and its parts – provided by subjects. And, finally, given the lack of correspondence with existing vocabularies (e.g., AAT), this research speaks to the need for developers of traditional text-based representational systems to explore the addition of visual structure vocabularies and to share with the CBIR community their experience with the structure of terminological approaches in order to close the semantic gap.

References

Abate, F. R. (Ed.). (2002). *The Oxford American dictionary of current English*. New York: Oxford University Press.

Arnason, H. H. (1968). *History of modern art: painting sculpture, architecture*. Englewood Cliffs, NJ: Prentice-Hall.

Arnheim, R. (1954). *Art and visual perception, a psychology of the creative eye*. Berkeley: University of California Press.

Arnheim, R. (1982). *The power of the center: A study of composition in the visual arts.* Berkeley: University of California Press.

Arnheim, R. (1969). *Visual thinking*. Berkeley: University of California Press.

Athitsos, V., Swain, M. J., & Frankel, C. (1997, June 20). *Distinguishing photographs and graphics on the World Wide Web.* Paper presented at the Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico.

Austin, D. L. (1994). *An image in not an object: but it can help.* Paper presented at the 16th International Essen Symposium, Essen.

Ayto, J., & Simpson, J. A. (2005). *The Oxford dictionary of modern slang*. Oxford: Oxford University Press.

Barnard, K., & Forsyth, D. (2001). *Learning the semantics of words and pictures.* Paper presented at the International Conference on Computer Vision.

Barthes, R. (1985). *The responsibility of forms, critical essays on music, art and representation*. New York: Hill and Wang.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80*(3, Part 2), 1-45.

Batty, D. (1989). Thesaurus construction and maintenance: A survival kit. *Database*, 13-20.

Beebe, C. (1986). *Computer graphics tools applied to message communication: The underlying structure.* Paper presented at the Midwest Regional Academic Microcomputer Conference, Indianapolis.

Beebe, C., & Jacob, E. K. (1998). Graphic language documents: Structures and functions. In W. M. el-Hadi, J. Maniez & A. S. Pollitt (Eds.), *Structures and relations in knowledge organization, proceedings 5th Int. ISKO Conference* (Vol. 6, pp. 244-255). Lille, France: Ergon Verlag.

Belkhatir, M., Mulhem, P., & Chiaramella, Y. (2005). *A full-text framework for the image retrieval signal/semantic integration.* Paper presented at the DEXA 2005.

Belkin, N. J. (1977). *The problem of 'matching' in information retrieval.* Paper presented at the Second International Research Forum in Information Science, London.

Belkin, N. J. (1990). The cognitive viewpoint in information science. *Journal of Information Science, 16*, 11-15.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). ASK for information retrieval Part 1. *Journal of Documentation, 38*(2), 61-71.

Belton, R. J. (2002). *Glossary*. Retrieved September 30, 2002, from http://www.arts.ouc.bc.ca/fina/glossary/gloshome.html

Berenstein, P. (1996). *Finding images online*. Wilton, CT: Pemberton Press.

Berger, A. A. (1989). *Seeing is believing, an introduction to visual communication*. Mountain View, CA: Mayfield Publishing.

Berlin, B., & Kay, P. (1969). *Basic color terms*. Berkeley: University of California Press.

Berryman, G. (1979). *Notes on graphic design and visual communication*. Los Altos, CA: Crisp.

Bertin, J. (1980). The basic test of the graph: A matrix theory of graph construction. In H. Bouma & M. Wrolstad (Eds.), *Processing visible language* (pp. 585-604). New York: Plenum.

Bertin, J. (1983/1967). *Semiology of graphics* (W. J. Berg, Trans.). Madison: University of Wisconsin Press.

Besser, H., & Snow, M. (1990). Access to diverse collections in university settings: The Berkeley dilemma. In T. Petersen & P. Molholt (Eds.), *Beyond the book: Extending MARC for subject access* (pp. 203-224). Boston: G.K.Hall.

Besser, H., & Yamashita, R. (1998). *The cost of digital image distribution: The social and economic implications of the production, distribution and usage of image data*: Andrew W. Mellon Foundation.

Biederman, I. (1972). Perceiving real-world scenes. *Science, 177*(4043), 77-80.

Biederman, I. (1987). Recognition-by-Components: A Theory of Human Image Understanding. *Psychological Review, 94*(2), 115-147.

Bober, M. (2001). MPEG-7 visual shape descriptors. *IEEE Transactions on Circuits and Systems for Video Technology, 11*(6), 716-719.

Bolter, J. D. (1996). Ekphrasis, virtual reality, and the future of writing. In G. Nunberg (Ed.), *The future of the book* (pp. 253-272). Berkeley: University of California Press.

Borgdona, G., Carreta, P., Gagliardi, I., Merelli, D., Mussio, P., Naldi, F., *et al*. (1990). Pictorial indexing for an integrated pictorial and textural IR environment. *Journal of Information Science, 16*(165-173).

Bothamley, J. (1993). *Dictionary of theories*. London: Gale Research International Ltd.

Boulding, K. (1968). *The image, knowledge in life and society*. Ann Arbor: University of Michigan Press.

Bragdon, C. (1922). *The beautiful necessity*. New York: Alfred Knopf.

Brolio, J., Draper, B. A., Beveridge, J. R., & Hansen, A. R. (1989). ISR: A database for symbolic processing in computer vision. *Computer*, 22-30.

Burnham, J. (1971). *The Structure of Art*. New York: George Braziller.

Busch, J. A. (1994). How to choose among alternative technologies for physical access to art images. *Computers and the History of Art, 4*(2), 3-16.

Caivano, J. L. (1990). Visual texture as a semiotic system. *Semiotica, 80*(3/4), 239-252.

Caivano, J. L. (1994). Towards an order system for visual texture. *Language of Design, 2*, 59-84.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.

Carroll, J. M. (1982). Structure in visual communication. *Semiotica, 40*(3/4), 371-392.

Carson, C., Belongie, S., Greenspan, H., & Malik, J. (1997, June 20). *Region-based image querying.* Paper presented at the Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico.

Carson, C., Belongie, S., Greenspan, H., & Malik, J. (2002). Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(8), 1026-1038.

Cataldo, J. M. (1966). *Graphic design & visual communication*. Scranton, PA: International Textbook.

Chang, S.-F., Smith, J. R., & Meng, J. (1997). Efficient techniques for feature-based image/video access and manipulation. In P. B. Heidorm & B. Sandore (Eds.), *Digital image access and retrieval* (pp. 86-99). Urbana-Champaign: Graduate School of Library and Information Science, University of Illinois.

Chang, S.-K., Ichikawa, T., & Ligomenides, P. A. (Eds.). (1986). *Visual languages*. New York: Plenum.

Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist, 167*(5), E117-E139.

Chapman, R. L. (Ed.). (1987). *American slang* (Abridged edition of New Dictionary of American Slang ed.). New York: Harper Paperbacks.

Chen, H.-l. (2001). An analysis of image queries in the field of art history. *Journal of the American Society for Information Science and Technology, 52*(3), 260-273.

Chen, Y., Bart Jr., H. L., & Teng, F. (2005). *A content-based image retrieval system for fish taxonomy.* Paper presented at the MIR '05, Singapore.

Chu, H. (2001). Research in image indexing and retrieval as reflected in the literature. *Journal of the American Society for Information Science and Technology, 52*(12), 1011-1018.

Cobb, M. A., & Petry, F. (1998). Modeling spatial relationships within a fuzzy framework. *Journal of the American Society for Information Science, 49*(3), 253-266.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ.: Erlbaum.

Collins, K. (1998). Providing subject access to images: A study of user queries. *The American Archivist, 61*, 36-55.

Collinson, D. (1987). *Fifty major philosophers, a reference guide*. London: Routledge.

Cordier, F., & Tijus, C. (2001). Object properties: A typology. *Current Psychology of Cognition, 20*(6), 445-472.

Corridoni, J. M., Del Bimbo, A., & Vicario, E. (1997). Painting retrieval based on color semantics. In A. W. M. Smeulders & R. Jain (Eds.), *Image and Multi-Media Search* (Vol. 8, pp. 13-23). River Edge, NJ: World Scientific.

Corridoni, J. M., Del Bimbo, A., & Vicario, E. (1998). Image retrieval by color semantics with incomplete knowledge. *Journal of the American Society for Information Science, 49*(3), 267-282.

Cox, I., Ghosn, J., & Miller, M. L. (1997). *Hidden annotation in content based image retrieval.* Paper presented at the Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico.

Crehange, M., Foucault, O., Halin, G., & Mouoaddib, N. (1989). Semantics of user interface for image retrieval: possibility theory and learning techniques. *Information Processing & Management, 25*(6), 615-127.

Danto, A. C. (1998). Caravaggio's gaze [Review of Caravaggio's Secrets]. *The Lingua Franca Book Review,* 10-11.

Datta, R., Li, J., & Wang, J. Z. (2005). *Content-based image retrieval - approaches and trends of the new age.* Paper presented at the MIR '05, Singapore.

Del Bimbo, A., He, W.-X., & Vicario, E. (1998). Using weighted spatial relationships in retrieval by visual contents. In E. Vicario (Ed.), *Image Description and Retrieval* (pp. 161-192). New York: Plenum.

Di Sciascio, E., Donini, F. M., & Mongiello, M. (2002). *Structured knowledge representation for image retrieval* (Vol. 16).

Doblin, J. (1980). A structure for nontextual communication. In P. A. Kolers, M. Wrolstad & H. Bouma (Eds.), *Processing visible language 2* (pp. 89-111). New York: Plenum.

Dondis, D. A. (1973). *A primer of visual literacy*. Cambridge, MA: The MIT Press.

Drabenstott, K. M., Simcox, S., & Fenton, E. G. (1999). End-User Understanding of Subject Headings in Library Catalogs. *Library Resources & Technical Services, 43*(3), 140-160.

Eakins, J. P., & Graham, M. E. (1999). *Content-based image retrieval: A report to the JISC Technology Applications Programme*. Newcastle: Institute for Image Data Research, University of Northumbria.

Eakins, J. P., & Graham, M. E. (2005). *Content-based image retrieval -- approaches and trends of the new age.* Paper presented at the MIR 05, Singapore.

Eco, U. (1996). Afterward. In G. Nunberg (Ed.), *The future of the book* (pp. 295-306). Berkeley: University of California Press.

Eggleston, P. (2000). Asset management and image overload: Handling 80 billion images a year. *Advanced Imaging*, 12-16.

Elkins, J. (1999). *The domain of images*. Ithaca, NY: Cornell University Press.

Enser, P. G. B., & McGregor, C. G. (1993). *Analysis of visual information retrieval queries* (No. British Library R&D Report 6104). London: British Library.

Fabun, D. (1968). *Communications: The Transfer of Meaning*. Beverly Hills, CA: Glencoe Press.

Fan, J., Gao, Y., Luo, H., & Xu, G. (2004). *Salient objects: Semantic building blocks for image concept interpretation.* Paper presented at the CIVR 2004.

Fan, X., Xie, X., Li, Z., Li, M., & Ma, W.-Y. (2005). *Photo-to-search: Using multimodal queries to search the web from mobile devices.* Paper presented at the MIR '05, Singapore.

Feldman, E. B. (1967). *Art as image and idea*. Englewood Cliffs, NJ: Prentice Hall.

Fidel, R. (1994). User centered indexing. *Journal of the American Society for Information Science, 45*(8), 572-576.

Fleming, M., & Levie, W. H. (1993). *Instructional message design* (2nd ed.). Englewood Cliffs: Educational Technology Publications.

Fleming, M. (1984, January). *Visual attention to picture and word materials as influenced by characteristics of the learners and design of the materials.* Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Dallas, TX.

Florack, L. (1997). *Image structure*. Dordrecht, The Netherlands: Kluwer Academic.

Focillon, H. (1948). *The life of forms* (C. B. Hogan & G. Kubler, Trans. 2nd revised ed.). New York: George Wittenborn, Inc.

Forsyth, D. A., Malik, J., Leung, T. K., Greenspan, H., & Fleck, M. M. (1997). Finding pictures of objects in large collections of images. In P. B. Heidorn & B. Sandore (Eds.), *Digital image access & retrieval* (pp. 118-139). Urbana-Champaign: Graduate School of Library and Information Science at the University of Illinois.

Foucault, M. (1970). *The order of things*. New York: Vintage Books.

Fussell, S. R., & Krauss, R. M. (1989a). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology, 25*, 203-219.

Fussell, S. R., & Krauss, R. M. (1989b). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology, 19*, 509-525.

Fussell, S. R., & Krauss, R. M. (1989). Understanding friends and strangers: The effects of audience design on message comprehension. *European Journal of Social Psychology, 19*, 509-525.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive Science, 29*, 737-767.

Gecsei, J., & Gecsei, D. M. (1989). Browsing access to information systems. *Optical Information Systems*, 237-241.

George, E. E. (1978). Ernst Cassier and Neo-Kantian aesthetics: A holistic approach to the problems of language and art. In R. W. Bailey, L. Matejka & P. Steiner (Eds.), *The sign, semiotics around the world* (pp. 132-145). Ann Arbor: Slavic Publications.

Giardetti, J. R., & Oller, J. W. (1995). A theory of photographic meaning. *Semiotica, 106*(1/2), 99-152.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

Goldsmith, T. E., & Davenport, D. M. (1990). Assessing structural similarity of graphs. In R. W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization*. Norwood, NJ: Ablex.

Goldstein, E. B. (1980). *Sense and perception*. Belmont, CA: Wadsworth.

Golomb, C. (1994). Drawing as representation: The child's acquisition of a meaningful graphic language. *Visual Arts Research, 20*(1), 14-28.

Gombrich, E. H. (1960). *Art and illusion*. Princeton, NJ: Princeton University Press.

Gombrich, E. H., Hochberg, J., & Black, M. (1972). *Art, perception, and reality*. Baltimore: The Johns Hopkins University Press.

Goodman, N. (1968). *Languages of art, an approach to a theory of symbols*. Indianapolis: Bobbs-Merrill.

Goodrum, A. A., Rorvig, M. E., Jeong, K.-T., & Suresh, C. (2001). An open source agenda for research linking text and image content features. *Journal of the American Society for Information Science and Technology, 52*(11), 948-953.

Gorman, M., & Winkler, P. W. (Eds.). (1998). *Anglo-American cataloguing rules* (revised ed.). Chicago: American Library Association.

Gouge, J. O., & Gouge, S. B. (1996). Is the pixel-based model for image processing simply wrong? *Advanced Imaging, 11*(2), 10-12, 66.

Green, R. (1995a). The expression of conceptual syntagmatic relationships: A comparative study. *Journal of Documentation, 51*(4), 315-338.

Green, R. (1995b). Syntagmatic relationships in index languages: A reassessment. *Library Quarterly, 65*(4), 365-385.

Green, R., & Bean, C. A. (1995). Topical relevance relationships. II. An exploratory study and preliminary typology. *Journal of the American Society for Information Science, 46*(9), 654-662.

Grosky, W. I. (1997). Managing multimedia information in database systems. *Communications of the ACM, 40*(12), 73-80.

Guba, E., Wolf, W., de Groot, S., Kneymeyer, M., Van Atta, R., & Light, L. (1964). Eye movements and TV viewing in Children. *AV Communication Review, 12*, 386-401.

Hagen, M. (1979). A new theory of the psychology of representational art. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation* (pp. 196-212). New York: Praeger.

Hammond, T. H. (1993). Toward a general theory of hierarchy: Books bureaucrats, basketball tournaments and administrative structure to the nation-state. *Journal of Public Administration Research and Theory, 3*(1), 120-145.

Hanson, J. (1987). *Understanding video*. Newbury Park, CA: Sage.

Harrison, E. R. (1992). Classification redefined: Research and implications for art education. *Visual Arts Research, 18*(1), 42-49.

Hastings, S. K. (1999). Evaluation of image retrieval systems: Role of user feedback. *Library Trends, 48*(2), 438-452.

Headings in Library Catalogs. *Library Resources & Technical Services, 43*(3), 140-160.

Hibler, J. N. D., Leung, C. H. C., & Mannock, K. L. (1992, February 13-14). *A system for content based storage and retrieval in an image database system.* Paper presented at the Image Storage and Retrieval Systems: Proceedings of the International Society for Optical Engineering, San Jose.

Hirschheim, R., Klein, H. K., & Lyytinen, K. (1995). *Information systems development and data modeling: Conceptual and philosophical foundations*. Cambridge: Cambridge University Press.

Hochberg, J. (1972). The representation of things and people. In E. H. Gombrich, J. Hochberg & M. Black (Eds.), *Art, perception, and reality*. Baltimore: The Johns Hopkins University Press.

Hochberg, J. (1983). Problems of picture perception. *Visual Arts Research*, 7-24.

Hochberg, J., & McAlister, E. (1953). A quantitative approach to figural “goodness.” *Journal of Experimental Psychology, 46*(5), 361-364.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction, processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Hortin, J. A. (1994). Theoretical foundations of visual learning. In D. M. Moore & F. M. Dwyer (Eds.), *Visual Literacy* (pp. 5-29). Englewood Cliffs, NJ: Educational Technology Publications.

Howard, V. A. (1980). Theory of representation: Three questions. In P. A. Kolers, M. Wrolstad & H. Bouma (Eds.), *Procesing visible language 2* (pp. 501-516). New York: Plenum.

Huang, T., Methotra, S., & Ramchandran, K. (1997). Multimedia analysis and retrieval system (MARS) project. In P. B. Heidorm & B. Sandore (Eds.), *Digital image access* (pp. 101-117). Urbana-Champaign: Graduate School of Library and Information Science, University of Illinois.

Hughley, J. H. (1989). Literature review for the influence of structural dimensions on similarity/dissimilarity judgment when categorizing paintings: A developmental study. *Visual Arts Research, 15*(2), 71-77.

Hunt, K. P., & Hodge, M. H. (1971). Category-item Frequency and Category-name meaningfulness (m'): Taxonomic Norms for 84 Categories. *Psychonomic Monograph Supplements, 4*(6), 97-121.

Hurlburt, A. (1977). *Layout: The design of the printed page*. New York: Watson-Guptill.

Iqbal, Q., & Aggarwal, J. (2002). Retrieval by classification of images containing large manmade objects using perceptual grouping. *Pattern Recognition, 35*(7), 1463-1479.

Itten, J. (1961). *Kunst der farbe: subjektives Erleben und objektives Erkennen als Wege zur Kunst*. Ravensburg: Otto Maier Verlag.

Iyer, H., & Giguere, M. (1995). Towards designing an expert system to map mathematics classificatory structures. *Knowledge Organization, 22*(3/4), 141-147.

Jacob, E. K. (1992). Classification and categorization: Drawing the line. In B. Kwasnik & R. Fidel (Eds.), *Advances in classification research: Volume 2* (pp. 67-83). Medford, NJ: Learned Information for the American Society for Information Science.

Jacob, E. K., & Shaw, D. (1998). Sociocognitive perspectives on representation. In M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 33, pp. 131-185). Medford, NJ: Information Today for the American Society for Information Science.

Jaimes, A., & Chang, S.-F. (2000). *A conceptual framework for indexing visual information at multiple levels.* Paper presented at the IS&T/SPIE Conference on Internet Imaging, San Jose, California.

Jorgensen, C. (1995a, October 8). *Classifying images: Criteria for grouping as revealed in a sorting task.* Paper presented at the 6th ASIS SIG/CR Classification Research Workshop, Chicago.

Jorgensen, C. (1995b). *Image attributes: An investigation.* Unpublished Dissertation, Syracuse University, Syracuse.

Jorgensen, C. (1996a). The applicability of selected classification systems to image attributes. In R. Green (Ed.), *Knowledge Organization and Change: Advances in Knowledge Organization* (Vol. 5, pp. 189-197). Frankfurt/Main: Indeks Verlag.

Jorgensen, C. (1996b). *Indexing images: Testing an image description template.* Paper presented at the ASIS.

Jorgensen, C. (1997a). *A framework for image indexing based on free description.* Paper presented at the ASIS SIG/CR.

Jorgensen, C. (1997b). The use of reference points in a sorting task. In *Advances in Classification Research* (Vol. 7, pp. 99-107): ASIS.

Jorgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management, 34*(2/3), 161-174.

Jorgensen, C., Jaimes, A., Benitez, A. B., & Chang, S.-F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology, 52*(11), 938-947.

Jorgensen, C., & Jorgensen, P. (2005). Image querying by image professionals. *Journal of the American Society for Information Science and Technology, 56*(12), 1346-1359.

Keister, L. (1994). User types and queries: Impact on image access systems. In R. Fidel, T. Hahn & E. M. Rasmussen (Eds.), *Challenges in indexing electronic text and images*. Medford, NJ: Learned Information.

Kelly, K. L., & Judd, D. B. (1976). *Color: Universal language and dictionary of names*. Washington, DC: U.S. Dept. of Commerce, National Bureau of Standards.

Kennedy, J. M. (1985). Arnheim, Gestalt theory and pictures. *Visual Arts Research, 11*(1), 23-45.

Kepes, G. (Ed.). (1965). *Structure in art and in science*. New York: George Braziller.

Kindborg, M., & Kollerbaur, A. (1987). *Visual languages and human computer interaction.* Paper presented at the 3rd Conference of British Computer Society, HCI Specialist Group.

Kirsch, J. L., & Kirsch, R. A. (1985, October). *Computers viewing art at work.* Paper presented at the 5th Symposium on Small Computers in the Arts, Philadelphia.

Kirsch, R. A., & Kirsch, J. L. (1985). The structure of paintings: Formal grammar and design. *Planning and Design, 12*(4).

Knowlton, J. Q. (1966). On the definition of picture. *AV Communication Review, 14*(2), 157-183.

Kohler, W. (1947). *Gestalt psychology*. New York: New American Library.

Kolers, P. A., & Smythe, W. E. (1979). Aspects of picturing. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation* (pp. 328-341). New York: Praeger.

Kolers, P. A., Wrolstad, M., & Bouma, H. (Eds.). (1980). *Processing visible language 2*. New York: Plenum.

Krampen, M. (1994). Gestalt. In T. A. Sebeok (Ed.), *The encyclopedic dictionary of semiotics, tome 1* (pp. 292-293). Berlin: Mouton de Gruyer.

Krause, M. G. (1988). Intellectual problems of indexing picture collections. *Audiovisual Librarian, 14*(2), 73-81.

Krauss, R. M., & Fussell, S. R. (1991). Constructing shared communicative environments. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on Socially Shared Cognition* (pp. 172-200). Washington, DC: American Psychological Association.

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: a preliminary study. *Psychonomic Science, 1*, 113-114.

Krauss, R. M., & Weinheimer, S. (1967). Effect of referent similarity and communication mode on verbal encoding. *Journal of Verbal Learning and Verbal Behavior, 6*, 359-363.

Kuhlthau, C. C. (1991). A process approach to library skills instruction. In *Information literacy: Learning how to learn* (pp. 35-40). Chicago: American Library Association.

Lakin, F. (1986). Spatial parsing for visual languages. In S.-K. Chang, T. Ichikawa & P. A. Ligomenides (Eds.), *Visual languages* (pp. 35-85). New York: Plenum.

Lancaster, F. W. (2003). *Indexing and abstracting in theory and practice* (3rd ed.). Champaign, IL: University of Illinois Press.

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences, 16*, 217-265.

Lau, I. Y.-M., & Chui, C.-y. (2001). I know what you know: Assumptions about other's knowledge and their effects on message construction. *Social Cognition, 19*(6), 587-600.

Layne, S. S. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science, 45*(8), 583-588.

Lester, P. M. (1995). *Visual communication, images with messages*. Belmont, CA: Wadsworth Publishing Co.

Leung, C. H. C., Hibler, J. N. D., & Mwara, N. (1992). Picture retrieval by content description. *Journal of Information Science, 18*(2), 111-119.

Levie, W. H. (1984). Research on pictures: A guide to the literature. *Journal of Visual Verbal Languaging*.

Lew, M. S., Sebe, N., & Huang, T. S. (2000). *Improving visual matching*, Hilton Head, SC.

Li, S., Lee, M.-C., & Adjeroh, D. (2005). Effective invariant features for shape-based image retrieval. *Journal of the American Society for Information Science and Technology, 56*(7), 729-740.

Lin, H.-C., Chui, C.-Y., & Yang, S.-N. (2001). *LinStar texture: a fuzzy logic CBIR system for textures'*. Paper presented at the Multimedia '01, Ottawa, Canada.

Lohse, G., Biolsi, K., Walker, N., & Reuter, H. (1994). A classification of visual representations. *Communications of the ACM, 37*(12), 36-49.

Lohse, G., Walker, N., Biolsi, K., & Reuter, H. (1991). Classifying graphical information. *Behaviour & Information Technology, 10*(5), 419-436.

Luescher, M. (1969). *The Luescher color test* (I. A. Scott, Trans.). New York: Random House.

Lunin, L. F. (1994). Analyzing art objects for an image database. In R. Fidel, T. B. Hahn, E. M. Rasmussen & P. J. Smith (Eds.), *Challenges in Indexing Electronic Text and Images* (pp. 57-72). Medford, NJ: American Society for Information Science.

Lupton, E., & Miller, J. A. (Eds.). (1993). *The ABCs of [triangle, square, circle]: The Bauhaus and design theory*. New York: Princeton Architectural Press.

Lupton, E., & Miller, J. A. (1996). *Design writing research, writing on graphic design*. New York: Princeton Architectural Press.

Ma, W.-Y., & Manjunath, B. S. (1998). A texture thesaurus for browsing large aerial photographs. *Journal of the American Society for Information Science, 49*(7), 633-648.

McCloud, S. (1993). *Understanding comics the invisible art*. New York: HarperPerennial.

McCreight, T. (1996). *Design language*. Cape Elizabeth, ME: Brynmorgen Press.

McDonald, S., Lai, T.-S., & Tait, J. (2001). *Evaluating a content based image retrieval system.* Paper presented at the SIGIR ‘01, New Orleans.

MacDonald-Ross, M. (1989). Towards a graphic ecology. In H. Mandl & J. Levin (Eds.), *Knowledge acquisition from text and pictures* (pp. 145-154).

Machines, C. o. P. L. (1969). *Picture language machines*, Canberra, Australia.

Maillot, N., Thonnat, M., & Boucher, A. (2004). Towards ontology-based cognitive vision. *Machine Vision and Applications, 16*, 33-40.

Martinent, J., Chiaramella, Y., & Mulhem, P. (2005, October 31 - November 5). *A model for weighting image objects in home photographs.* Paper presented at the CIKM ‘05, Bremen, Germany.

Mervis, C. B., & Greco, C. (1984). Parts and early conceptual development: Comment on Tversky and Hemenway. *Journal of Experimental Psychology: General, 113*(2), 194-197.

Mitchell, W. T. J. (1980). *The language of images*. Chicago: University of Chicago Press.

Mojsilovic, A., Gomes, J., & Rogowitz, B. E. (2004). Semantic-friendly indexing and querying of images based on the extraction of the objective semantic cues. *International Journal of Computer Vision, 56*(1/2), 79-107.

Morgan, J., & Welton, P. (1992). *See what I mean? an introduction to visual communication* (second ed.). London: Edward Arnold.

Moriarty, S. E., & Kenney, K. (1997, January). *A philosophical discussion of representation.* Paper presented at the 28th Annual Conference of the International Visual Literacy Association, Cheyenne, WY.

Morrison, J. B., & Tversky, B. (2005). Bodies and their parts. *Memory and Cognition, 33*(4), 696-710.

Mostafa, J. (1994). Digital image representation and access. In M. E. Williams (Ed.), *Annual review of information science and technology (ARIST)* (Vol. 29, pp. 91-135). Medford, NJ: American Society for Information Science.

Mountrakis, G., Stefanidis, A., Schlaisich, I., & Agouris, P. (2004). Supporting quality-based image retrieval through user preference learning. *Photogrammetric Engineering and Remote Sensing, 70*(8), 793-981.

Nake, F. (1994). Human-computer interaction- signs and signals interfacing. *Language of Design, 2*, 193-205.

Nohara-LeClair, M. (2001). A direct assessment of the relation between shared knowledge and communication in a referential communication task. *Language and Speech, 44*(2), 217-236.

O'Connor, B. C. (1996). *Explorations in indexing and abstracting: Pointing, virtue, and power*. Englewood, CO: Libraries Unlimited.

O'Donnell, T., Aharon, S., Halliburton, S. S., & Gupta, A. (2000). *Multi-modality model based registration in the cardiac domain.* Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head, SC.

Ohlgren, T. H. (1980). Subject indexing of visual resources: A survey. *Visual Resources, 1*(1), 67-73.

Pacey, P. (1983). Information technology and the universal availability of images. *IFLA Journal, 9*(3), 230-235.

Paek, S., Benitez, A. B., & Chang, S.-F. (1999). *Self-describing schemes for interoperable MPEG-7 multimedia content descriptions.* Paper presented at the Conference on Visual Communications and Image Processing '99, San Jose, CA.

Panofsky, E. (1972). *Studies in iconology* (Icon Editions ed.). New York: Harper & Row.

Papadias, D., Mantzourogiannis, M., Kalnis, P., Mamoulis, N., & Ahmad, I. (1999, August). *Content-based retrieval using heuristic search.* Paper presented at the SIG/IR '99, Berkley.

Pappas, M. (1994). Information skills process models and electronic resources. *Indiana Media Journal, 16*(4), 39-44.

Park, H. S., & Ra, J. B. (1999, January 22-25). *Morphological image segmentation preserving semantic object shapes.* Paper presented at the Visual Communications and Image Processing '99, San Jose.

Petersen, T. (1994). *Art & architecture thesaurus*. Retrieved June 20, 2006, from http://www.getty.edu/research/tools/vocabulary/aat/

Pettersson, R. (1998). Circles, triangles and squares: Sacred symbols in the Mayan culture. *Journal of Visual Literacy, 18*(2), 137-148.

Pisciotta, H., Dooris, M., Frost, J., & Halm, M. (2005). Penn State's visual image user. *Portal Libraries and the Academy, 5*(1), 33-58.

Pollitt, A. S. (1999). *Interactive information retrieval based on faceted classification using views*. Retrieved September 7, 1999, from www.hud.ac.uk/schools/cedar/dorking.htm

Pomerantz, J. R. (1991). The structure of visual configurations: Stimulus versus subject contributions. In G. R. Lockhead & J. R. Pomerantz (Eds.), *The perception of structure* (pp. 195-210). Washington, D.C.: American Psychological Association.

Preble, D., & Preble, S. (1985). *Artforms, an introduction to the visual arts*. New York: Harper Row.

Priss, U., & Jacob, E. K. (1999). Utilizing faceted structures for information systems design. In L. Woods (Ed.), *Knowledge, Creation, Organization and Use: Proceedings of the 62nd ASIS Annual Meeting* (pp. 203-212). Medford, NJ: Information Today.

Quack, T., Monich, U., Thiele, L., & Manjunath, B. S. (2004). *Cortina: A system for large-scale, content-based web image retrieval.* Paper presented at the MM '04, New York.

Quinn, B. (1994). Recent theoretical approaches in classification and indexing. *Knowledge Organization, 21*(3), 140-147.

Ragans, R. (1988). *Arttalk*. Glencoe: Macmillan/McGraw-Hill.

Rahmani, R., Goldman, S. A., Zhang, H., Krettek, J., & Fritts, J. E. (2005, November). *Localized content based image retrieval.* Paper presented at the MIR '05, Singapore.

Rankin, R. O. (1990). A taxonomy of graphs. *Information Design Journal, 6*(2), 147-159.

Rasmussen, E. M. (1997). Indexing images. In M. E. Williams (Ed.), *Annual review of information science and technology (ARIST)* (Vol. 32, pp. 169-196). Medford, NJ: Information Today.

Ratan, A. L., & Grimson, W. E. L. (1997). *Training templates for scene classification using a few examples.* Paper presented at the Workshop on Content-Based Access of Image and Video Laibraries, San Juan, Puerto Rico.

Ravela, S., & Manmatha, R. (1997). *Retrieving images by similarity of visual appearance.* Paper presented at the Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico.

Razel, M., & Vat-Sheva, E. (1986). Developing visual language skills: The Agam Program. *Journal of Visual Verbal Languaging*, 49-54.

Riegl, A. (1992/1893). *Problems of style, foundations for a history of ornament* (E. Kain, Trans.). Princeton, NJ: Princeton University Press.

Rock, I. (1984). *Perception*. New York: Scientific American Books.

Rogowitz, B. E., Frese, T., Smith, J. R., Bouman, C. A., & Kalin, E. (1998, January 26-29). *Perceptual image similarity experiments.* Paper presented at the Human Vision and Electronic Imaging III, San Jose, CA.

Rorvig, M. E. (1993). A method for automatically abstracting visual documents. *Journal of the American Society for Information Science, 44*(1), 40-56.

Rorvig, M. E. (1998). *Image retrieval methods*. Retrieved November 4, 1998, from http://129.120.9.76/corbis

Rorvig, M. E. (1999a). Images of similarity: A visual exploration of optimal similarity metrics and scaling properties of TREWC topic-document sets. *Journal of the American Society for Information Science, 50*(8), 639-651.

Rorvig, M. E. (1999b). A visual exploration of the orderliness of TREC relevance judgments. *Journal of the American Society for Information Science, 50*(8), 552-560.

Rorvig, M. E., & Hemmje, M. (1999). Conference notes - 1996: Foundation of advanced information visualization for visualization information (retrieval) systems. *Journal of the American Society for Information Science, 50*(9), 835-837.

Roussopoulous, N., Faloursos, C., & Sellis, T. (1988). An efficient pictorial database system for PSQL. *IEEE Transaction on Software Engineering, 14*(5), 639-650.

Rudnick, M. F., & Elsuydam, P. (1973). Pictorial stimulus variables. *Viewpoints, 49*(2), 24.

Rui, Y., Huang, T. S., & Chang, S.-F. (1999). Image retrieval: Current techniques, promising directions, and open issues. *Journal of Visual Communication and Image Representation, 10*(1), 39-62.

Rui, Y., Huang, T. S., Methrotra, S., & Ortega, M. (1997, June 20). *A relevance feedback architecture for content-based multimedia information retrieval systems.* Paper presented at the Workshop on Content-Based Access of Image and Video Libraries, San Juan, Puerto Rico.

Saint-Martin, F. (1990). *Semiotics of visual language*. Bloomington, IN: Indiana University Press.

Santini, S., Gupta, A., & Jain, R. (2001). Emergent semantics through interaction in image databases. *IEEE Transactions on Knowledge and Data Engineering, 13*(3), 337-351.

Santini, S., & Jain, R. (1997). *Image databases are not databases with images.* Paper presented at the International Conference on Image Analysis & Processing, Firenze, Italy.

Santini, S., & Jain, R. (1998). *Beyond query by example.* Paper presented at the IEEE Workshop on Multimedia Signal Processing, Los Angeles, CA.

Scheiderman, R. A. (1996). *Why librarians should rule the net*. Retrieved June, 1998, from enodelist@garnet.berkeley.edu

Sclaroff, S. (1997). Distance to deformable prototypes: Encoding shape categories for efficient search. In A. W. M. Smeulders & R. Jain (Eds.), *Image databases and multi-media search* (Vol. 8, pp. 149-159+). River Edge, NJ: World Scientific.

Schwering, A., & Raubal, M. (2005). *Measuring semantic similarity between geospatial conceptual regions.* Paper presented at the GeoS 2005.

Searle, J. R. (1980). Las Meninas and the paradoxes of pictorial representation. In W. T. J. Mitchell (Ed.), *The language of images* (pp. 247-258). Chicago: University of Chicago Press.

Shank, R., & Kass, A. (1988). Knowledge representation in people and machines. In U. Eco, M. Santambrogio & P. Viola (Eds.), *Meaning and mental representation* (pp. 181-200). Bloomington: Indiana University Press.

Shatford, S. (1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging & Classification Quarterly, 6*(3), 39-61.

Small, J. P. (1991). Retrieving images verbally: No more key words and other heresies. *Library Hi Tech, 33*(1), 51-60+67.

Smeulders, A. W. M., Worring, M., Santini, S., Gupta, A., & Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(12), -1-32.

Smith, J. P., & Warner, A. J. (1984). A taxonomy of representations in information retrieval system design. *Journal of Information Science, 8*(3), 113-121.

Smith, J. R. (2001). Quantitative assessment of image retrieval effectiveness. *Journal of the American Society for Information Science and Technology, 52*(11), 969-979.

Smoliar, S. W., Baker, J. D., Nakayama, T., & Wilcox, L. (1997). Multi-media search: An authoring perspective. In A. W. M. Smeulders & R. Jain (Eds.), *Image databases and multi-media search* (Vol. 8, pp. 3-10). River Edge, NJ: World Scientific.

Soergel, D. (1985). *Organizing information*. San Diego, CA: Academic Press.

Sonesson, G. H. (1988). *Pictorial concepts, inquiries into the semiotic heritage and its relevance to the analysis of the visual world.* Lund, Sweden: Lund University Press.

Sonesson, G. H. (1989). *Pictorial concepts* (Vol. 89). Lund, Sweden: Lund University Press.

Sonesson, G. H. (1995). *The multimediation of the lifeworld*. Retrieved January 23, 1998, from http://www.wblab.lu.se/extern/arthist/sonesson/media_1.html

Spink, A. (2001). Searching the web: The public and their queries. *Journal of the American Society for Information Science and Technology, 52*(3), 226-234.

Steinberg, S. G. (1996). Seek and ye shall find (maybe). *Wired*, 108-114ff.

Stemler, S. E. (2004). *A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability*. Retrieved June 15, 2006, from http://pareonline.net/getvn.asp?v=9&n=4

Stenvert, R. (1992). Bridging the gap between pixel and picture. *Computers and the History of Art, 2*(2), 19-24.

Stephenson, C. (1999). Recent developments in cultural heritage databases: Directions for user-centered design. *Library Trends, 48*(2), 410-437.

Steuer, S. (1993). Image databases for the desktop. *Computer Graphics World*, 55-58.

Torres, R. S., Silva, C. G., & Rocha, M. a. H. V. (2003, November). *Visual structures for image browsing.* Paper presented at the CIKM ‘03, New Orleans.

Traina Jr., C., Figueiredo, J. M., & Traina, A. J. M. (2005). *Image domain formalization for content-based image retrieval.* Paper presented at the SAC ‘05, Santa Fe, New Mexico.

Tsai, C. F., McGarry, K., & Tait, J. (2006). Qualitative evaluation of automatic assignment of keywords to images. *Information Processing & Management, 42*(1), 136-154.

Tufte, E. R. (1983). *Visual display of quantitative information*. Cheshire, CT: Graphics Press.

Turner, J. M. (1993). Subject access to pictures: Considerations in the surrogation and indexing of visual documents for storage and retrieval. *Visual Resources, IX*(3), 241-271.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113*(2), 169-193.

Twyman, M. (1979). A schema for the study of graphic language (tutorial paper). In P. A. Kolers & M. E. Wrolstad (Eds.), *Processing of visible language volume I* (pp. 117-150). New York: Plenum Press.

Twyman, M. (1985). Using pictorial language: A discussion of the dimensions of the problem. In T. Duffy & R. Waller (Eds.), *Designing usable text* (pp. 245-312). Orlando: Academic Press.

van den Broek, E. L., Kisters, P. M. F., & Vuurpijl, L. G. (2004, June). *Design guidelines for a content-based image retrieval color-selection interface.* Paper presented at the Dutch Directions in HCI, Amsterdam.

van Rijsbergen, C. J. (1979). *Information retrieval* (second ed.). London: Butterworth's. Retrieved June 30, 2006, from http://www.dcs.gla.ac.uk/Keith/Preface.html

Van Sommers, P. (1984). *Drawing and cognition*. Cambridge: Cambridge University Press.

Vaughn, W. (1989). Paintings by number: Art history and the digital image. In A. Hamber, J. Miles & W. Vaughn (Eds.), *Computers and the History of Art* (pp. 74-97). London: Mansell.

Vaughn, W. (1992). Automated picture referencing: A further look at Morelli. *Computers and the History of Art, 2*(2), 7-18.

Veltkamp, R. C., & Tanase, M. (2001). *Content-based image retrieval systems: A survey* (Technical Report No. UU-CS-2000-34): Utrecht University.

Veltrusky, J. (1973). Some aspects of the pictorial sign. In L. Matejka & I. R. Titunik (Eds.), *Semiotics of art, Prague school contributions* (pp. 245-264). Cambridge, MA: MIT Press.

Ventura, A. D., Gagliardi, I., & Schettini, R. (1998). Indexing color-texture image patterns. In E. Vicario (Ed.), *Image description and retrieval* (pp. 105-120). New York: Plenum.

Wang, J. Z., Jia, L., Chan, D., & Weiderhold, G. (1999). Semantics-sensitive retrieval for digital picture libraries. *D-Lib Magazine, 5*(11).

Wartofsky, M. W. (1979). Picturing and representing. In C. F. Nodine & D. F. Fisher (Eds.), *Perception and pictorial representation* (pp. 272-283). New York: Praeger.

Wellisch, H. (1995). *Indexing from A to Z* (Second ed.). New York: H. W. Wilson.

Williams, R. (1979). *The geometrical foundation of natural structure: A sourcebook of design*. New York: Dover Publications.

Wolf, W. (1970). *A study of eye movement in television viewing* (Final report. No. ERIC No. 046254). Columbus: Ohio State University.

Wollheim, R. (1993). *The mind and its depths*. Cambridge, MA: Harvard University Press.

Worth, S. (1981). Pictures can't say ain't. In L. Gross (Ed.), *Studying Visual Communication* (pp. 179?-184). Philadelphia: University of Pennsylvania Press.

Wu, A., Xu, D., Yang, X., & Zheng, J. (2005a). *Cognition theory motivated image semantics and image language.* Paper presented at the FSKD 2005.

Wu, A., Xu, D., Yang, X., & Zheng, J. (2005b). *Generic solution for image object recognition based on vision cognition theory.* Paper presented at the FSKD 2005.

Xing, W., Liu, W., & Yuan, B. (2005). Extracting superquadratic-based geon description for 3D object recognition. *Chinese Journal of Electronics, 14*(2), 198-202.

Yang, K., Jacob, E. K., Loehrlein, A., Lee, S., & Yu, N. (2004). *The best of both worlds: A hybrid approach to the construction of faceted vocabularies.* Paper presented at the Asian Information Retrieval Symposium, Beijing.

Yu, K., Ma, W.-Y., Tresp, V., Xu, Z., He, X., Zhang, H., et al. (2003). *Knowing a tree from the forest: Art image retrieval using a society of profiles.* Paper presented at the MM '03, Berkeley.

Zhang, L., Hu, Y., Li, M., Ma, W.-Y., & Zhang, H. (2004). *Efficient propagation for face annotation in family albums.* Paper presented at the MM '04, New York.

Zhang, Q.-L., & Chang, S.-K. (1998). Iconic indexing for visual databases. In E. Vicario (Ed.), *Image Description and Retrieval*: Plenum Press.

Appendix A
Natural Language Words

Natural language words are the exact words spoken by the subjects and include all variant forms. Homographs are indicated by an appended number, and homograph definitions are attached as footnotes.

0
1
2
3
4
5
6
7
8
9
10
11
12
13
15
20
22
23
24
25
30
40
45
50
60
70
80
100
1000
1/2
1/3
1/4
1/5
1/6
1/8
1/9
2/3
3/4
3/5
5/8
a
about
above
abstract
abstractly
accent
accented
accurate
across
action
activity
actual
actually
additional
addition-sign
adjacent
aerial
Afghanistan
Africa
after
after2[73]
again
against
ahead
air
airplane
airport
align
aligned
alignment
all
all-over
allowed
all-the-way
almost
alone
along
alphabet
also
alternating
although
always
am
American
amoeba
amoebae
amoebas
amorphicky
amorphous
amount
amphitheater
amplified
an
analogy
and
angle
angled
angles
angling
angular
animal
ankle
another
ant
antenna
antennae
antennas
ants
anvil
any

[73] After2: Indicates time.

anyone
anything
anywhere
apart
apart-from
apex
appear
appearance
appearing
appears
appendage
appendages
apple
approach
approximate
approximately
aqua
aquarium
arc
arch
arched
archipelago
architect
arcing
arcs
are
area
areas
arid
arm
arms
army
around
around2[74]
arranged
arrow
art
arteries
artery
articles
artist
artistic

[74] Around2: Approximate.

artists
as
as-far-as
aside-from
as-if
as-long-as
as-opposed-to
aspect
asterisk
as-well
as-well-as
asymmetrical
at
Atlantic-Ocean
attach
attached
avalanche
avenue
avenues
average
away
awhile
axe
baby
back
back2[75]
Back3[76]
backdrop
background
backward
backwards
bacteria
badge
bag
bags
ball
bam
banana
bananas
band
bands

[75] Back2: Underneath.

[76] Back3: Return.

bank
bar
bare
barely
barren
barrier
bars
base
baseball
baseball-bat
basic
basically
basin
basis
bathroom
battery
bay
be
beach
bead
beaded
beads
beak
beams
bean
bearing
beautiful
becomes
bed
been
before
before2[77]
begin
beginning
begins
behind
beige
beigish
belong
below
beltway
bend

[77] Before2: Indicates time.

bends
beneath
beside
besides
better
between
beyond
big
Big-Bend
bigger
biggest
billow
billowing
biological
biology
bird
bisect
bisecting
bisection
bisects
bit
bits
bitty
black
blacken
blackness
blacksmithing
blade
blank
blanket
blankness
blast
bleacher
bleachers
blend
blends
blister
blob
blobby
blobs
block
block2[78]

[78] Block2: Barrier.

block2ing
block2s
blocks
blocky
Blood
blood-vessel
blood-vessels
blot
blotch
blotches
blotchy
blowing
blue
blueish
blueness
blues
blunt
blur
blurred
blurring
blurry
board
boards
boat
boats
boatyard
bodies
body
body2[79]
bombed-out
bone
bones
bonnet
book
boot
border
bordering
borders
botany
both
bottle
bottom

[79] Body2: Quantity.

boundaries
boundary
bounded
bowing
bowl
bows
bowtie
box
boxes
bracket
brain
branch
branches
branching
bread
break
breaks
breeze
brick
bridge
bridges
bright
brighter
brightest
brightly
brilliant
bring
broad
broadest
broken
brown
browns
brush
bubble
bubbles
bubbly
bug
bugs
building
buildings
built
bulb
bulbous

bulby
bulge
bulges
bulging
bullet
bull's-eye
bump
bumps
bumpy
bunch
burning
burnt
bush
busted
busy
but
butterfly
buzzard
buzz-cut
by
by2[80]
bypass
C
cactus
call
campfire
can
canal
candycane
cantaloupe
canvas
cap
capillary
capital
Capitol
captain
captures
car
caricature
carpet
cars
cartoon
cartoons
carved
carving
cat
CAT-scan
caught-up
cauliflower
cave
cavern
cavernous
cavities
cavity
cell
cells
cellular
celluloid
cement
center
centered
centimeter
centimeters
central
certain
chamber
chambers
change
channel
characteristic
characteristics
charcoal
charcoally
checkerboard
checkered
checkmark
chemistry
chest
Chicago
child
chin
chip
chipped
chop
chopped
choppy
chromosome
chunk
Church
Cincinnati
circle
circlely
circles
circuit
circuitry
circular
circumference
city
city-block
cityscape
clam
class
classic
claw
clay
clear
clearly
cliff
climb
clip
clipped
clips
clock
close
closed
closely
closer
closest
closeup
cloth
cloud
clouds
cloudy
clover
cloverleaf
clump
clumps
cluster

[80] By2: Indicates causation.

clustered
clusters
clutter
coarse
coast
coastal
coastline
coffin
collage
collection
collectively
color
coloration
colored
colorful
coloring
colors
color-scheme
column
columns
combination
combine
come
comes
comic-book
coming
community
compared
compartment
compartments
compass
compensate-for
complete
completely
complex
composed-of
composite
computer
computerchip
computer-graphic
concave
concavity
concentrated
concentration
concentric
conch
concrete
condensed
cone
cones
confine
confines
conjoin
conjoined
connect
connected
connecting
connection
connects
consecutive
consecutively
consistency
consistent
consistently
consists-of
constant
construction
construction-area
contain
contained
container
containing
contains
contemporary
context
continent
continue
continues
contour
contours
contrast
contrasts
convertible
convex
coral
cord
corn
corner
corners
correct
could
counterclockwise
couple
courtyard
cover
covered
covering-up
covers
cover-up
cow
cows
crack
cracked
cracks
cranium
crash
crater
craters
craziness
crazy
cream
creams
create
created
creation
creature
creep
creeping
creep-up-on
crept
crescent
crest
crevice
crevicey
crisscross
crochet
crooked
cross

cross2[81]
crossed
crossing
cross-section
crossways
crosswise
crown
crude
crudely
crumple
crumpled
crystal
crystally
crystals
cube
cubes
cubism
cubist
cup
cupcake
curl
curling
curlish
curly
curlyQ
cursive
curtain
curvature
curve
curved
curves
curving
curvy
cut
cut2[82]
cutoff
cutout
cutouts
cuts
cutting
cytoplasm
cytoplasmic
dab
dabbed
dagger
Dali
dancer
dancing
dark
darkened
darker
darkest
darkly
darkness
dash
dead
dead-end
decent
deep
deep2
deep2ly
defined
definite
definitely
definition
degree
degrees
delicate
demonstrate
demonstration
dense
densely
dent
depiction
depressed
depth
descended
desert
deserty
design
designs
desolate
detail
detailed
details
developed
developing
development
deviated
diagonal
diagonally
diameter
diamond
difference
different
digital
dime
dimension
dimensional
dimensions
dip
dipped
dipping
dips
direct
direction
directions
directly
dirt
disappear
disappearance
disappearing
disappears
discern
discerning
dissected
dissipated
distance
distances
distinct
distinction
distinctly
distinguish
distorted
divide
divided
divider

[81] Cross2: Plus-sign.

[82] Cut2: Traverse.

divides
dividing
division
divvy
divvy-up
DNA
do
dock
docks
does
dog
dollar-sign
dolphin
dome
domes
dominated
dominates
done
door
doorknob
dot
dots
dotted
dotting
dotty
double
doughnut
down
downward
downwards
dozen
dragon
drainage-pipe
drapery
drastic
draw
drawing
drawn
drawn
dried-out
driftwood
driftwoods
drip
dripping
drips
drive
drives
driveway
driveways
drop
droplet
dropped
dropping
droppings
drop-shadow
dry
drying-out
dry-out
due
dumps
dust
dyed
each
ear
ears
earth
earthy
easily
east
echo
echoes
eclipse
edge
edged
edges
effect
egg
eggs
Egypt
either
elbow
electricity
electron-microscope
elephant
elephant-trunk
elevated
elevation
ellipse
elliptical
elongate
else
emanate
emanated
emanating
embryo
emerges
emerging
emitting
emphasis
empty
encircling
enclosed
enclosing
enclosure
encompass
encompasses
encompassing
encroached
encroaching
end
ended
ending-up
ends
elongated
enough
entire
entirely
entirety
equal
equally
equal-sign
equidistance
equidistant
equilateral
equilibrium
eraser
especially
essential
essentially

even
evenly
eventual
eventually
ever
Eveready
every
everything
everywhere
exact
exactly
examining
except
exit
exits
expanded
expanding
exploding
explosion
exposed
extend
extended
extending
extends
extent
exterior
extract
extracted
extreme
extremely
eye
eyeball
eyebrow
eyelashes
eyelid
eyes
fabric
face
face2[83]
facing
factories
factory

[83] Face2: Surface.

fade
faded
faint
fairly
fall
fall-color
falling
far
faraway
farm
farmer
farm-land
farms
farther
farthest
fashion
fat
fat2[84]
fat2ish
fat2test
fatty
feather
feathering
feathers
feathery
feature
features
feed
feeling
feet
feet2[85]
female
femur
fence
fences
festive
few
fewer
fiber
fibers
field

[84] Fat2: Thick.

[85] Feet2: Body part.

fields
fifth
figure
figure8
figures
fill
fill2[86]
fill2er
filled
filled-in
filled-up
fill-in
filling
fills
fin
fine
finely
finger
fingered
fingerprint
fingerprints
fingers
finish
finishing
finish-off
fire
fire-hydrant
first
fish
fist
fit
fits
fitted
five-year-old
flag
flags
flame
flare
flared
flares
flat
flatten

[86] Fill2: Substance.

flattens
fleck
flecks
flesh
flies
flipper
floating
flood
floor
Florida
flow
flower
flowers
flowing
fluffy
fluid
fluidous
fluorescent
fly
flying
focal-point
focus
focused
fold
folds
follicle
follicles
follow
followed
following
follows
food
foot
foot2[87]
football
foothill
for
forefinger
foreground
forest
forestry
fork
forks
form
formation
formed
forming
forms
fossil
fossilized
fountain
fragment
frame
frames
frayed
free-form
freeway
freeways
french-fries
frequent
fries
frill
fringe
fringey
from
front
fruit
fruit-cocktail
full
full-moon
fully
funky
funny
fur
furry
further
furthest
fused
fuses
fuzzy
game
gap
gaping
gaps
gas
gasoline
gave
gel
general
generally
gentle
gently
geographic
geographical
geologic
geological
geometric
germs
get
giant
give
Giza
glass
glasses
glisten
glob
globe
globs
globular
glowing
go
goalpost
goes
going
gold
golden
golf-course
good
goop
goopy
got
gouge
Gradius-Three
gradual
gradually
grainy
grand
granular

[87] Foot2: Body part.

grape
grass
grasses
grassy
gravelly
gray
great
greater
Great-Wall-of-China
Greece
Greek
green
greenish
grid
grids
gridscape
gridwork
gritty
ground
group
grout
grow
growing
grows
guess
guideline
gulf
Gulf-of-Mexico
gully
gumdrop
gun
gun-sight
gushing
guy
H
hail
hair
hair2[88]
hairy
half
halfway
hallmark
hallowed
Hand
hand2
Handle
handmade
handwriting
hanging
haphazard
haphazardly
happening
Harbor
hard2
hardly
harsh
has
hash-mark
hatchery
have
haze
hazy
he
head
headdress
headed
heading
headline
heart
heatwave
heaviest
heavily
heavy
height
helicopter
helipad
her
here
hexagon
hidden
hide
hides
high
higher
highlight
high-school
Highway
hill
hilltop
hilly
him
hint
his
hit
hits
hitting
holding
hole
hollow
home
honeycomb
hook
hooking
horizon
horizontal
horizontally
horn
horrible
horse
horseshoe
horseshoeish
hot-air-balloon
Hot-Wheels
hourglass
house
housed
housing-development
hovering
how
however
hub
huge
hugeass
human
hump
humped
hypotenuse
I

[88] Hair2: Tiny.

I-465
ice
icicles
idea
identical
identifiable
if
image
imaginary
imagine
immediate
immediately
imperfect
important
impression
impressionism
impressionistic
in
in2[89]
in-between
inch
inches
incorrect
increase
increasing
Indentation
index-finger
index-fingers
India
Indian
Indianapolis
Indiana-University
indicate
individual
industrial
in-fact
infrared
infrequent
in-front-of
in-general
initial
initially

[89] In2: Indicates condition.

ink
inkiness
inky
Inlet
inline
inner
insect
inside
instead
intact
intercepting
interconnected
interconnects
interest
interesting
interior
intermittent
interpretation
intersect
intersected
Intersecting
intersection
intersects
intersperse
interspersed
interstate
interstice
interstitial
intertwined
intertwining
intestine
intestiney
into
intricate
invade
invaded
invades
invading
inverse
inverted
Invisible
inward
iridescent

irregular
irregularly
is
island
islands
it
Italy
itty-bitty
jagged
jaggedly
jaggedy
jaggy
jail-bars
jam
Japan
Japanese
jaw
Jefferson-Memorial
jelly
jelly-bean
Jesus
jet
jetting
jigsaw
jog
jogging
jogs
join
joined
joining
joins
joint
joints
junk
just
just2[90]
just3[91]
jut
juts
jutting
keep

[90] Just2: About.

[91] Just3: Merely.

keyhole
kid-would-draw
kid-making-pictures
kidney
kidney-bean
kimono
kind
kind-of
kind-of-like
kind-of-looks-like
king
Klee
Kleeish
kneecap
knife
knobbly
Kool-Aid
L
lace
lagoon
laid
lake
Lake-Monroe
lakes
land
landscape
landscape2[92]
lane
language
large
largely
larger
largest
last
lateral
laterally
latitude
lava
lava-lamp
lay
layer
layers

[92] Landscape2: Scenery.

laying
layout
leached
lead
leading
leaf
lean
leaning
leans
least
leave
leaves
left
left2[93]
left-hand
leftover
leg
leg2[94]
legs
lemon
lemon2[95]
length
lengths
lengthways
lengthwise
less
let
letter
level
levels
lie
light
light2[96]
light2er
light2est
light2ly
lightning
like

[93] Left2: Remaining.

[94] Leg2: Side.

[95] Lemon2: Yellow.

[96] Light2: Pale.

likewise
lima-bean
lime
line
lined-up
lines
line-up
lining
lip
lips
liquid
liquidy
literally
little
lobster
location
locations
long
longer
longest
looked
looking
look-like
loop
loopy
loose
loosely
Los-Angeles
lose
lot
lots
Louvre
low
lower
lowermost
Lowes
lowest
lowlight
lowlights
luminescent
lying
M
made-of

maggot
magnified
main
mainly
maintain
major
majority
make
male
Malevich
man
manmade
manner
many
map
marble
march
marching
margin
margins
marina
mark
marked
marker
mark-out
marks
mashed
mass
massive
match
matched
matching
material
matte
matter
mauve
maybe
McDonalds
me
meandering
meanders
meaning
median
medium
meet
meeting
meets
meet-up
melted
membrane
membranes
memorial
merger
merges
merging
messed-up
messy
meteor
meteors
Mexico
mice
microbial
microphone
microscope
microscopic
mid-
middle
middle-ground
might
mile
miles
millimeter
millimeters
millions
mimic
mimicking
mine
mineral
minerals
mini-
miniature
mining-train
minutes
mirror
miscellaneous
misses
misshapen
missing
miter
mitered
mitochondria
mix
mixed
mixture
mobile-things-people-
live-in
modern
Modern-art
modernesque
modernist
Mohawk
mold
molten
monument
Monument-Circle-Indy
moon
moonish
moons
moonscape
moonsurface
more
more-or-less
mosaic
mosaics
mosaicy
mosquito
moss
most
mostly
motherboard
motion
mottled
mound
mountain
mountainous
mountain-range
mountains
mouth
move

movies
much
mud
multicolor
multilane
multilanes
multiple
mummy
muscle
museum
mushed
mushroom
muzzle
N
narrow
narrower
natural-resources
nature
navy-blue
near
nearby
nearest
nearly
neat
neatly
neatness
neck
needle
needles
negation
negative
neighborhood
neon
nestle
nestled
network
neuron
never
new
New-Mexico
newspaper
next
next-to
nice
night
night-vision-goggles
no
nodule
no-man's-land
non-
none
north
northeast
northwest
nose
not
nothing
nothingness
now
nowhere
nuclear
nucleus
Number
O
object
objects
oblong
observatory
obtuse
obvious
obviously
occur
occurrence
ocean
Ocean-Spray
oceany
oclock
octagon
octagonal
odd
of
off
offset
off-white
Ohio-State
oil
oil-paint
OK
old
olive
on
once-in-a-while
onion
only
open
openwork
opposite
or
orange
oranges
orangey
orbs
order
organic
organism
orient
orientated
orientating
orientation
oriented
orienting
orients
origin
original
originally
or-so
other
other2[97]
otherwise
our
out
outer
outer-space
outline
outlined
outlining
out-of-focus
outside

[97] Other2: Opposite.

outward
oval
ovally
over
over2[98]
over3[99]
overall
over-and-over
overhead
overlap
overlapped
overlapping
overlaps
overview
ovum
pack
packed
Pac-man
page
paint
Painting
pair
pairs
pale
palette
pan
panel
panels
paper
parabola
parallel
paralleling
parch
parched
Paris
park
parking-garage
parking-lot
part
partial
partially

[98] Over2: Above.

[99] Over3: Superimposed.

particular
partition
partitioned
partly
parts
pass
passing
past
pastel
patch
patches
patchy
pattern
patterns
paw
pea
peace-symbol
peach
peachy
peak
peaks
peanut
pear
pearls-on-a-string
pebble
pebbly
peek
peeling
peep
peeping
pencil
penciled
peninsula
people
percent
perfect
perfectly
perimeter
perpendicular
perpendicularly
per-se
person
perspective
perspectives
petal
petals
photo
photograph
Picasso
picket
picking-up
pick-up
picture
pie
piece
pieces
pier
pig
pillar
pillars
pin
pink
pinkish
pinks
pins
pins
pipe
pipeline
pipes
pitch
Pitcher
pixel
pixelated
pixels
place
placed
placement
places
plain
plane
planet
planetarium
plant
plants
plastic
plate

plateau
plateaus
platform
play
playing
plaza
plot
plotted
plug
plus-sign
pod
point
point2[100]
point3ed[101]
point3iest
pointing-to
points-to
pointy
polygon
polygons
ponytail
pool
pool-cue
pop-art
popping
populated
port
portion
portrait
portrait-format
position
positioned
possible
possibly
postage-stamp
Postcard
potato
pour
poured
pouring
powerful

100 Point2: Highlight.

101 Point3ed: Sharpened.

predominant
present
pretty
pretty-much
previous
primitive
prism
prismatic
prisms
probable
probably
profile
progression
progressive
progressively
projection
prominent
proper
proportion
proportions
protecting
protection
protoplasm
protoplasms
protrude
protrudes
protruding
psychedelic
pull
pupil
purple
put
puzzle
pyramid
pyramids
Q
quadrant
quadrants
quality
quarter
quarter2[102]
quasi-

102 Quarter2: One-fourth.

quick
quickly
quintessential
quite
rabbit
radiant
radiate
radiates
radiating
radius
railroad
railroads
railway
rain
rainbow
raindrop
raindrops
raised
Ramen-noodle
ran
random
randomly
range
ran-out
rather
ravaged
ravine
ray
reach
reaching
readily
real
really
recognition
recognizable
recognizably
recognize
recognized
recovers
rectangle
rectangular
rectangularish
rectilinear

red
reddish
reds
reef
reference
reflecting
reflection
region
regions
regular
regularly
relative
relatively
remaining
remarkable
remind
reminiscent
Renoir
repeat
repeated
represent
representation
resemblance
resemble
reservoir
residential
rest
rest-on
retrace
reverse
reversed
rhinoceros
rib
ribbon
ribbons
rib-cage
ribs
rice
ridge
ridges
right
right2[103]

[103] Right2: Correct.

right3[104]
right4[105]
right-hand
right-round
rind
ring
ripple
ripples
river
rivers
rivery
road
roads
roadway
roam
roaming
rock
rocky
roller
rollers
Rome
roof
roofs
room
rooster
root
roots
rotate
rotated
rotten
rotunda
Rouault
rough
rough2[106]
rough2ly
round
roundabout
rounded
roundish
roundness

[104] Right3: Almost.

[105] Right4: Exact.

[106] Rough2: Approximate

row
rows
rug
rumpled
run
rung
running
runs
runway
runways
rural
rusty
S
sailboat
same
sample
sand
sandwich
sandwiched
sandy
San-Fernando-Valley
satellite
say
scaffolding
scale
scallop
scallops
scarf
scarves
scattered
scene
Science-fiction
screen
scribbles
scribbling
scuba-diving
sea
sealed
seam
seashells
seashore
sea-urchin
seaweed

second
secondary-road
section
sectioned
sections
sector
see
seed
seeing
seem
seen
see-through
segment
segmented
semi-circle
Senate
sense
separate
separated
separates
separating
separation
sequence
sequencing
series
set
sets
setting
settlement
setup
several
shade
shaded
shades
shading
shadow
shadowy
shaft
shallow
shape
shape6
shaped
shapely
shapes
shard
shards
share
sharp
sharper
sharply
shattered
Shedd-Aquarium
sheet
shelf
shell
shells
shifted
shingle
shining
ship
ships
shoe
shooting
shoots
shore
shoreline
short
shorter
shortest
shot
shoulder
show
showing
shows
shrank
shrimp
shrinking
shrubbery
shrunken
shy
side
side-by-side
sided
sides
sideways
sight
sign
significant
silhouette
Sim-City
similar
simple
single
singular
sit
sits
sitting
situated
sizable
size
sizes
skeletal
skeleton
sketch
sketched
sketchy
skew
skewed
skim
skimming
skin
skinnier
skinny
skip
skipping
skirts
skull
sky
skyscraper
slant
slanted
slanting
slants
slapped
slate
sleeve
slender
slide

slide2[107]
slight
slightly
slinky
slip
slope
slot
slow
slowly
smack
smackbang
smack-dab
small
smaller
smallerish
smallest
smear
smeared
smears
smidgen
smile
smiley-face
smoke
smokestack
smoky
smokiness
smooth
smoothed
smudged
snail
snake
snakes
snaking
snaky
snapshot
snip
snipped
snout
snow
snowflake
snowy
so

[107] Slide2: Inclined plane.

soccer
soccer-ball
soft
soil
solar-flare
solid
some
somebody
someone
something
sometimes
somewhat
so-much
sorry
sort
sort-of
sort-of-like
sort-of-looks-like
so-to-speak
source
south
southeast
southwest
space
spaced
space-shuttle
spaghetti
sparrow
spatial
specific
specifically
speck
speckled
specks
spectrum
speech-bubble
Spensers
sperm
sperms
spewed
spewing
sphere
spheres

spherical
spider
spidery
spike
spikes
spikey
spiking
spill
spinal-cord
spiral
spit
splash
splashed
splat
splatter
splattered
splattering
split
splitting
splitup
splot
splotch
splotchy
spoke
spongy
spoon
sporting-complex
spot
spots
spouting
sprawled
spray
spraying
spread
spreading
spreads
sprinkled
sprouting
square
squared
squares
squarish
squashed

squeezed
squibble
squiggle
squiggled
squiggles
squigglies
squiggly
squigglyness
squirt
squished
squishy
stack
stacked
stadium
stained-glass
stair
stairs
stalagmite
stand
standard
standing
standout
star
starburst
starbursty
starfish
Starlight
start
started
starter
starting
start-off
Startrek
starts
state
stay
staying
stays
steadily
steady
steep
steeply
steer
stem
stemming
step
stick
stick2[108]
stick2-to
stick-figure
sticking-out
stick-man
stick-out
stick-up
sticky
still
stitching
stone
stood
stop
stopped
stops
stop-sign
storm
storms
straddle
straight
straightly
strand
strange
strap
strawberry
streak
streaks
stream
stream2[109]
stream2ing
street
streets
stretched
stretches
stretching
striation
striking
string
strip
stripe
striped
stroke
strong
structure
structures
stubbier
stuck
stuff
style
stylized
subdivision
submarine
substance
subtle
suburb
sucker
summer-color
summing-up
sun
sunburst
sunbursts
sunflower
sunlight
sunnyside-up
sunrays
sunset
sunshine
sunspot
superhighway
superimposed
Supernintendo
supposed
surface
surfboard
surround
surrounding
surrounds
swastika
sweeping
sweeps

[108] Stick2: Joined.

[109] Stream2: Flowing.

swimming-pool
swirl
swirling
swirls
swirly
swish
swishes
swiss-cheese
switch
swoop
swooping
swoops
swoosh
symbol
symmetrical
symmetry
system
systematically
table
tail
take
takes-up
take-up
taking-up
talk
talking
tall
taller
tallest
tallways
tan
tap
taper
tapered
tapering
tapers
target
teacup
tear
teardrop
technical
technically
teench
teensy
teeny
teeth
telephone
temple
tend
tendency
tentacle
tentacles
tepee
terrace
terrain
terribly
Texas
Textile
texture
textured
textures
texturized
than
that
the
their
them
theme
then
theoretical
there
thermal
thermometer
these
they
thick
thicker
thickest
thickish
thickly
thickness
thin
thing
think
thinner
this
thorn
thoroughfare
those
though
thread
threads
thready
three-dimensional
throat
through
throughout
thrown
thrusted
thrusting
thumb
tibia
tidal-wave
tide
tied
tight
tightly
tile
tiling
tilt
tilted
tilty
times
tiny
tip
tiptop
tire
tissue
to
toast
toe
together
ton
tone
tons
too
took
top
topmost

topographic
tops
toptop
torn
torturous
toss-in
total
totally
toucan
touch
touches
touching
toward
towards
town
track
track2
track2s
track3
track3s
traffic
trail
trailer
trailer-park
trailing
trails
transparent
travel
traveling
tree
trees
trench
triangle
triangular
trippy
Trivial-Pursuit
true
trunk
try
trying
tube
tubular
tumbleweed
tunnel
turkey-vulture
turn
turn-around
turquoise
turquoisey
turtle
TV
twice
twisty
two-dimensional
type
types
U
ugly
unaligned
under
under-construction
underneath
undulating
uneven
unfocussed
uniform
uniformly
university
unravel
until
up
up-against
upon
upper
upright
upside-down
upswing
upward
upwards
urban
us
use
V
vacant
vacuole
valley
Van-Gogh
vanishing-point
vapor-trail
variation
varies
variety
various
vary
varying
vault
vector
vectors
veer
vegetation
vein
veins
version
vertical
vertically
very
vibrant
vice-versa
view
viscous
visible
vision
vivid
vividly
void
Volcano
volume
vulture
W
walking
walkway
wall
walls
wandering
wanders
wart
warts
Washington-DC
water

watercolor
watery
wave
waves
wavy
wavyish
way
we
weather
weave
weaves
weaving
web
wedge
weird
weirdest
weirdness
welded
well
well2
were
west
whale
what
whatever
wheel
wheels
when
where
which
white
whitecap
whitecaps
White-House
whole
whole-bunch
whole-works
whorl
whorly
wide
widen
widens
wider
width
wig
wiggly
wild
will
wind
windblown
window
wings
wiped
wire
wires
wise
wisp
wisps
wispy
with
within
without
woman
wood
wooden
woods
word
words
works
worm
worms
would
woven
wrap
wraparound
wrapped
wrapping
wrapsaround
wrinkles
wrist
wrong
X
x-coordinates
xray
Y
yacht
yachts
yards
yarn
yellow
yellowish
yellows
yellowy
yet
yin-yang
yolk
yolks
you
your
yucky
zag
zebra
zig
zigzag
zigzaggy
zillion
zipper
zone

Appendix B
Stop word list

Homographs are defined in Appendix A.

after2
again
allowed
alone
also
although
am
and
anyone
anything
anywhere
apart-from
are
as
as-far-as
aside-from
as-long-as
as-well
as-well-as
at
awhile
back3
bam
be
becomes
been
before2
belong
besides
bring
built
but
by2
call
canvas
certain
change
climb
come
comes
coming
composed-of
consists-of
continue
continues
create
created
creation
creep
creeping
creep-up-on
crept
cut2
demonstrate
demonstration
depiction
do
does
done
draw
drawn
dropped
dropping
easily
else
eventual
eventually
examining
fall
falling
feed
fifth
first
for
from
gave
get
give
go
goes
going
got
guess
hanging
happening
has
have
he
headed
heading
her
here
him
his
how
however
I
if
image
imagine
in2
indicate
instead
is
it
just3
keep
last
lead
leading
leave
let
likewise
lose
made-of
maintain
make
march
marching
matter
me
meaning
minutes
move
never
now
occur
occurrence
of
OK
or
original
originally
otherwise
our
overview
page
per-se
picking-up
pick-up
picture
placed
previous
pull
put
ran
reach
reaching
readily
recovers
retrace
run
running
runs
scene

second
seem
sit
sits
sitting
somebody
someone
something
sorry
stay
staying
stays
steer
still
sucker
summing-up
supposed
switch
take
tap
than
that
the
their
them
then
there
these
they
thing
think
this
those
though
to
took
toss-in
toward
towards
travel
traveling
try
trying
turn
turn-around
until
us
use
walking
way
we
were
what
whatever
when
where
which
will
wise
with
works
working
would
yet
you
your

Appendix C
Word List

Normalization for variant form is indicated by "used for" (UF).

0
1
2
3
4
5
6
7
8
9
10
11
12
13
15
20
22
23
24
25
30
40
45
50
60
70
80
100
1000
1/2
1/3
1/4
1/5
1/6
1/8
1/9
2/3
3/4
3/5
5/8
a
about
above
abstract-art, [UF: abstract, abstractly]
accent, [UF: accented]
accurate
across
action, [UF: activity]
actual, [UF: actually]
additional
addition-sign
adjacent
aerial
Afghanistan
Africa
after
against
ahead
air
airplane
airport
align, [UF: aligned, alignment]
all
all-over
all-the-way
almost
along
alphabet
alternating
always
American
amoeba, [UF: amoebae, amoebas]
amorphous, [UF: amorphicky]
amount
amphitheater

amplified
an
analogy
angle, [UF: angled, angling, angles, angular]
animal
ankle
another
ant, [UF: ants]
antenna, [UF: antennae, antennas]
anvil
any
apart
apex
appear, [UF: appearance, appearing, appears]
appendage, [UF: appendages]
apple
approach
approximate, [UF: approximately]
aqua
aquarium
arc, [UF: arcing, arcs]
arch, [UF: arched]
archipelago
architect
area, [UF: areas]
arid
arm, [UF: arms]
army
around
around2
arranged
arrow
art
artery, [UF: arteries]
article, [UF: articles]
artist, [UF: artistic, artists]
as-if
as-opposed-to
aspect
asterisk
asymmetrical
Atlantic-Ocean
attach, [UF: attached]
avalanche
avenue, [UF: avenues]
average
away
axe
baby
back
back2
backdrop
background
backward, [UF: backwards]
bacteria
badge
bag, [UF: bags]
ball
banana, [UF: bananas]
band, [UF: bands]
bank
bar, [UF: bars]
bare
barely
barren
barrier
base
baseball
baseball-bat
basic, [UF: basically]
basin
basis
bathroom
battery
bay
beach
bead, [UF: beaded, beads]
beak
beams
bean
bearing
beautiful
bed
before

beginning, [UF: begin, begins]
behind
beige, [UF: beigish]
below
beltway
bend, [UF: bends]
beneath
beside
better
between
beyond
big
big-bend
bigger
biggest
billow, [UF: billowing]
biology, [UF: biological]
bird
bisection, [UF: bisect, bisecting, bisects]
bit, [UF: bits]
bitty
black, [UF: blacken, blackness]
blacksmithing
blade
blank, [UF: blankness]
blanket
blast
bleacher, [UF: bleachers]
blend, [UF: blends]
blister
blob, [UF: blobs, blobby]
block, [UF: blocky]
blocking, [UF: block2, block2s]
blood
blood-vessel, [UF: blood-vessels]
blot, [UF: blotch, blotches, blotchy]
blowing
blue, [UF: blueish, blueness, blues]
blunt
blurry, [UF: blur, blurred, blurring]
board, [UF: boards]
boat, [UF: boats]
boatyard
body, [UF: bodies]
body2
bombed-out
bone, [UF: bones]
bonnet
book
boot
border, [UF: bordering, borders]
botany
both
bottle
bottom
boundaries, [UF: boundary, bounded]
bowing, [UF: bows]
bowl
bowtie
box, [UF: boxes]
bracket
brain
branch, [UF: branches, branching]
bread
break, [UF: breaks]
breeze
brick
bridge, [UF: bridges]
bright, [UF: brightly]
brighter
brightest
brilliant
broad
broadest
broken
brown, [UF: browns]
brush
bubble, [UF: bubbles, bubbly]
bug, [UF: bugs]
building, [UF: buildings]
bulb, [UF: bulbous, bulby]
bulge, [UF: bulging, bulges]
bullet
bull's-eye
bump, [UF: bumps, bumpy]
bunch

burning
burnt
bush
busted
busy
butterfly
buzzard
buzz-cut
by
bypass
C
cactus
campfire
canal
candycane
cantaloupe
cap
capillary
capital
Capitol
captain
captures
car, [UF: cars]
caricature
carpet
cartoon, [UF: cartoons]
carved, [UF: carving]
cat
CAT-scan
caught-up
cauliflower
cave
cavern, [UF: cavernous]
cavity, [UF: cavities]
cell, [UF: cells, cellular]
celluloid
cement
center, [UF: central, centered]
centimeter, [UF: centimeters]
chamber, [UF: chambers]
channel
characteristic, [UF: characteristics]
charcoal, [UF: charcoally]
checkered, [UF: checkerboard]
checkmark
chemistry
chest
Chicago
child
chin
chip, [UF: chipped]
chop, [UF: chopped]
choppy
chromosome
chunk
church
Cincinnati
circle, [UF: circlely, circles, circular]
circuit, [UF: circuitry]
circumference
city
city-block
cityscape
clam
class
classic
claw
clay
clear, [UF: clearly]
cliff
clip, [UF: clipped, clips]
clock-orientation, [UF: clock]
close, [UF: closely]
closed
closer
closest
closeup
cloth
cloud, [UF: clouds, cloudy]
clover
cloverleaf
clump, [UF: clumps]
cluster, [UF: clustered, clusters]
clutter
coarse

coastline, [UF: coastal, coast]
coffin
collage
collection, [UF: collectively]
color, [UF: coloration, colored, coloring, colors]
colorful
color-scheme
column, [UF: columns]
combine, [UF: combination]
comic-book
community
comparison, [UF: compared]
compartment, [UF: compartments]
compass-orientation, [UF: compass]
compensate-for
complete, [UF: completely]
complex
composite
computer
computerchip
computer-graphic
concave, [UF: concavity]
concentration, [UF: concentrated]
concentric
conch
concrete
condensed
cone, [UF: cones]
confine, [UF: confines]
conjoin, [UF: conjoined]
connect, [UF: connected, connecting, connection, connects]
consecutive-order, [UF: consecutive, consecutively]
consistency
consistent, [UF: consistently]
constant
construction
construction-site, [UF: construction- area]
container
containment, [UF: contain, contained, containing, contains]
contemporary
context
continent
contour, [UF: contours]
contrast, [UF: contrasts]
convertible
convex
coral
cord
corn
corner, [UF: corners]
correct
could, [UF: can]
counterclockwise
couple
courtyard
cover, [UF: covered, covers]
cover-up, [UF: covering-up]
cow, [UF: cows]
crack, [UF: cracked, cracks]
cranium
crashing, [UF: crash]
crater, [UF: craters]
crazy, [UF: craziness]
cream, [UF: creams]
creature
crescent
crest
crevice, [UF: crevicey]
crisscross
crochet
crooked
cross, [UF: crossed, crossing, crossways, crosswise]
cross2
cross-section
crown
crudely, [UF: crude]
crumpled, [UF: crumple]
crystal, [UF: crystally, crystals]
cube, [UF: cubes]
cubism, [UF: cubist]
cup

cupcake
curl, [UF: curling, curlish, curly,
curlyQ]
cursive-writing, [UF: cursive]
curtain
curve, [UF: curvature, curved, curves,
curving, curvy]
cut, [UF: cutting, cuts]
cutoff
cutout, [UF: cutouts]
cytoplasm, [UF: cytoplasmic]
dab, [UF: dabbed]
dagger
Dali
dancer, [UF: dancing]
dark, [UF: darkly, darkened,
darkness]
darker
darkest
dash
dead
dead-end
decent
deep
deep2, [UF: deep2ly]
definite, [UF: defined, definitely]
definition
degree, [UF: degrees]
delicate
dense, [UF: densely]
dent
depressed
depth
descended
desert, [UF: deserty]
design, [UF: designs]
desolate
detail, [UF: detailed, details]
developing
development, [UF: developed]
deviation, [UF: deviated]
diagonal, [UF: diagonally]
diameter
diamond
different, [UF: difference]
digital
dime
dimension, [UF: dimensions,
dimensional]
dip, [UF: dipped, dipping, dips]
direct, [UF: directly]
direction, [UF: directions]
dirt
discern, [UF: discerning]
disappearing, [UF: disappearance,
disappear, disappears]
dissected
dissipated
distance, [UF: distances]
distinct, [UF: distinctly]
distinction
distinguish
distorted
division, [UF: divided, divide,
divider, divides, dividing]
divvy, [UF: divvy-up]
DNA
dock, [UF: docks]
dog
dollar-sign
dolphin
dome, [UF: domes]
dominant, [UF: dominated, dominates]
door
doorknob
dot
dotted, [UF: dotting, dots, dotty]
double
doughnut
down, [UF: downward, downwards]
dozen
dragon
drainage-pipe
drapery
drastic
drawing, [UF: drawn]

dried-out, [UF: dry-out, drying-out]
driftwood, [UF: driftwoods]
drive, [UF: drives]
driveway, [UF: driveways]
drop, [UF: drip, dripping, drips, droplet, droppings]
drop-shadow
dry
due
dump, [UF: dumps]
dusty, [UF: dust]
dyed
each
ear, [UF: ears]
earth, [UF: earthy]
east
echo, [UF: echoes]
eclipse
edge, [UF: edged, edges]
effect
egg, [UF: eggs]
Egypt
either
elbow
electricity
electron-microscope
elephant
elephant-trunk
elevated
elevation
ellipse, [UF: elliptical]
elongate, [UF: elongated]
emanate, [UF: emanated, emanating]
embryo
emerging, [UF: emerges]
emitting
emphasis
empty
encircling
enclosing, [UF: enclosed, enclosure]
encompass, [UF: encompasses, encompassing]
encroaching, [UF: encroached]
end, [UF: ended, ending-up, ends]
enough
entire, [UF: entirely, entirety]
equal, [UF: equally]
equal-sign
equidistant, [UF: equidistance]
equilateral
eraser
especially
essentially, [UF: essential]
even, [UF: evenly]
ever
Eveready
every
everything
everywhere
exact, [UF: exactly]
except
exit, [UF: exits]
expanding, [UF: expanded]
exploding, [UF: explosion]
exposed
extend, [UF: extended, extending, extends]
extent
exterior
extracted, [UF: extract]
extremely, [UF: extreme]
eye, [UF: eyes]
eyeball
eyebrow
eyelash, [UF: eyelashes]
eyelid
fabric
face
face2
facing
factory, [UF: factories]
fade, [UF: faded]
faint
fairly
fall-color
far

faraway
farm, [UF: farms]
farmer
farm-land
farther, [UF: further]
farthest, [UF: furthest]
fashion
fat, [UF: fatty]
fat2, [UF: fat2ish]
fat2test
feather, [UF: feathers, feathering, feathery]
feature, [UF: features]
feeling
feet
feet2
female
femur
fence, [UF: fences]
festive
few
fewer
fiber, [UF: fibers]
field, [UF: fields]
figure, [UF: figures]
figure8
filled, [UF: fill1, filled-up, fillin-, filled-in, filling, fills]
filler, [UF: fill2]
fin
fine, [UF: finely]
finger, [UF: fingered, fingers]
fingerprint, [UF: fingerprints]
finish, [UF: finishing, finish-off]
fire
fire-hydrant
fish
fist
fit, [UF: fitted, fits]
five-year-old
flag, [UF: flags]
flame
flared, [UF: flare, flares]
flat, [UF: flatten, flattens]
fleck, [UF: flecks]
flesh
flipper
floating
flood
floor
Florida
flower, [UF: flowers]
flowing, [UF: flow]
fluffy
fluid, [UF: fluidous]
fluorescent
flying, [UF: fly, flies]
focus, [UF: focal-point, focused]
folded, [UF: fold, folds]
follicle, [UF: follicles]
follow, [UF: followed, following, follows]
food
foot
foot2
football
foothill
forefinger
foreground
forest, [UF: forestry]
fork, [UF: forks]
form, [UF: formation, formed, forming, forms]
fossil, [UF: fossilized]
fountain
fragment
frame, [UF: frames]
frayed
free-form
freeway, [UF: freeways]
french-fries
frequent
fries
frill
fringe, [UF: fringey]
front

fruit
fruit-cocktail
full, [UF: fully]
full-moon
funky
funny
furry, [UF: fur]
fused, [UF: fuses]
fuzzy
game
gap, [UF: gaps]
gaping
gasoline, [UF: gas]
gel
general, [UF: generally, in-general]
gentle, [UF: gently]
geography, [UF: geographic, geographical]
geology, [UF: geologic, geological]
geometry, [UF: geometric]
germ, [UF: germs]
giant
Giza
glass
glasses
glisten
glob, [UF: globs, globular]
globe
glowing
goalpost
gold, [UF: golden]
golf-course
good
goopy, [UF: goop]
gouge
Gradius-Three
gradual, [UF: gradually]
grainy, [UF: granular]
grand
grape
grass, [UF: grasses, grassy]
gravelly
gray
great
greater
Great-Wall-of-China
Greece, [UF: greek]
green, [UF: greenish]
grid, [UF: grids, gridscape, gridwork]
gritty
ground
grouped, [UF: group]
grout
growing, [UF: grow, grows]
guideline
gulf
Gulf-of-Mexico
gully
gumdrop
gun
gun-sight
gushing
guy
H
hail
hair, [UF: hairy]
hair2
halfway
hallmark
hand
hand2
handle
handmade
handwriting
haphazard [UF: haphazardly]
harbor
hard2
hardly
harsh
hash-mark
hatchery
hazy, [UF: haze]
head
headdress
headline

heart
heatwave
heaviest
heavy, [UF: heavily]
height
helipad, [UF: helicopter]
hexagon
hidden, [UF: hide, hides]
high
higher
highlight
high-school
highway
hill, [UF: hilly]
hilltop
hint
hit, [UF: hits, hitting]
holding
hole
hollow, [UF: hollowed]
home
honeycomb
hook
hooking
horizon
horizontal, [UF: horizontally]
horn
horrible
horse
horseshoe, [UF: horseshoeish]
hot-air-balloon
Hot-Wheels
hourglass
house
housed
housing-development
hovering
hub
huge, [UF: hugeass]
human
hump, [UF: humped]
hypotenuse
I-465
ice
icicle, [UF: icicles]
idea
identical
identifiable
imaginary
immediate, [UF: immediately]
imperfect
important
impression
impressionism, [UF: impressionistic]
in-between
inch, [UF: inches]
incorrect
increase, [UF: increasing]
indentation
index-finger, [UF: index-fingers]
India, [UF: Indian]
Indiana-University
Indianapolis
individual
industrial
in-fact
infrared
infrequent
in-front-of
initial, [UF: initially]
ink, [UF: inkiness, inky]
inlet
inline
insect
inside-of, [UF: in, inner, inside,
into, inward]
intact
intercepting
interconnection, [UF: interconnected,
interconnects]
interest, [UF: interesting]
interior
intermittent
interpretation
intersecting, [UF: intersect,
intersected, intersection,

intersects]
intersperse, [UF: interspersed]
interstate
interstice, [UF: interstitial]
intertwined, [UF: intertwining]
intestine, [UF: intestiney]
intricate
invade, [UF: invaded, invading, invades]
inverted, [UF: inverse]
invisible
iridescent
irregular, [UF: irregularly]
island, [UF: islands]
italy
itty-bitty
jagged, [UF: jaggedy, jaggedly, jaggy]
jail-bar, [UF: jail-bars]
jam
Japan, [UF: Japanese]
jaw
Jefferson-Monument
jelly
jelly-bean
Jesus
jet, [UF: jetting]
jigsaw
jog, [UF: jogging, jogs]
joined, [UF: join, joining, joins]
joint, [UF: joints]
junk
just
just2
jut, [UF: jutting, juts]
keyhole
kid-would-draw, [UF: kid-making-pictures]
kidney
kidney-bean
kimono
kind
kind-of
kind-of-like
kind-of-looks-like
king
Klee, [UF: Kleeish]
kneecap
knife
knobbly
Kool-Aid
L
lace
lagoon
lake, [UF: lakes]
Lake-Monroe
land
landscape
landscape2
lane
language
large, [UF: largely]
larger
largest
lateral, [UF: laterally]
latitude
lava
lava-lamp
layer, [UF: layers]
laying, [UF: laid, lay]
layout
leached
leaf, [UF: leaves]
lean, [UF: leaning, leans]
least
left, [UF: left-hand]
leftover, [UF: left2]
leg, [UF: legs]
leg2
lemon
lemon2
length, [UF: lengths, lengthways, lengthwise]
less
letter
level, [UF: levels]

lie, [UF: lying]
light
light2, [UF: light2ly]
light2er
light2est
lightning
like
lima-bean
lime
line, [UF: lines]
line-up, [UF: lined-up]
lining
lip, [UF: lips]
liquid, [UF: liquidy]
literally
little
lobster
location, [UF: locations]
long
longer
longest
look-like-(ll), [UF: looked, looking]
loop, [UF: loopy]
loose, [UF: loosely]
Los-Angeles
lot, [UF: lots]
Louvre
low
lower
Lowes
lowest, [UF: lowermost]
lowlight, [UF: lowlights]
luminescent
M
maggot
magnified
main, [UF: mainly]
major
majority
male
Malevich
man
manmade
manner
many
map
marble
margin, [UF: margins]
marina
mark, [UF: marked, marks]
marker
mark-out
mashed
mass
massive
match, [UF: matched, matching]
material
matte
mauve
maybe
McDonalds
meandering, [UF: meanders]
median
medium
meet, [UF: meeting, meets, meet-up]
melted
membrane, [UF: membranes]
memorial
merger, [UF: merging, merges]
messy, [UF: messed-up]
meteor, [UF: meteors]
Mexico
microbe, [UF: microbial]
microphone
microscope, [UF: microscopic]
mid-
middle
middle-ground
might
mile, [UF: miles]
millimeter, [UF: millimeters]
million, [UF: millions]
mimic, [UF: mimicking]
mine
mineral, [UF: minerals]
mini-

miniature
mining-train
mirror
miscellaneous
misses
misshapen
missing
miter, [UF: mitered]
mitochondria
mixture, [UF: mix, mixed]
mobile-things-people-live-in
modern, [UF: modernesque, modernist]
modern-art
mohawk
mold
molten
monument
Monument-Circle-Indy
moon, [UF: moonish, moons]
moonscape, [UF: moonsurface]
more
more-or-less
mosaic, [UF: mosaics, mosaicy]
mosquito
moss
most, [UF: mostly]
motherboard
motion
mottled
mound
mountain, [UF: mountainous, mountains]
mountain-range
mouse, [UF: mice]
mouth
movie, [UF: movies]
much
mud
multicolor
multilane, [UF: multilanes]
multiple
mummy
muscle
museum
mushed
mushroom
muzzle
N
narrow
narrower
natural, [UF: nature]
natural-resources
navy-blue
near, [UF: nearby]
nearest
nearly
neat, [UF: neatly, neatness]
neck
needle, [UF: needles]
negation
negative
neighborhood
neon
nestle, [UF: nestled]
network
neuron
new
New-Mexico
newspaper
next, [UF: next-to]
nice
night
night-vision-goggles
nodule
no-man's-land
none
north
northeast
northwest
nose
not, [UF: no, non-]
nothing, [UF: nothingness]
nowhere
nuclear
nucleus
number

O
object, [UF: objects]
oblong
observatory
obtuse-angle, [UF: obtuse]
obvious, [UF: obviously]
ocean, [UF: oceany]
Ocean-Spray
oclock
octagon, [UF: octagonal]
odd
off
offset
off-white
Ohio-State
oil
oil-paint
old
olive
on
once-in-a-while
onion
only
open
openwork
opposite
orange, [UF: oranges, orangey]
orb, [UF: orbs]
ordered, [UF: order]
organic
organism
orientation, [UF: orient, orientated, orientating, oriented, orienting, orients]
origin
or-so
other
other2
outer-space
outline, [UF: outlined, outlining]
out-of-focus
outside-of, [UF: out, outer, outside, outward]
oval, [UF: ovally]
over
over2
over3
overall
over-and-over
overhead
overlap, [UF: overlapped, overlapping, overlaps]
ovum
pack
packed
Pac-man
paint
painting
pair, [UF: pairs]
pale
palette
pan
panel, [UF: panels]
paper
parabola
parallel-to, [UF: parallel, paralleling]
parch, [UF: parched]
Paris
park
parking-garage
parking-lot
part, [UF: partial, partially, partly, parts]
particular
partition, [UF: partitioned]
pass, [UF: passing, past]
pastel
patch, [UF: patches, patchy]
pattern, [UF: patterns]
paw
pea
peace-symbol
peach, [UF: peachy]
peak, [UF: peaks]
peanut
pear
pearls-on-a-string

pebble, [UF: pebbly]
peek, [UF: peep, peeping]
peeled, [UF: peeling]
pencil, [UF: penciled]
peninsula
people
percent
perfect, [UF: perfectly]
perimeter
perpendicular-to, [UF: perpendicular, perpendicularly]
person
perspective, [UF: perspectives]
petal, [UF: petals]
photograph, [UF: photo]
Picasso
picket
pie
piece, [UF: pieces]
pier
pig
pillar, [UF: pillars]
pin, [UF: pins]
pink, [UF: pinkish, pinks]
pipe, [UF: pipes]
pipeline
pitch
pitcher
pixel, [UF: pixelated, pixels]
place, [UF: places]
placement
plain
plane
planet
planetarium
plant, [UF: plants]
plastic
plate
plateau, [UF: plateaus]
platform
play, [UF: playing]
plaza
plot, [UF: plotted]
plug
plus-sign
pod
point
point2
point3iest
point3y, [UF: point3ed]
pointing-to, [UF: pointing-to, pointed-to, points-to]
polygon, [UF: polygons]
ponytail
pool-cue
pop-art
popping
populated
port
portion
portrait
portrait-format
position, [UF: positioned]
possible, [UF: possibly]
postage-stamp
postcard
potato
pouring, [UF: pour, poured]
powerful
predominant
present
pretty
pretty-much
primitive
prism, [UF: prismatic, prisms]
probable, [UF: probably]
profile
progressive, [UF: progression, progressively]
projection
prominent
proper
proportion, [UF: proportions]
protecting, [UF: protection]
protoplasm, [UF: protoplasms]
protrusion, [UF: protrude,

protruding, protrudes]
psychedelic
pupil
purple
puzzle
pyramid, [UF: pyramids]
Q
quadrant, [UF: quadrants]
quality
quarter
quasi-
quick, [UF: quickly]
quintessential
quite
rabbit
radiating, [UF: radiant, radiate,
radiates]
radius
railway, [UF: railroad, railroads]
rain
rainbow
raindrop, [UF: raindrops]
raised
Ramen-noodle
random, [UF: randomly]
range
ran-out
rather
ravaged
ravine
ray
real
really
recognize, [UF: recognized,
recognition, recognizable, recognizably]
recreational-facility
rectangle, [UF: rectangular,
rectangularish]
rectilinear
red, [UF: reddish, reds]
reef
reference
reflection, [UF: reflecting]
region, [UF: regions]
regular, [UF: regularly]
relative, [UF: relatively]
remaining
remarkable
remind, [UF: reminiscent]
Renoir
repeat, [UF: repeated]
represent, [UF: representation]
resemble, [UF: resemblance]
reservoir
residential
residential-building
rest
rest-on
reversed, [UF: reverse]
rhinoceros
rib, [UF: ribs]
ribbon, [UF: ribbons]
rib-cage
rice
ridge, [UF: ridges]
right, [UF: right-hand]
right2
right3
right4
right-round
rind
ring
ripple, [UF: ripples]
river, [UF: rivers, rivery]
road, [UF: roads, roadway]
roam, [UF: roaming]
rock, [UF: rocky]
roller, [UF: rollers]
Rome
roof, [UF: roofs]
room
rooster
root, [UF: roots]
rotated, [UF: rotate]
rotted, [UF: rotten]
rotunda

Rouault
rough
rough2, [UF: rough2ly]
round, [UF: rounded, roundish,
 roundness]
roundabout
row, [UF: rows]
rug
rumpled
rung
runway, [UF: runways]
rural
rusted, [UF: rusty]
S
sailboat
same
sample
sand
sandwich
sandwiched
sandy
San-Fernando-Valley
satellite
say
scaffolding
scale
scallop, [UF: scallops]
scarf, [UF: scarves]
scattered
science-fiction
screen
scribble, [UF: scribbles, scribbling]
scuba-diving
sea
sealed
seam
seashells
seashore
sea-urchin
seaweed
secondary-road
section, [UF: sectioned, sections]
sector
see, [UF: seeing, seen]
seed
see-through
segment, [UF: segmented]
semi-circle
Senate
sense
separate, [UF: separated, separates,
 separation, separating]
sequence, [UF: sequencing]
series
set, [UF: sets]
setting
settlement
setup
several
shade, [UF: shades]
shaded, [UF: shading]
shadow, [UF: shadowy]
shaft
shallow
shape, [UF: shaped, shapely, shapes]
shape6
shard, [UF: shards]
share
sharp, [UF: sharply]
sharper
shattered
Shedd-Aquarium
sheet
shelf
shell, [UF: shells]
shifted
shingle
shining
ship, [UF: ships]
shoe
shooting, [UF: shoots, shot]
shore
shoreline
short
shorter
shortest

shot
shoulder
show, [UF: showing, shows]
shrimp
shrubbery
shrunken, [UF: shrank, shrinking]
shy
side, [UF: sided, sides]
side-by-side
sideways
sight
sign
significant
silhouette
Sim-City
similar
simple
single, [UF: singular]
situated
sizable
size, [UF: sizes]
skeleton, [UF: skeletal]
sketch, [UF: sketched, sketchy]
skew, [UF: skewed]
skim, [UF: skimming]
skin
skinnier
skinny
skipping, [UF: skip]
skirts
skull
sky
skyscraper
slanting, [UF: slant, slanted, slants]
slapped
slate
sleeve
slender
slide
slide2
slight, [UF: slightly]
slinky
slip
slope, [UF: sloping, slopes]
slot
slow, [UF: slowly]
smack, [UF: smackbang, smack-dab]
small
smaller, [UF: smallerish]
smallest
smear, [UF: smeared, smears]
smidgen
smile
smiley-face
smoke, [UF: smoky, smokiness]
smokestack
smooth, [UF: smoothed]
smudged
snail
snake, [UF: snakes, snaking, snaky]
snapshot
snip, [UF: snipped]
snout
snow, [UF: snowy]
snowflake
so
soccer
soccer-ball
soft
soil
solar-flare
solid
some
sometimes
somewhat
so-much
sort
sort-of
sort-of-like
sort-of-looks-like
so-to-speak
source
south
southeast
southwest

space, [UF: spaced, spatial]
space-shuttle
spaghetti
sparrow
specific, [UF: specifically]
speck, [UF: speckled, specks]
spectrum
speech-bubble
Spensers
sperm, [UF: sperms]
spewed, [UF: spewing]
sphere, [UF: spheres, spherical]
spider, [UF: spidery]
spikey, [UF: spike, spikes, spiking]
spill
spinal-cord
spiral
spit
splash, [UF: splashed]
splatter, [UF: plat, splattered, splattering]
split, [UF: splitting, splitup]
splot, [UF: splotch, splotchy]
spoke
spongy
spoon
sporting-complex
spot, [UF: spots]
spouting
sprawled
spraying, [UF: spray]
spread, [UF: spreads, spreading]
sprinkled
sprouting
square, [UF: squared, squares, squarish]
squashed
squeezed
squibble
squiggle, [UF: squiggled, squiggles, squigglies, squiggly, squigglyness]
squirt
squished, [UF: squishy]
stack, [UF: stacked]
stadium
stained-glass
stair, [UF: stairs]
stalagmite
standard
standing, [UF: stand, stood]
standout
star
starburst, [UF: starbursty]
starfish
Starlight
start, [UF: started, start-off, starter, starting, starts]
Startrek
state
steady, [UF: steadily]
steep, [UF: steeply]
stem
stemming
step
stick
stick2, [UF: stick2-to, stuck]
stick-figure, [UF: stick-man]
stick-out, [UF: sticking]
stick-up
sticky
stitching
stone
stop, [UF: stops, stopped]
stop-sign
storm, [UF: storms]
straddle
straight, [UF: straightly]
strand
strange
strap
strawberry
streak, [UF: streaks]
stream
stream2, [UF: stream2ing]
street, [UF: streets]
stretched, [UF: stretches,

stretching]
striation
striking-out
string
strip
striped, [UF: stripe]
stroke
strong
structure, [UF: structures]
stubbier
stuff
style
stylized
subdivision
submarine
substance
subtle
suburb
summer-color
sun
sunburst, [UF: sunbursts]
sunflower
sunlight, [UF: sunrays, sunshine]
sunnyside-up
sunset
sunspot
superhighway
superimposed
Supernintendo
surface
surfboard
surrounding, [UF: surround,
surrounds]
swastika
sweeping, [UF: sweeps]
swimming-pool, [UF: pool]
swirl, [UF: swirling, swirls, swirl]
swish, [UF: swishes]
swiss-cheese
swoop, [UF: swooping, swoops]
swoosh
symbol
symmetrical, [UF: symmetry]
system, [UF: systematically]
table
tail
take-up, [UF: taking-up, takes-up]
talk, [UF: talking]
tall, [UF: tallways]
taller
tallest
tan
taper, [UF: tapered, tapering,
tapers]
target
teacup
teardrop
technical, [UF: technically]
teench
teensy, [UF: teeny]
teeth
telephone
temple
tend, [UF: tendency]
tentacle, [UF: tentacles]
tepee
terrace
terrain
terribly
texas
textile
texture, [UF: textured, textures,
texturized]
theme
theoretical
thermal
thermometer
thick, [UF: thickly, thickish,
thickness]
thicker
thickest
thin
thinner
thorn
thoroughfare
thread, [UF: threads, thready]

three-dimensional
throat
through
throughout
thrown
thrusted, [UF: thrusting]
thumb
tibia
tidal-wave
tide
tied
tight, [UF: tightly]
tile, [UF: tiling]
tilt, [UF: tilted, tilty]
times
tiny
tip
tiptop, [UF: toptop]
tire
tissue
toast
toe
together
ton, [UF: tons]
tone
too
top, [UF: topmost, tops]
topography, [UF: topographic]
torn, [UF: tear]
torturous
total, [UF: totally]
toucan
touching, [UF: touch, touches]
town
track
track2, [UF: track2s]
track3, [UF: track3s]
traffic
trail, [UF: trails]
trailer
trailer-park
trailing
transparent
tree, [UF: trees]
trench
triangle, [UF: triangular]
trippy
Trivial-Pursuit
true
trunk
tube, [UF: tubular]
tumbleweed
tunnel
turkey-vulture
turquoise, [UF: turquoisey]
turtle
TV
twice
twisty
two-dimensional
type, [UF: types]
U
ugly
unaligned
under
under-construction
underneath
undulating
uneven
unfocused
uniform, [UF: uniformly]
university
unravel
up, [UF: upper, upward, upwards]
up-against
upon
upright
upside-down
upswing
urban
V
vacant
vacuole
valley
Van-Gogh
vanishing-point

vapor-trail
variation, [UF: vary, varying,varies]
various, [UF: variety]
vault
vector,[UF: vectors]
veer
vegetation
vein, [UF: veins]
version
vertical, [UF: vertically]
very
vibrant
vice-versa
view
viscous
visible
vision
vivid, [UF: vividly]
void
volcano
volume
vulture
W
walkway
wall, [UF: walls]
wandering, [UF: wanders]
wart, [UF: warts]
Washington-DC
water, [UF: watery]
watercolor
wave, [UF: waves, wavy, wavyish]
weather
weaving, [UF: weave, weaves]
web
wedge
weird, [UF: weirdness]
weirdest
welded
well
well2
west
whale
wheel, [UF: wheels]
white
whitecap, [UF: whitecaps]
White-House
whole
whole-bunch
whole-works
whorl, [UF: whorly]
wide, [UF: width]
wider, [UF: widen, widens]
wig
wiggly
wild
wind
windblown
window
wings
wiped
wire, [UF: wires]
wise
wisp, [UF: wisps, wispy]
within
without
woman
wood, [UF: wooden, woods]
word, [UF: words]
worm, [UF: worms]
woven
wrap, [UF: wrapped, wrapping]
wraparound, [UF: wrapsaround]
wrinkled, [UF: wrinkles]
wrist
wrong
X
x-coordinates
xray
yacht, [UF: yachts]
yard, [UF: yards]
yarn
yellow, [UF: yellowish, yellows,
yellowy]
yin-yang
yolk, [UF: yolks]
yucky

Z
zebra
zig, [UF: zag]
zigzag, [UF: zigzaggy]
zillion
zipper
zone

Appendix D
Term List

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
0	19	10	[UF: nothing]		
1	286	21			
2	631	21	[UF: both, couple, dou-ble, either, pair, twice]		
3	216	21			
4	115	20			
5	98	18			
6	61	16			
7	53	13			
8	34	11			
9	17	8			
10	25	11			
11	8	5			
12	6	5	[UF: dozen]		
13	3	1			
15	5	2			
20	6	2			
22	1	1			
23	1	1			
24	1	1			
25	4	3			
30	10	5			
40	4	3			
45	21	9			
50	2	2			
60	1	1			
70	3	1			
80	1	1			
100	1	1		[SN: Used with nu-meric values.]	
1000	4	1		[SN: Used with nu-meric values.]	
1/2	537	20	[UF: half]		
1/3	254	15			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
1/4	131	15	[UF: quarter2]		
1/5	11	3			
1/6	7	2			
1/8	21	6			
1/9	1	1			
2/3	54	12			
3/4	28	10			
3/5	1	1			
5/8	1	1			
a	3695	21	[UF: an, another, any, each, individual, only, particular, single]	[SN: Indicates the article not the letter.]	
<Abnormal>	0	0			381
above	129	20	[UF: elevated, over2, raised]	[SN: Refers to above on a x-y axis.]	
abstract-art	49	15	[UF: cubism, impression-ism, stylized]		
<Accessory>	0	0			
<Action>	9	6	[UF: motion]		387
adult	0	0			
aerial	89	20	[UF: airplane, hot-air-bal-loon, overhead]	[SN: Refers to type of view.]	
affirmation	0	0			
Afghanistan	1	1			
Africa	2	1			
air	4	4			
<Aircraft>	0	0			59
<Airport>	13	4			230
<Air-travel-environ-ment>	0	0			
aligned	15	7	[UF: inline, lined-up]		
<Alignment>	0	0			336
all	521	21	[UF: always, entire, ever, every, everything, over-all, total, whole-works]		
almost	325	20	[UF: fairly, just2, nearly, pretty much, quasi-, right3]		
<Alphabet>	3	1			13

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
alternating	2	1			
amoeba	19	5			
<Amphitheater>	4	1			237
<Angle>	148	16		[SN: Specific angles are preceded by a numeral, e.g., 90 degrees.]	23
<Animal-body-part>	0	0			
<Animal-life>	2	2	[UF: animal, creature]	[SN: Includes mythical creatures]	145
<Ankle>	2	1			189
ant	2	2			
antenna	5	2			
anvil	2	1			
<Appendage>	1	1			181
apple	4	1			
approximate	1066	20	[UF: about, around, general, loosely, more or less, relatively, rough, say]		
<Aquarium>	1	1			238
<Arc>	17	2			29
arch	16	2			
archipelago	1	1			
<Architect>	2	2			224
<Area>	130	20			313
<Arm>	55	9			182
army	2	1			
<Arrangement>	0	0		[SN: Indicates proximity in gestalt theory.]	
arrow	7	3			
art	6	3	[UF: capture]		
<Art-and-craft-process>	0	0			
artery	8	2			
<Article>	1	1			110
<Artifact>	0	0			
<Artist>	10	2			225

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Art-Style>	31	12	[UF: appearance, approach, effect, fashion, feeling, impression, manner, style, stylized]		394
<Aspects-of-living-thing>	0	0			
asterisk	1	1			
asymmetrical	1	1			
Atlantic-Ocean	1	1			
<Atmosphere>	0	0		[SN: Phenomena not specifically related to weather conditions.]	126
<Attribute>	0	0			
avalanche	1	1			
average	2	1	[UF: medium]		
<Axe>	6	2			51
back	10	6		[SN: Refers to the rear side.]	
background	56	12	[UF: backdrop]		
bacteria	2	1	[UF: germ]		
bad	0	0			
<Badge>	1	1			98
<Bag>	6	1			78
banana	2	1			
bank	2	1			
barely	190	19	[UF: hardly, just3, slightly]		
baseball	8	3			
baseball-bat	1	1			
bathroom	1	1			
battery	3	1	[UF: Eveready]		
<Bay>	27	5			291
beach	13	5			
beaded	4	1			
beak	6	1			
bean	5	4			
beautiful	1	1			
<Bed>	1	1			102

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
behind	38	15		[SN: Refers to a place to the right of an object.]	
below	163	19	[UF: under, beneath]		
<Beltway>	8	2			273
Beltway-465	2	2	[UF: I-465.]		
<Beside>	188	20	[UF: adjacent, along, follow, lining, nestle, next, side-by-side, skirting, trailing]		450
<Between>	151	20	[UF: in-between, sandwiched]		451
<Beverage>	0	0			87
Big-Bend	2	1			
biology	11	4			
<Bird>	9	6			149
bisection	23	4			
bit	209	19	[UF: dab, fleck]		
black	150	19	[UF: pitch]		
<Blacksmithing>	1	1			49
<Blade>	2	1			50
blanket	2	1			
bleacher	2	2			
blend	8	4	[UF: leached, composite]		
blob	103	15	[UF: glob]		
<Block>	1	1		[SN: A rectangular shaped solid.]	39
blocking	7	5	[UF: barrier, protecting]		
<Blood>	9	5			204
<Blood-vessel>	6	4			205
blowing	5	3	[UF: windblown]		
blue	122	13	[UF: navy-blue, turquoise]		
blunt	2	1			
blurry	23	8	[UF: fuzzy, hazy, out-of-focus, unfocused, smudged]		
board	5	1			
<Board-game>	0	0			67

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Boat>	10	1	[UF: ship, yacht]		60
boatyard	1	1		[SN: Refers to a dry-dock.]	
<Body>	16	7			168
<Body-of-water>	0	0			
<Bone>	1	1			208
bonnet	6	1			
<Book>	6	2			108
boot	9	3			
botany	1	1			
<Bottle>	1	1			
bottom	1078	21	[UF: base]		
<Bowl>	19	3	[UF: basin]		75
bowtie	2	1			
bracket	1	1			
brain	1	1			
branch	17	6	[UF: stemming]	[SN: A line off to side of a straight line at less than 90 degrees.]	
<Bread>	3	1			89
breeze	1	1			
<Brick>	2	1			248
<Bridge>	23	9			275
bright	23	9	[UF: brighter, brightest, brilliant, glisten, shining, vibrant, vivid]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Brightness>	0	0			423
broken	16	6	[UF: busted]		
brown	40	13	[UF: tan, sandy, beige]		
brush	32	5			
bubble	32	6	[UF: blister]		
<Building>	86	19	[UF: complex]	[SN: Types of buildings.]	231
<Building-component>	0	0			
<Building-element>	0	0			
bulb	6	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
bullet	4	1			
bump	32	10	[UF: bulge, knobbly, nodule]		
<Burial-artifact>	0	0			106
<Burnt>	1	1			386
bush	3	2	[UF: shrubbery]		
<Business>	0	0			
<Business-structure>	0	0			244
butterfly	10	2			
buzz-cut	1	1			
<By-object>	0	0			413
bypass	5	2			
<By-position>	0	0			412
C	14	1			
cactus	3	1			
canal	9	1			
<Candy>	0	0			86
candycane	1	1			
cantaloupe	1	1			
cap	2	1			
capillary	2	1			
capital	1	1			
Capitol	2	2			
captain	1	1			
<Car>	6	5			61
<Cartoon>	5	4	[UF: caricature, comic book]		399
<Carved>	2	1			405
<Case>	0	0			15
cat	1	1			
cauliflower	6	2			
cave	9	2	[UF: cavern]		
<Cell>	114	15			202
<Cement>	1	1			250
center	937	21	[UF: equidistant, halfway, middle, median, mid]		
centimeter	206	8			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Certitude>	0	0			356
<Change-in-condition>	0	0			
channel	5	2			
charcoal	2	1			
checkered	9	5			
checkmark	12	1			
<Cheese>	0	0			88
chemistry	1	1			
chest	1	1			
Chicago	4	2			
child	3	2	[UF: five-year-old]		
<Children's-art>	3	3	[UF: kid-would-draw]	[SN: Refers to a child like picture style.]	400
children's-slide	1	1	[UF: slide2]		
<Chimney>	0	0			263
<Chin>	2	1			174
chip	6	4	[UF: chunk]		
chromosome	4	1			
<Church>	2	2	[UF: temple]		233
Cincinnati	1	1			
circle	682	21	[UF: ring, round]		
circuit	2	1			
circumference	2	1			
city	87	18			
city-block	27	7	[UF: block1]		
cityscape	5	2			
clam	3	1			
<Clarity>	0	0			365
<Classic>	1	1			396
claw	1	1			
clay	5	1			
cliff	3	1			
<Clock-orientation>	28	4	[UF: hand2, oclock]	[SN: Refers to o'clock and the clock hand position]	465
closed	7	3	[UF: sealed]		
<Closed-curve>	0	0			25

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Closure>	0	0			329
<Cloth>	2	2	[UF: fabric]		92
<Clothing>	0	0			
<Cloud>	18	6			129
clover	2	2			
cloverleaf	2	2			
<Cnidarian>	0	0		[SN: Previously known as Coelenter-ate.]	159
coarse	1	1			
coastline	35	5			
coffin	2	1			
<Collage>	1	1			5
<Color>	140	20	[UF: dye, shade]		415
<Color-quality>	0	0			
column	24	7	[UF: stack]		
<Comparative-mea-sure>	0	0			
<Comparison>	3	2	[UF: contrast		351
compartment	5	2			
<Compass-orientation>	3	2			464
<Completeness>	0	0			
complex	12	6	[UF: busy, intricate, torturous]	[SN: Complicated.]	
computerchip	1	1			
<Computer-game>	0	0			68
<Concentration>	8	4			362
concentric	10	2			
conch	2	1			
<Condition>	0	0		[SN: State, or, figural characteristics.]	
<Cone>	1	1			43
consecutive-order	1	1			
<Consistency>	18	9	[UF: texture]		361
<Constructed-environ-ment>	0	0			
<Constructing>	0	0			48
<Construction-material>	0	0			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Construction-site>	2	1		[SN: Transliterated from “construction area.”]	
<Container>	2	1			72
<Containment>	8	4			441
<Continent>	1	1			309
<Continuation>	0	0			
<Contour>	59	15	[UF: outline, profile, silhouette]		408
convertible	1	1			
<Cooking>	0	0			52
coral	2	2			
cord	1	1			
corn	2	2			
corner	690	21		[SN: A 90 degree angle.]	
correct	9	6	[UF: accurate, proper, right2, true, well]		
<Correctness>	0	0			377
<Country>	0	0			310
<Courtyard>	1	1			284
<Cover>	100	18	[UF: all-over, cover-up, mark-out]		455
<Covering>	0	0		[SN: The coverings on animal objects, and organic things occurring on skin.]	
cow	8	2			
crack	43	12	[UF: crevice, split]		
crashing	1	1			
crater	15	3			
crazy	13	4			
<Creative-source>	0	0			
crescent	6	2			
crisscross	4	4			
crochet	1	1			
crooked	12	7	[UF: distorted, mis-shapen]		
<Cross>	285	21	[UF: across, straddle]		439

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
cross-section	11	4			
crown	2	1			
<Crustacean>	0	0			156
crystal	44	8			
cube	1	1		[SN: A square shaped solid.]	
<Cultural-building>	0	0			
<Cup>	1	1			73
<Cupcake>	3	1			82
<Curl>	27	9	[UF: swirl]		28
<Cursive-writing>	7	1	[UF: handwriting]		
curtain	3	1	[UF: drapery]		
<Curve>	278	19	[UF: bend, bow]		24
<Cut>	67	16	[UF: chopped, clip, cut-off, dissected, snip]		404
cutout	2	2			
<Cutting-tool>	0	0			
<Cylinder>	3	2	[UF: roller]		44
cytoplasm	2	1			
dagger	6	2			
Dali	1	1			
<Damaged>	0	0			384
<Dancer>	3	1			226
dark	126	19	[UF: darker, darkest, deep]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
dash	3	3			
<Decayed>	0	0			385
<Decrease>	0	0			353
deep	5	3			
<Degree>	48	11	[UF: extent]	[SN: Adverbs must be included due their prevalence in the raw data.]	355
<Delicacy>	0	0			364
delicate	10	3	[UF: fine]		

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
dense	28	10	[UF: condensed, heavy, heaviest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Depth>	2	2			350
desert	17	6	[UF: no man's land]		
<Design>	16	6	[UF: structure]	[SN: Designs contribute to the visual illusion of tactile texture.]	402
<Destroyed>	0	0			
developing	1	1	[UF: under-construction]		
<Deviation>	15	5	[UF: jog, veer, zag]		371
<Device>	0	0			57
<Diagonal>	222	17			462
diameter	15	5			
<Diamond>	17	3			33
different	144	21	[UF: other, as opposed to]		
digital	9	4	[UF: computer, computer graphic]		
dime	2	1			
<Dimension>	21	5			342
<Dimensionality>	0	0			344
<Direction>	43	15	[UF: bearing, -wise]	[SN: Directional location from viewer perspective.]	458
<Disappearing>	4	4			
<Dish>	0	0			
<Distance>	54	15	[UF: away, leave]		435
distinct	32	9	[UF: clear, definition]		
<Division>	140	19	[UF: apart, divvy, extracted, partition, separate]		325
DNA	5	2			
dock	2	1		[SN: Refers to a wharf.]	
dog	12	5			
dollar-sign	1	1			
dolphin	2	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Domain>	12	6	[UF: context, idea, theme]		319
dome	24	7			
<Domesticated>	0	0			147
dominant	106	19	[UF: especially, focus, main, major, predominant, significant]		
<Door>	4	1			259
doorknob	1	1			
<Dot>	442	21	[UF: point]		21
dotted	43	9			
<Doughnut>	1	1			81
down	853	21	[UF: descended]		
dragon	2	1			
<Drawing>	8	7			2
<Drawing-technique>	0	0			406
driftwood	5	1			
<Driveway>	2	1			268
drop	9	4			
drop-shadow	1	1			
dry	4	4			
<Dryness>	0	0			366
dull	0	0			
dusty	1	1			
<Ear>	9	6			178
<Earth>	5	3			134
east	5	3			
<Echinoderm>	0	0			158
<Eco-system>	0	0			308
edge	637	21	[UF: border, boundary, frame, margin]		
<Edge-condition>	0	0			
<Edging>	0	0		[SN: Refers specifically to clothing.]	95
<Educational-institution>	0	0			
egg	5	4	[UF: ovum		
Egypt	1	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Elbow>	2	1			183
<Electrical-component>	0	0			58
<Electricity>	2	1			143
<Electronics>	0	0			
elephant	1	1			
elephant-trunk	1	1			
<Elevation>	2	1			
embryo	1	1			
emerging	17	7	[UF: emanating, emitting, popping, sprouting]		
<Emphasis>	22	10	[UF: accent, feature, hallmark]		360
empty	41	12	[UF: bare, barren, blank, desolate, missing, vacant, void]		
<Enclosed-body>	0	0			
encroaching	4	1	[UF: invade]		
end	307	21	[UF: cutoff, dead end, finish, stop, ran out]	[SN: Refers to the end of a shape.]	
enough	4	3			
<Environment-component>	0	0			
<Equal>	51	16	[UF: equilateral, even]		358
<Equality>	0	0		[SN: Refers to a comparison and not a quantity.]	
equal-sign	1	1			
<Equilibrium>	1	1	[UF: compensate-for]	[SN: Refers to balance and order (the condition in which every part is in its right place).]	330
eraser	3	1			
everyday	0	0			
exact	285	21	[UF: dead,definite, just, right4, smack, specific, technical]		
<Existence>	0	0			318
exit	8	3		[SN: Indicates highway exit.]	

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
exploding	8	4	[UF: blast]		
exposed	8	4	[UF: showing]		
<Extension>	0	0			370
extremely	13	13	[UF: drastic, hard, too]		
<Extremity>	0	0			434
<Eye>	16	6			171
eyeball	2	1			
eyebrow	1	1			
eyelash	1	1			
eyelid	1	1			
<Face>	19	8			170
<Facial-expression>	0	0			177
<Facing>	37	13			449
factory	5	3			
fall-color	1	1			
far	83	16	[UF: beyond, faraway, farther, past]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
farm	5	5	[UF: farm-land]		
<Farmer>	1	1			223
<Fastening>	0	0			46
fat	2	1			
<Feather>	9	5			197
female	3	2			
femur	1	1			
<Fence>	27	3			266
festive	1	1			
few	23	8	[UF: fewer]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Fiber>	3	3	[UF: strand]		93
field	27	11			
<Figure>	6	6		[SN: Refers to the gestalt principle of figure-ground.]	11
figure8	1	1		[SN: Indicates the shape.]	

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
filled	158	19	[UF: filler, full, packed, populated, take-up]		
<Film>	0	0		[SN: Refers to still images only.]	9
fin	1	1			
<Finger>	16	6			187
fingerprint	7	4			
<Fire>	17	3	[UF: burning, campfire, flame]		142
<Fire-fighting>	0	0			56
fire-hydrant	2	1			
<Fish>	22	6			155
<Fist>	1	1			186
flag	11	3			
flared	6	4			
<Flesh>	1	1			203
flipper	4	1			
floating	5	4	[UF: hovering]		
<Flood>	1	1			141
<Floor>	6	3			255
Florida	7	2			
<Flower>	3	2			215
flowing	10	6	[UF: slinky, streaming]		
<Fluid-movement>	0	0			391
flying	8	6			
<Folded>	3	2			382
<Follicle>	2	2			195
<Food>	1	1		[SN: When notation is added this facet can be combined with the notation of the Plant facet to indicate plants used as food.]	80
foot	9	2	[UF: feet]	[SN: Unit of linear measure.]	
<Foot>	4	3	[UF: feet2]	[SN: The lower extremity of the leg.]	190
football	15	3			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
foothill	2	1			
foreground	5	4			
forest	22	4			
fork	3	3		[SN: A line breaking into two directions at less than 180 degrees.]	
<Format>	0	0			427
fossil	3	1			
<Fraction>	7	3	[UF: percent, proportion]	[SN: Specific values from this study are enumerated in this list.]	328
frayed	2	1	[UF: unravel]		
<Free-form>	6	3	[UF: amorphous]		37
french-fries	10	1	[UF: fries]		
<Frequency>	0	0			337
frequent	1	1			
frill	4	1			
fringe	2	2			
front	29	11			
<Fruit>	4	2			164
<Fruit-cocktail>	1	1			84
<Furniture>	0	0			
<Gaiety>	0	0			379
<Game>	1	1			65
gasoline	2	1			
<Gender>	0	0			218
<General-concept>	0	0			
<General-part>	0	0			433
<Generic-place>	0	0			
geography	2	1			
geology	2	1			
geometry	17	5			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Gestalt>	0	0		[SN: Attributes that contribute to the theory that the form of an image as a whole is different from and greater than the sum of its parts.]	
<Gift>>	2	1	[UF: present]	[SN: Refers to a gift-wrapped box.]	79
Giza	1	1			
<Glass>>	9	5			247
glasses	4	2			
glowing	11	4	[UF: fluorescent, iridescent, luminescent, neon, nuclear]	[SN: Refers to the color quality of nuclear energy.]	
goalpost	5	1			
<Golf-course>	1	1			279
good	11	7	[UF: better, decent, nice]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Goodness>	0	0			374
goopy	10	5	[UF: gel, jelly, sticky, viscous]		
Gradius-Three	1	1			
gradual	21	8	[UF: gentle, slow]		
<Grain>	0	0			166
grape	3	1			
grass	7	4			
gray	18	9	[UF: slate]		
Great-Wall-of-China	2	1			
Greece	1	1			
green	45	11	[UF: lime]		
grid	15	4			
<Grouped>	33	11	[UF: class, clump, cluster, collection, pack, set2]		334
grout	1	1			
growing	11	3			
<Gulf>	1	1			290

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
Gulf-of-Mexico	1	1			
gumdrop	1	1			
<Gun>	3	2			45
gun-sight	1	1	[UF: sight]	[SN: The symbol of a circle with a plus-sign in it.]	
H	11	1			
hail	1	1			
<Hair>	28	5	[UF: fur]		194
<Hair-style>	0	0			196
<Hand>	15	6			185
<Handle>	7	4			114
handmade	3	2			
<Harbor>	6	1		[SN: Refers to a natural harbor.]	292
hard	0	0			
<Hardness>	0	0			367
hash-mark	5	2			
hatchery	1	1			
<Head>	80	17			169
<Headdress>	2	1			101
headline	2	2			
heart	6	1			
heatwave	2	1			
<Height>	27	9		[SN: Refers to vertical axis.]	349
<Helipad>	2	1			229
<Hexagon>	1	1			35
hidden	4	3	[UF: eclipse]		
high	63	15	[UF: higher, tall, taller, tallest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
highlight	17	7	[UF: important, interest, point2, prominent, remarkable, standout, striking]		
<High-school>	1	1			234

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Highway>	58	16	[UF: freeway, superhighway]		270
<Hill>	20	9	[UF: mound]		304
hilltop	1	1			
<Holding>	1	1			443
hole	9	5	[UF: hollow]	[SN: Indents completely through.]	
honeycomb	1	1			
<Hook>	1	1			30
hooking	1	1			
horizon	5	1			
<Horizontal>	218	17	[UF laying, level, lie, lying, x-coordinate]		461
<Horizontal-perspective>	0	0			452
horn	3	2		[SN: Hard permanent outgrowth on an animal's head.]	
horse	6	2			
horseshoe	21	6			
Hot-Wheels	8	1			
hourglass	2	1			
house	50	16	[UF: home]		
hub	8	2			
<Hue>	12	8	[UF: shade]	[SN: Individual colors (hues) occurring in the raw data are grouped according to the eleven of the basic colors delineated by Berlin and Kay as typical of developed color vocabularies.]	417
huge	23	7	[UF: massive]		
<Human>	5	2			146
hump	12	3			
hypotenuse	3	2			
<Ice>	1	1			132
icicle	5	2			
<Image>	0	0			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Image-foundation>	0	0		[SN: Surface on which an image appears to reside.]	
imaginary	9	3	[UF: theoretical]	[SN: In the sense of visual structure and not interpretation.]	
imperfect	1	1	[UF: crude]		
improbable	0	0			
inch	544	17			
inconsequential	0	0		[SN: Lack of emphasis.]	
incorrect	1	1	[UF: wrong]		
<Increase>	13	9	[UF: amplified, expanding, repeat, over-and-over, times]		352
<Indentation>	19	5	[UF: cavity, concave, dent, depressed, gouge]		373
index-finger	12	2	[UF: forefinger]		
<Indexical>	0	0		[SN: Refers to the position of the object relative to itself within the context of the image.]	
India	1	1			
Indianapolis	6	2			
Indiana-University	1	1			
<Industrial>	1	1			243
infrared	1	1		[SN: A type of color film.]	
infrequent	3	2	[UF: once-in-a-while, sometimes]		
In-front-of	8	4		[SN: Refers to a place to the left of an object.]	
ink	6	3			
<Inlet>	1	1			293
<Insect>	6	3	[UF: bug]		151
<Inside-of>	1763	21	[UF: interior, within]		442
<Interconnection>	4	2			436
intermittent	2	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Intersecting>	41	11	[UF: intercepting]		437
<Interstate>	5	2			272
intertwined	5	4	[UF: caught-up, tied, twisty]		
intestine	6	2			
<Invisible>	8	2			322
irregular	28	9			
island	9	5			
Italy	5	1			
<Item-of-clothing>	0	0			94
jagged	45	16	[UF: choppy, torn]		
<Jail-bar>	1	1			264
<Jam>	4	1			85
Japan	4	3			
<Jaw>	3	1			173
Jefferson-Monument	3	2		[SN: Transliterated from Jefferson Memorial.]	
jelly-bean	1	1			
Jesus	1	1			
<Jewelry>	0	0			99
jigsaw	2	1			
<Joined>	347	21	[UF: against, attach, combine, conjoin,connect, fit, fuse, hit, meet, merger, skim, stick, together, touching, up-against, welded]		440
<Joint>	2	1			315
<Judgment>	0	0		[SN: Inherent conditions, not transitory.]	
keyhole	1	1			
kidney	1	1			
kidney-bean	2	1			
kimono	1	1			
<Kinds-of-change>	0	0		[SN: An indication of comparison.]	
<Kinds-of-image>	0	0			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Kinds-of-insect>	0	0			152
king	2	1			
Klee	1	1			
kneecap	1	1			
knife	4	2			
Kool-Aid	5	1			
L	6	2			
lace	4	1			
<Lagoon>	1	1			297
<Lake>	22	6			296
Lake-Monroe	1	1			
<Lamp>	0	0			105
<Land>	99	20	[UF: ground]		302
<Land-based-environment>	0	0			
<Land-based-substance>	0	0			
<Land-formation>	0	0			307
landscape	8	4		[SN: Refers to natural scenery.]	
landscape-format	4	3	[UF: landscape2]	[SN: Type of format.]	
<Land-water-formation>	0	0			300
large	462	21	[UF: big, bigger, biggest, giant, grand, great, greater, larger, largest, sizable]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
latitude	1	1			
lava	5	2			
lava-lamp	1	1			
<Layer>	26	7			454
<Leader>	0	0			221
<Leaf>	10	6			214
left	1524	21	[UF: before]	[SN: Analogous to the position of west on a person's body when they face north.]	
<Leg>	27	11		[SN: Animal part from hip to ankle.]	188

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
lemon	1	1			
<Length>	49	14		[SN: Refers to either the x or y axis.]	348
less	26	10	[UF: least]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Letter>	14	2		[SN: The set of 26 letters that constitute the Roman alpha-bet.]	14
lightning	8	2			
lima-bean	7	1			
<Line>	1286	21	[UF: guideline, trail]		22
<Line-curve>	0	0			
<Linguistic-quantity>	0	0		[SN: Non-specific.]	340
lip	6	2			
<Liquid>	8	5	[UF: fluid, melted, mol-ten]		119
<Living-organism>	4	2	[UF: organism]	[SN: Animals and plants]	144
lobster	1	1			
<Locale>	0	0			
long	242	20	[UF: longer, longest, bar]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
longitude	0	0			
<Loop>	2	2			27
Los-Angeles	2	1			
Louvre	3	3			
<Loveliness>	0	0			378
low	169	19	[UF: lower, lowest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
lowercase	0	0			
Lowes	1	1			
lowlight	1	1			
M	8	2			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
maggot	2	1			
magnification	7	4	[UF: closeup]		
male	6	5			
Malevich	1	1			
<Mammal>	0	0		[SN: By general category.]	
<Manmade>	1	1			392
many	214	21	[UF: bunch, lot, millions, multiple, so-much, ton, whole-bunch, zillion]		
<Map>	20	6			111
marble	3	2			
marina	1	1			
marker	3	3			
<Marking>	0	0			53
mashed	10	4	[UF: mushed, squished, squashed, squeezed]		
mass	9	6			
matte	16	1		[SN: Use of a material to mask or frame part of an image.]	
<Maturity>	0	0			219
McDonalds	3	1			
meandering	17	4	[UF: roam, wandering]		
<Measurement-unit>	0	0		[SN: Specific distance intervals.]	345
<Measuring>	0	0			54
<Mechanical-part>	0	0			
<Membrane>	4	2			193
messy	9	7	[UF: clutter, slapped, yucky]		
<Meteor>	5	1			125
Mexico	2	1			
microbe	1	1			
<Micro-organism>	0	0			160
microphone	2	1			
microscope	50	16	[UF: electron-microscope]		
middle-ground	1	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
mile	1	1			
<Military>	0	0			222
millimeter	19	4			
mine	1	1			
<Mineral>	1	1			137
mining-train	1	1			
miter	10	1			
mitochondria	2	2			
<Mixture>	7	4	[UF: miscellaneous]		120
<Modern-art>	9	3	[UF: contemporary, modern]		397
mohawk	3	1			
mold	2	1			
<Mollusk>	0	0			157
<Money>	0	0			107
<Monument>	8	5	[UF: memorial]		241
Monument-Center	3	2	[UF: Monument-Circle-Indy]		
<Moon>	17	7	[UF: full-moon]		123
moonscape	2	1			
more	309	21	[UF: additional, majority, most, much]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Mosaic>	11	3			4
mosquito	1	1			
moss	1	1			
motherboard	1	1			
mottled	1	1			
<Mountain>	47	14			305
mountain-range	22	7			
mouse	2	1	[UF: mice]		
<Mouth>	9	4			175
<Movement-reflecting-growth>	0	0			388
<Movie>	1	1			7
mud	1	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Multicolor>	6	5	[UF: colorful]	[SN: Indicates a selection of colors.]	418
<Multilane>	2	2			271
mummy	1	1			
<Municipality>	0	0			312
muscle	2	2			
<Museum>	1	1			236
mushroom	5	1			
muzzle	1	1			
<Mythical>	0	0			161
N	2	2			
<Natural>	6	5	[UF: natural-resources]		393
<Naturally-occurring-phenomena>	0	0		[SN: place it originates]	
<Natural-place>	0	0			
near	95	18	[UF: by, close, closer, closest, immediate, misses, nearest, shy]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
neat	2	2			
<Neatness>	0	0			375
<Neck>	20	7			179
<Neckwear>	0	0			100
needle	5	2			
negation	701	21	[UF: except, not, without]		
negative	6	4	[UF: CAT-scan, x-ray]		
neighborhood	4	2			
neuron	1	1			
new	13	3			
New-Mexico	1	1			
<Newness>	0	0			363
<Newspaper>	2	2			109
night	2	2			
night-vision-goggles	1	1			
<Non-domesticated>	0	0			148
none	10	4			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Non-edible>	0	0		[SN: Refers to normal human consumption.]	167
<Non-human>	0	0			
<Non-human-animal-appendage>	0	0			191
<Non-living-thing>	0	0			
<Normal>	0	0			
<Normalcy>	0	0			
north	6	3			
northeast	5	2			
northwest	4	2			
<Nose>	8	5			172
no-shadow	0	0			
<Nowhere>	1	1			316
nucleus	6	4			
<Number>	2	2		[SN: Isolates consist of numerical values.]	341
<Numeral>	0	0		[SN: The set of ten Arabic numerals, not values.]	16
O	1	1		[SN: Letter.]	
<OBJECT>	20	8			1
<Observatory>	5	2			240
obtuse-angle	3	2			
<Ocean>	13	8	[UF: sea]		289
Ocean-Spray	1	1			
<Octagon>	21	3			36
<Off>	271	21			447
Ohio-State	1	1			
oil	4	2			
old	1	1			
olive	11	1		[SN: Color.]	
<On>	913	21	[UF: upon, rest-on]		448
onion	3	2			
opaque	0	0			
open	13	9	[UF: gaping]		
<Opening>	0	0			314

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
openwork	1	1			
opposite	21	10	[UF: other2		
orange	43	10	[UF: peach]	[SN: Color.]	
<Ordered>	11	8	[UF: arranged, layout, network, setup, system]	[SN: Put in a certain position.]	333
<Organ>	0	0		[SN: Indicates body tissue.]	210
<Organic>	2	2			138
<Orientation>	25	8			430
<Outer-space>	2	2			121
<Outside-of>	355	21	[UF: exterior]		445
oval	88	12	[UF: ellipse]		
overlap	21	9	[UF: peek]		
Pac-man	1	1			
paint	48	7	[UF: oil paint]		
<Painting>	13	8			3
pale	73	18	[UF: faded, faint, light2, light2er, light2est]	[SN: Refers to light in color. Notation for comparatives: add -x2, or, -x3 to root word.]	
<Palette>	3	3	[UF: color scheme]	[SN: Indicates specific selection of colors used in an image.]	416
<Pan >	1	1			76
panel	13	1			
<Paper>	2	2			8
parabola	2	2			
parallel-to	88	16			
Paris	2	2			
<Park>	10	5			278
parking-garage	1	1			
parking-lot	8	4			
<Part>	354	21	[UF: portion, share]		324
<Partial-coverage>	0	0			457
<Part-of-clothing>	0	0			
pastel	4	1			
patch	39	13	[UF: blotch, splotch]		

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Pattern>	19	8		[SN: A repeated design.]	403
paw	1	1			
pea	3	2			
peace-symbol	2	1			
peak	21	8			
peanut	1	1			
pear	1	1			
pearl-necklace	1	1	[UF: pearls-on-a-string]		
<Peeled>	1	1			383
pencil	2	1			
peninsula	20	5			
perfect	27	12			
<Perfection>	0	0			376
perimeter	1	1			
<Perpendicular-to>	13	6			463
person	7	4	[UF: person]		
<Perspective>	5	4			414
petal	3	2			
<Photograph>	55	19	[UF: shot, snapshot]		6
Picasso	3	2			
picket	10	3			
<Pie>	4	1			83
<Piece>	61	14	[UF: fragment, sample]		327
pier	3	2			
pig	4	4			
<Pillar>	2	2			257
pin	28	1			
pink	15	6			
<Pitcher>	9	1			74
pixel	4	3			
<PLACE>	27	12	[UF: setting]		227
<Placement>	2	2			446
<Planet>	1	1			124
<Planetarium>	1	1			239

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Plant>	6	4	[UF: vegetation]	[SN: Notation can be combined with Food facet notation to indicate plants used as food.]	162
<Plant-body-part>	0	0			
<Plant-form>	0	0			163
<Plastic>	2	2	[UF: celluloid]		252
<Plate>	1	1			77
plateau	2	1			
<Platform>	5	2			256
<Play>	1	1			66
<Playground-equip-ment>	0	0			70
<Plaza>	1	1			283
plot	6	2			
plug	1	1			
plus-sign	11	4	[UF: addition sign, cross2]	[SN: The symbol +.]	
<Pod>	1	1			
<Pointing-to>	54	16			459
<Political-building>	0	0			242
<Polygon>	1	1		[SN: Closed figure with straight lines for all sides.]	31
ponytail	1	1			
pool-cue	3	1			
pop-art	8	1			
port	1	1		[SN: Refers to con-structed harbor.]	
portrait-format	3	2	[UF: portrait]		
<Position>	5	3	[UF: situate]	[SN: Position of ob-jects with reference to the frame and each other.]	428
possible	106	16	[UF: could, maybe, might]		
<Postage-stamp>	1	1			113
<Postcard>	4	1			112

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
potato	4	1			
pouring	7	2	[UF: dump, gushing]		
<Precipitation>	0	0			
<Prepared-food>	0	0			90
<Presentation>	0	0			429
<Primitive>	1	1			395
<Prism>	33	3			42
<Probability>	0	0			357
probable	37	15			
<Production-media>	0	0		[SN: Refers to media applied to a surface (image-foundation).]	401
<PROPERTY>	4	3	[UF: aspect, characteristic]		317
protoplasm	2	1			
<Protrusion>	43	11	[UF: convex, jut, projection, stick out, stick up]		372
psychedelic	6	3	[UF: trippy]		
<Publication>	0	0		[SN: Includes videos, movies, books, etc.]	
<Punctuation>	0	0			17
pupil	2	1			
pure	0	0			
<Pure-substance>	0	0			
<Purity>	0	0		[SN: Saturation.]	421
purple	9	3	[UF: mauve]		
<Puzzle>	4	1			69
<Pyramid>	92	12			41
Q	9	1			
quadrant	25	6			
<Quantity>	22	9	[UF: amount, body]		339
quarter	6	1		[SN: A 25 cent coin.]	
rabbit	1	1			
radiating	28	7			
radius	3	2			
railway	11	5	[UF: track3		
<Rain>	3	2			130
rainbow	4	3			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
raindrop	3	2			
Ramen-noodle	1	1			
<Random>	18	8	[UF: haphazard]		335
<Random-movement>	0	0			389
<Rate-of-change>	0	0			354
ravaged	4	1	[UF: bombed-out]		
ravine	3	2	[UF: gully]		
ray	70	16	[UF: beam, shaft, vector]		
real	70	19	[UF: actual, concrete, in-fact, literally]	[SN: In the sense of 'not artificial.']	
<Recreation>	0	0			
<Recreational-facility>	1	1	[UF: sporting-complex]		277
<Rectangle>	860	21	[UF: box, oblong, recti-linear]		34
red	108	16	[UF: red-dye]		
reef	1	1			
region	6	3			
regular	118	17	[UF: basic, basis, es-sential, quintessential, standard]	[SN: Typical form.]	
<Regularity>	0	0			332
<Relational>	0	0		[SN: Perspectives between objects regardless of frame orientation.]	
<Religious-figure>	0	0			220
Renoir	1	1			
<Reproduction>	0	0			209
<Reptile>	0	0			150
reservoir	2	1		[SN: Refers to con-structed lake.]	
<Residential-building>	4	2	[UF: residential]		232
rest	64	18	[UF: leftover, remaining]		
<Reversed>	22	8	[UF: backward, mirror, reflection, reverse]		432
rhinoceros	3	1			
rib	9	1			
ribbon	6	4			
rib-cage	2	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
rice	6	2			
ridge	13	4			
right	1458	21	[UF: after, counterclock-wise]	[SN: Analogous to the position of east on a person's body when they face north.]	
<Rind>	1	1			217
ripple	8	3	[UF: undulating]		
<River>	77	9			294
<Road>	266	19	[UF: thoroughfare, sec-ondary-road]		269
<Rock>	20	7	[UF: grit, pebble, gravel, stone]		135
<Role>	0	0			
Rome	1	1			
<Roof>	12	6			262
<Room>	5	3	[UF: chamber]		253
rooster	1	1			
<Root>	2	1			211
<Rotated>	6	1			431
rotted	1	1			
rotunda	3	3			
Rouault	1	1			
rough	7	3	[UF: grainy, harsh]		
roundabout	7	3			
row	21	8	[UF: range]		
rug	2	2	[UF: carpet]		
<Rung>	1	1			115
runway	8	1			
<Rural>	2	1			286
rusted	1	1			
S	10	4			
sailboat	1	1			
same	224	20	[UF: consistent, con-stant, echo, identical, identifiable, kind, match, sort, type, uniform]		
sand	6	2			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
sandwich	1	1			
San-Fernando-Valley	1	1			
satellite	20	8			
scaffolding	1	1			
<Scale>	3	1			198
scallop	19	2			
scarf	1	1			
scattered	19	9	[UF: dissipate, intersperse, sprinkled, spread, sprawled]		
<Science-fiction>	1	1			398
<Screen>	1	1		[SN: Surface of computer monitor.]	10
scribble	6	2			
scuba-diving	1	1			
seam	1	1			
<Seasonal-color>	0	0			420
sea-urchin	1	1			
seaweed	1	1			
<Section>	228	17	[UF: detail, sector, segment]		326
<Seed>	3	2			216
semi-circle	13	5			
Senate	1	1		[SN: Refers to specific place where U.S.Senate meets.]	
series	32	12	[UF: sequence, progression, alternating]		
serious	0	0			
<Sewing>	0	0			47
<Shaded>	71	13		[SN: Refers to darkened or graded tones]	425
<Shadow>	3	2			409
shallow	5	2			
<Shape>	686	21	[UF: form]		20
shape6	1	1		[SN: Indicates the shape.]	
shard	4	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
sharp	63	12	[UF: pointy, pointiest, sharper, steep]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Sharpness>	0	0			368
Shedd-Aquarium	1	1			
<Shelf>	1	1			104
<Shell>	12	4	[UF: seashell]		199
shingle	2	1			
shoe	23	7			
shooting	22	10	[UF: thrown, thrusted]		
<Shore>	11	3	[UF: shoreline]		299
short	44	15	[UF: quick, shorter, shortest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
shoulder	7	1			
shrimp	6	1			
shrunken	2	1			
side	941	21	[UF: leg2]		
sideways	114	14	[UF: lateral, over1]		
<Sign>	59	8	[UF: mark, track2]	[SN: Conventional figures.]	12
Sim-City	1	1			
similar	2670	21	[UF: analogy, as-if, interpretation, kind-of, kind-of-like, kind-of-looks-like, like, look-like, mimic, or-so, quasi-, recognizable-as, rather, reference, represent, re-mind, resemble, sense, so-to-speak, somewhat, sort-of, sort-of-like, sort-of-looks-like, tendency, variation, version]		
<Similarity>	0	0		[SN: Describes rela-tionship.]	338
simple	8	5	[UF: plain]		
<Simplicity>	0	0			331
<Size>	90	19			346

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Skeleton>	2	2			206
<Sketch>	2	2			407
<Skin>	9	3			192
skipping	3	3			
<Skull>	2	2	[UF: cranium]		207
<Sky>	11	8			288
skyscraper	2	2			
slanting	109	14	[UF: dip, lean, slope, skew, tilt]		
<Sleeve>	2	1			96
slide	18	4		[SN: For use under a microscope.]	
slip	3	1		[SN: Refers to a mooring.]	
slot	1	1			
small	877	21	[UF: little, mini-, miniature, smaller, smallest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
smile	4	2			
smiley-face	1	1			
smoke	11	5			
smokestack	1	1			
smooth	14	7	[UF: flat]		
<Smoothness>	0	0			369
snail	6	3			
snake	30	9			
snout	4	2			
<Snow>	4	1			131
snowflake	6	1			
soccer	2	2			
soccer-ball	1	1			
<Sociopolitical-location>	0	0		[SN: Proper nouns.]	
soft	3	2	[UF: fluffy, spongy]		
<Soil>	7	2	[UF: dirt]		136
solar-flare	1	1			
<Solid>	13	5			118
some	198	21	[UF: several, various]		

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
south	10	4			
southeast	6	2			
southwest	13	3			
space	95	16	[UF: break, gap, interstice]		
space-shuttle	1	1			
spaghetti	3	2			
sparrow	1	1			
sparse	0	0			
<SPATIAL-LOCATION>	2	2	[UF: location]	[SN: Directional or referential positions.]	426
<Specific-business>	0	0			245
<Specific-dimension>	0	0			343
<Spectrum>	1	1		[SN: Indicates a rainbow palette of seven colors.]	419
Spensers	1	1			
sperm	1	1			
<Sphere>	23	3	[UF: ball, globe, orb]		40
spider	22	8			
spikey	15	5			
spinal-cord	5	1			
<Spiral>	5	4	[UF: whorly]		26
spit	11	1			
splash	4	3			
splatter	27	5			
spoke	1	1			
spoon	7	1			
<Sports>	0	0			
<Sports-equipment>	0	0			63
spot	34	15	[UF: speck]		
spraying	10	4	[UF: fountain, jet, spew, spouting, squirt]		
spring-color	0	0			
square	367	21		[SN: A rectangle with four equal sides.]	
squiggle	185	20	[UF: squibble, wiggle]		
<Stadium>	13	4			282

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Stages-of-insect>	0	0			153
stained-glass	14	2			
<Stair>	2	1	[UF: step]		254
stalagmite	2	1			
star	15	6			
starburst	13	5			
starfish	1	1			
Starlight	1	1		[SN: Title of Van-Gogh painting.]	
start	374	21	[UF: beginning, initial, origin, source]	[SN: Refers to the beginning of a shape.]	
Startrek	1	1			
<State>	7	2			311
steady	2	2			
<Stem>	2	2			213
stick	11	3			
stick-figure	7	3			
stitching	13	1			
stop-sign	2	1			
<Storm>	3	1			133
straight	305	21	[UF: direct, due]		
strange	35	12	[UF: funky, funny, odd, weird, weirdest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
strap	2	1			
strawberry	3	1			
streak	43	8	[UF: smear, stroke, wiped]		
<Stream>	5	2			295
street	88	19	[UF: avenue, drive, lane]		
<Strength>	0	0			380
stretched	159	15	[UF: elongated, extend]		
string	1	1			
strip	15	3			
striped	27	6	[UF: striation]		
strong	5	3	[UF: powerful]		

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
subdivision	9	4	[UF: development, housing-development]		
submarine	1	1			
<Substance>	27	9	[UF: filler, junk, material, stuff]		117
subtle	1	1			
suburb	1	1			
summer-color	1	1			
<Sun>	8	4			122
sunburst	2	2			
sunflower	1	1			
sunlight	9	5			
sunset	1	1			
sunspot	1	1			
superimposed	30	9	[UF: ahead, in front of, over3]	[SN: Refers to one thing covering another.]	
Supernintendo	1	1			
surface	27	8	[UF: face2, plane, sheet, slick, spill]		
surfboard	11	1			
<Surrounding>	218	21	[UF: around, confine, encircling, enclosing, encompass, housed, wrap, wraparound]		444
swastika	2	1			
sweeping	22	7	[UF: billow, swish, swoop, swoosh]		
<Sweet>	0	0			
<Swimming-pool>	3	2			280
swiss-cheese	1	1			
<Symbol>	5	3		[SN: Kind of figure other than numeric, alphabetic, or punctuation.]	18
symmetrical	5	4			
<Table>	1	1			103
tail	8	5			
tapered	3	2			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
target	10	5	[UF: bull's-eye]		
teacup	1	1			
teardrop	10	2			
<Technique>	0	0			
<Teeth>	10	3			176
telephone	3	1			
tentacle	2	1			
tepee	1	1			
terrace	3	1			
<Terrain>	4	3			301
Texas	7	2			
<Textile>	1	1		[SN: Textile does not include clothing.]	91
<Textile-method>	0	0			410
<Text-string>	0	0		[SN: Combinations of numerals, letters, and/or punctuation.]	19
thermal	3	2		[SN: A type of color film.]	
thermometer	1	1			
thick	99	19	[UF: fat2, fat2test, thicker, thickest]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
thin	99	18	[UF: narrow, skinny, slender, thinner]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
thorn	1	1			
thread	6	4			
three-dimensional	7	6			
<Three-dimensional-shape>	1	1	[UF: volume]		38
throat	3	3			
<Through>	63	17			438
thumb	2	1			
tibia	1	1			
tidal-wave	3	1			
<Tide>	1	1			140

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Tile>	4	2			249
tiny	104	20	[UF: baby, hair2, hint, itty-bitty, smidgen, te-ench, teeny]	[SN: Very small quantity.]	
<Tissue>	6	5		[SN: Refers to coher-ent collections of specialized cells.]	201
toast	5	1			
toe	6	4			
<Tone>	23	6	[UF: light-degree]	[SN: Refers to intensity (treatment of light and dark in image.)]	422
<Tool>	0	0		[SN: Based on what is done with it.]	
top	1152	21	[UF: apex, crest, tip, tiptop]		
topography	9	1			
<Torso>	0	0			180
toucan	1	1			
town	28	8	[UF: community, settle-ment]		
<Toy>	0	0			71
<Track>	2	2		[SN: Prepared course for racing.]	281
<Traffic>	1	1		[SN: Refers to land-based travel.]	267
<Traffic-junction>	0	0		[SN: Related to the intersections of roads and highways]	274
trailer	4	1	[UF: mobile-things-peo-ple-live-in]		
trailer-park	2	1			
<Train>	0	0			62
<Transparency>	0	0			424
transparent	4	3	[UF: see-through		
tree	34	12			
trench	1	1			
<Triangle>	345	20	[UF: wedge]		32
Trivial-Pursuit	1	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Trunk>	2	2			212
tube	36	7	[UF: drainage-pipe, pipe, pipeline]		
tumbleweed	2	1			
<Tunnel>	1	1			276
turtle	1	1			
TV	1	1			
two-dimensional	1	1			
<Two-dimensional-shape>	0	0			
<Type-of-sport>	0	0			64
U	10	5			
ugly	1	1	[UF: horrible]		
unaligned	5	3	[UF: offset, shifted]		
<Unchanged>	0	0			
underneath	32	12	[UF: back3]	[SN: Refers to the location of a thing being covered.]	
<Unenclosed-Body>	0	0		[SN: Refers to bodies of water.]	
<Unequal>	1	1	[UF: uneven]		359
United-States	2	1			
<University>	2	1			235
up	825	20	[UF: upswing]		
upright	18	7	[UF: standing, sunny-side-up]		
upside-down	36	10	[UF: inverted, right-round, vice-versa]		
<Urban>	4	2			287
V	29	8			
vacuole	11	2			
<Validation>	0	0			320
<Valley>	5	2			303
Van-Gogh	1	1			
vanishing-point	8	1			
vapor-trail	2	1			
<Vault>	1	1			261
<Vegetable>	0	0			165

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Vehicle>	0	0			
vein	12	5			
<Vertical>	145	16			460
<Vertical-perspective>	0	0			453
very	581	21	[UF: pretty, quite, really, so, terribly, terribly, well2]	[SN: To considerable extent.]	
<View>	66	17			411
<Viewing>	0	0			55
<Violent-movement--non-fluid>	0	0			390
<Visibility>	0	0			
<Visible>	20	10	[UF: discern, obvious, seen, vision]		321
<Volcano>	5	2			306
vulture	5	2	[UF: buzzard, turkey vulture]		
W	1	1			
<Walkway>	2	1			265
<Wall>	43	7			258
<Wart>	3	2			200
Washington-DC	5	4			
water	108	14			
<Water-based-environ-ment>	0	0			228
watercolor	8	5			
<Water-event>	0	0			
<Wave>	70	11			139
weak	0	0			
<Weather>	3	2			127
weaving	3	2	[UF: woven]		
web	4	2			
west	10	4			
wet	0	0			
whale	4	2			
<Wheel>	12	4	[UF: tire]		116
white	84	15	[UF: off-white]		
whitecap	3	1			
White-House	4	1			

Term [<> indicates facet]	Freq. count	Pair count	Synonyms and near synonyms indicated by used-for [UF]	Scope Note	Concept number
<Whole>	132	21	[UF: all-the-way, complete, everywhere, intact, throughout]		323
<Whole-coverage>	0	0			456
wide	108	19	[UF: broad, broadest, stubbier, thick, thicker, thickest, wider]	[SN: Notation for comparatives: add -x2, or, -x3 to root word.]	
<Width>	0	0			347
wig	1	1			
wild	4	1			
<Wind>	4	2			128
<Window>	12	3			260
wing	7	3			
winter-color	0	0			
<Wire>	2	2			251
wisp	5	4			
<Within-water-body>	0	0			298
<Wood>	16	8			246
word	3	2	[UF: language, talk]	[SN: A meaningful text-string.]	
word-balloon	4	1	[UF: speech-bubble]		
<Worm>	8	4			154
wrinkled	3	3	[UF: crumple, rumpled]		
<Wrist>	2	1			184
X	94	8			
Y	24	2			
<Yard>	1	1			285
yarn	3	1			
yellow	59	11	[UF: gold, lemon2]		
yin-yang	1	1			
yolk	1	1			
zebra	12	1			
zigzag	5	4			
<Zipper>	2	1			97
zone	7	1			

Appendix E
Controlled vocabulary for describing image structure

This controlled vocabulary for describing image structure was created from the natural-language generated by 42 subjects when describing images from three domains: satellite imagery, abstract art, and photo-microscopy. It begins with a facet outline for ease of reference.

Angle brackets <> are used to indicate facet names, *italic* is used to indicate organizing-superordinates and conceptual-antonyms that were added by the researcher to complete the logical structure or syntactic sensibility.

The main facets are Object, Place, Property, and Spatial Location. The vocabulary is listed in hierarchical form with order in array. The primary guideline for the order in array was to organize by what is most important in the image structure domain. Within the facets, sub-facets and isolates the order was based on any of the following strategies applied as appropriate to the concept group: big to little (mold to microbe), general to specific (skeleton to tibia), top to bottom (thorn to root), most common to least common (rooster to toucan). Alphabetical order was only used for certain groups of proper nouns, such as *<Computer-game>* or *<Municipality>*.

OBJECT

- ***<Image>***
  - *<Kinds-of-image>*
    - <Drawing>
    - <Painting>
      - Starlight [SN: Title of painting by Van-Gogh.]
    - <Mosaic>
    - <Collage>
    - <Photograph>
    - <Movie>
      - Startrek
  - <*Image-foundation*> [SN: Surface on which an image appears to reside.]
    - <Paper>
    - <*Film*> [SN: Refers to still images only.]
      - Negative
      - Thermal [SN: A type of color film.]
      - Infrared [SN: A type of color film.]
    - <Screen> [SN: Surface of computer monitor.]
- ***<Non-living-thing>***
  - <Figure> [SN: Refers to the gestalt principle of figure-ground.]
    - <Sign> [SN: Conventional figures.]
      - <Alphabet>
        - <Letter> [SN: The set of 26 letters that constitute the Roman alphabet.]
        - <*Case*>
          - Capital
          - *Lowercase*
      - <*Numeral*> [SN: The set of ten Arabic numerals, not quantities.]
        - Shape6
        - Figure8
      - <*Punctuation*>
        - Asterisk
        - Dash
        - Hash-mark
        - Bracket
        - Checkmark
        - Plus-sign
        - Equal-sign
        - Dollar-sign
      - <Symbol> [SN: Kind of figure that is not numeric, alphabetic, or punctuation.]
        - Yin-yang
        - Peace-symbol
        - Smiley-face
        - Gun-sight
        - Word-balloon
        - Target
        - Stop-sign
        - Heart
        - Teardrop
        - Keyhole

        - Horseshoe
        - Star
        - Starburst
        - Sunburst
        - Swastika
        - Stick-figure
        - Arrow
    - <*Text-string*> [SN: Combination of numerals, letters, and/or punctuation.]
        - Word [SN: A meaningful string.]
- <Shape>
    - <*Two-dimensional-shape*>
        - <Dot>
        - <Line>
            - Wisp
            - Ray
            - Streak
            - Squiggle
            - Ripple
            - Zigzag
            - Jagged
        - <Angle> [SN: Specific angles are preceded by a numeral, e.g., 90 degrees.]
            - Branch [SN: A line off to side of a straight line at less than 90 degrees.]
            - Fork [SN: A line breaking into two directions at less than 180 degrees.]
            - Corner [SN: A 90-degree angle.]
            - *Obtuse-angle* [SN: Angles greater than 90 and less than 180 degrees.]
        - <Curve>
            - <*Closed-curve*>
                - Circle
                - Oval
                - Crescent
            - <*Line-curve*>
                - <Spiral>
                - <Loop>
                - <Curl>
                - <Arc>
                    - Arch
                    - Semi-circle
                    - Parabola
                - <Hook>
        - <Polygon> [SN: Closed figures with straight lines for all sides.]
            - <Triangle>
            - <Diamond>
            - <Rectangle>
                - Square [SN: A rectangle with four equal length sides.]
            - <Hexagon>
            - <Octagon>
        - <Freeform>
            - Spot
            - Blob
            - Splash
            - Splatter
            - Patch

- - - - Mass
    - <Three-dimensional-shape> [SN: Represented two-dimensionally.]
      - <Block> [SN: A rectangular shaped solid.]
        - Cube [SN: A square shaped solid.]
      - <Sphere>
        - Bubble
      - <Pyramid>
      - <Prism>
      - <Cone>
      - <Cylinder>
        - Tube
- <*Artifact*> [SN: Material objects.]
  - <Gun>
    - Bullet
  - <*Tool*> [SN: Based on what is done with it.]
    - <*Fastening*>
      - Strap
      - Pin
    - <*Sewing*>
      - Needle
    - <*Constructing*>
      - Scaffolding
    - <Blacksmithing>
      - Anvil
    - <*Cutting-tool*>
      - <Blade>
        - Knife
        - Dagger
      - <Axe>
    - <*Cooking*>
      - Spoon
    - <*Marking*>
      - Marker
      - Brush
      - Eraser
    - <*Measuring*>
      - Hourglass
      - Thermometer
    - <*Viewing*>
      - Glasses
      - Night-vision-goggles
      - Microscope
      - Slide [SN: For use under a microscope.]
    - <*Fire-fighting*>
      - Fire-hydrant
  - <*Electronics*>
    - <*Device*>
      - Microphone
      - Telephone
      - TV
    - <*Electrical-component*>

- Plug
- Battery
- Motherboard
- Circuit
- Computer-chip
- Mouse

- *<Vehicle>*
  - *<Aircraft>*
    - Satellite
    - Space-shuttle
  - <Boat>
    - Sailboat
    - Submarine
  - <Car>
    - Convertible
  - *<Train>*
    - Mining-train
    - Railway
- *<Recreation>*
  - *<Sports>*
    - *<Sports-equipment>*
      - Baseball-bat
      - Goalpost
      - Soccer-ball
      - Flipper
      - Surfboard
      - Pool-cue
    - *<Type-of-sport>*
      - Baseball
      - Football
      - Soccer
      - Scuba-diving
  - <Game>
    - <Play>
    - *<Board-game>*
      - Trivial-Pursuit
    - *<Computer-game>*
      - Gradius-Three
      - Pac-man
      - Sim-City
      - Supernintendo
  - <Puzzle>
    - Jigsaw
  - *<Playground-equipment>*
    - Children’s-slide
  - *<Toy>*
    - Hot-Wheels
- <Container>
  - *<Dish>*
    - <Cup>
      - Teacup
    - <Pitcher>

- <Bottle>
- <Bowl>
- <Pan>
- <Plate>

<Bag>

<Gift> [SN: Refers to a gift-wrapped box.]

<Food> [SN: When notation is added this facet can be combined with the notation of the Plant facet to indicate plants used as food.]

- <*Sweet*>
  - <Doughnut>
  - <Cupcake>
  - <Pie>
  - <Fruit-cocktail>
  - <Jam>
  - <*Candy*>
    - Gumdrop
    - Jelly-bean
    - Candy-cane
- <*Beverage*>
  - Kool-aid
  - Ocean-Spray
- <*Cheese*>
  - Swiss-cheese
- <Bread>
- <*Prepared-food*>
  - Toast
  - Sandwich
  - French-fries
  - Ramen-noodles
  - Spaghetti

<Textile> [SN: Textile does not include clothing.]

- <Cloth>
  - Rug
  - Blanket
  - Curtain
  - Flag
- <Fiber>
  - Thread
  - Cord
  - Ribbon
  - String
  - Yarn

<*Clothing*>

- <*Item-of-clothing*>
  - Kimono
  - Shoe
  - Boot
- <*Part-of-clothing*>
  - <*Edging*> [SN: Refers specifically to clothing.]
    - Frill
    - Fringe
    - Scallop

      - Lace
    - <Sleeve>
    - <Zipper>
  - <*Accessory*>
    - <Badge>
    - <*Jewelry*>
      - Pearl-necklace [SN: Transliterated from pearls-on-a-string.]
    - <*Neckwear*>
      - Scarf
      - Bowtie
    - <Headdress>
      - Bonnet
      - Crown
      - Wig
- <*Furniture*>
  - <Bed>
  - <Table>
  - <Shelf>
  - <*Lamp*>
    - Lava-lamp
- <*Burial-artifact*>
  - Mummy
  - Coffin
- <*Money*>
  - Quarter2 [SN: A 25 cent coin.]
  - Dime
- <*Publication*> [SN: Includes videos, movies, books, etc.]
  - <Book>
  - <Newspaper>
    - Headline
  - <Article> [SN: Piece of writing on a specific topic.]
  - <Map>
  - <Postcard>
  - <Postage-stamp>

<*Mechanical part*>

- <Handle>
- <Rung>
- <Wheel>
  - Spoke

<Substance>

- <*Pure-substance*>
  - <Solid>
  - <Liquid>
    - Water
    - Oil
- <Mixture>
  - Gasoline

<*Naturally-occurring-phenomena*> [SN: Place it originates.]

- <Outer-space>
  - <Sun>

- Sunspot
- Solar-flare
- <Moon>
- <Planet>
- <Meteor>

- <*Atmosphere*> [SN: Phenomena not specifically related to weather conditions.]
  - Air
  - Smoke
  - Vapor-trail
  - Heat-wave
  - Sunlight
  - Sunset
  - Night
- <Weather>
  - <Wind>
    - Breeze
  - <Cloud>
  - *<Precipitation>*
    - <Rain>
      - Raindrop
      - Hail
      - Rainbow
    - <Snow>
      - Snowflake
      - Avalanche
    - <Ice>
      - *Icicle*
  - <Storm>
- <Earth>
  - *<Land-based-substance>*
    - <Rock>
      - Lava
      - Marble
      - Fossil
    - <Soil>
      - Clay
      - Mud
      - Sand
    - <Mineral>
      - Crystal
    - <Organic>
      - Driftwood
  - <*Water-event*>
    - <Wave>
      - Whitecap
      - Tidal-wave
    - <Tide>
    - <Flood>
  - <Fire>
- <Electricity>
  - Lightning

- **<Living-Organism> [SN: Animals and plants.]**
  - <Animal-life> [SN: Delineated by general category, includes mythical creatures.]
    - <*Mammal*>
      - <Human>
        - Person
      - <*Non-human*>
        - <*Domesticated*>
          - Cat
          - Dog
          - Horse
          - Cow
          - Pig
          - Rabbit
        - <*Non-domesticated*>
          - Dolphin
          - Whale
          - Elephant
          - Zebra
          - Rhinoceros
    - <Bird>
      - Rooster
      - Sparrow
      - Vulture
      - Toucan
    - <*Reptile*>
      - Snake
      - Turtle
    - <Insect>
      - <*Kinds-of-insect*>
        - Butterfly
        - Mosquito
        - Ant
        - Spider
      - <*Stages-of-insect*>
        - Maggot
    - <Worm>
    - <Fish>
    - <*Crustacean*>
      - Lobster
      - Shrimp
    - <*Mollusk*>
      - Snail
      - Clam
    - <*Echinoderm*>
      - Sea-urchin
      - Starfish
    - <*Cnidarian*> [SN: Previously known as Coelenterate.]
      - Coral
    - <*Micro-organism*>
      - Mold
      - Amoeba

    - Bacteria
    - Microbe
  - <*Mythical*>
    - Dragon

- <Plant> [SN: When notation is added, this facet can be combined with the notation of the Food facet to indicate plants used as food.]
  - <*Plant-form*>
    - Tree
    - Bush
    - Grass
  - <Fruit>
    - Apple
    - Pear
    - Grape
    - Strawberry
    - Banana
    - Cantaloupe
    - Lemon
  - <*Vegetable*>
    - Corn
    - Pea
    - Kidney-bean
    - Lima-bean
    - Olive
    - Cauliflower
    - Onion
    - Peanut
    - Potato
    - Mushroom
    - Seaweed
  - <*Grain*>
    - Rice
  - <*Non-edible*> [SN: Refers to normal human consumption.]
    - Cactus
    - Sunflower
    - Clover
    - Tumbleweed
    - Moss

- <Body> [SN: Includes all parts of living organisms.]
  - <*Animal-body-part*>
    - <Head>
      - <Face>
        - <Eye>
          - Eyebrow
          - Eyelid
          - Eyelash
          - Eyeball
          - Pupil
        - <Nose>
          - Beak

      - Snout
      - Elephant-trunk
      - Muzzle
    - <Jaw>
    - <Chin>
    - <Mouth>
      - Lip
    - <Teeth>
    - <*Facial-expression*>
      - Smile
  - <Ear>
- <Neck>
  - Throat
- <*Torso*>
  - Shoulder
  - Chest
  - Rib-cage
- <Appendage>
  - <Arm>
    - <Elbow>
    - <Wrist>
    - <Hand>
      - <Fist>
      - <Finger>
        - Thumb
        - Index-finger
        - Fingerprint
  - <Leg>
    - <Ankle>
    - <Foot> [SN: The lower extremity of the leg.]
      - Toe
  - <*Nonhuman-animal-appendage*>
    - Antenna
    - Horn
    - Hump
    - Paw
    - Claw
    - Tentacle
    - *Wing*
    - Fin
    - Tail
- <*Covering*> [SN: The coverings on animal objects and organic things occurring on those coverings.]
  - <Skin>
  - <Membrane>
  - <Hair>
    - <Follicle>
    - <*Hair-style*>
      - Buzz-cut
      - Mohawk
      - Ponytail
  - <Feather>

    - <Scale>
    - <Shell>
      - Conch
    - <Wart
  - <Tissue> [SN: Refers to coherent collections of specialized cells.]
    - <Cell>
      - Protoplasm
      - Cytoplasm
      - Vacuole
      - Nucleus
      - Chromosome
      - DNA
      - Mitochondria
      - Neuron
    - <Flesh>
      - Muscle
      - Fat
  - <Blood>
    - <Blood-vessel>
      - Artery
      - Vein
      - Capillary
  - <Skeleton>
    - <Skull>
    - <Bone>
      - Rib
      - Femur
      - Kneecap
      - Tibia
  - *<Reproduction>*
    - Sperm
    - Egg
    - Embryo
    - Yolk
  - *<Organ>*
    - Brain
    - Spinal-cord
    - Kidney
    - Intestine
- *<Plant-body-part>*
  - <Root>
    - Bulb
  - <Trunk>
  - <Stem>
    - Thorn
  - <Leaf>
  - <Flower>
    - Petal
  - <Seed>
    - Bean
  - <Pod>
  - <Rind>

- *<Aspects-of-living-thing>*
  - *<Gender>*
    - Female
    - Male
  - *<Maturity>*
    - Child
    - Adult
  - *<Role>*
    - *<Religious-figure>*
      - Jesus
    - *<Leader>*
      - King
      - Captain
    - *<Military>*
      - Army
    - <Farmer>
    - <Architect>
    - <Artist>
      - Dali
      - Klee
      - Malevich
      - Picasso
      - Renoir
      - Rouault
      - Van-Gogh
    - <Dancer>

PLACE

- ***<Constructed-environment>***
  - *<Water-based-environment>*
    - Canal
    - Reservoir [SN: Refers to constructed lake.]
    - Port [SN: Refers to constructed harbor.]
    - Marina
    - Pier
    - Dock [SN: Refers to a wharf.]
    - Slip [SN: Refers to a mooring.]
    - Boatyard [SN: Refers to a dry-dock.]
  - *<Land-based-environment>*
    - *<Air-travel-environment>*
      - <Helipad>
      - <Airport>
        - Runway
    - <Building> [SN: Types of buildings.]
      - <Residential-building>
        - House
        - Trailer
        - Tepee
      - *<Cultural-building>*
        - <Church>

    - <*Educational-institution*>
      - <High-school>
      - <University>
        - Indiana-University
        - Ohio-State
    - <Museum>
      - Louvre
    - <Amphitheater>
    - <Aquarium>
      - Shedd-Aquarium
    - <Planetarium>
    - <Observatory>
    - <Monument>
      - Jefferson-Monument
      - Monument-Center [SN: Indianapolis.]
      - Great-Wall-of-China
  - <*Political*>
    - Capitol
    - White-House
    - Senate [SN: Refers to where the Senate meets.]
  - <*Business*>
    - <Industrial>
      - Mine
      - Factory
      - Hatchery
    - <*Business-structure*>
      - Skyscraper
      - Parking-garage
      - Parking-lot
    - <*Specific-business*>
      - Lowes
      - McDonalds
      - Spensers
- <Construction-site>
- <*Building-element*>
  - <*Construction-material*>
    - <Wood>
      - Board
    - <Glass>
      - Stained-glass
    - <Brick>
    - <Tile>
    - <Cement>
      - Grout
    - <Wire>
    - <Plastic>
  - <Room>
    - Bathroom
    - Rotunda
  - <*Building-component*>
    - <Stair>
    - <Floor>

    - <Platform>
    - <Pillar>
    - <Wall>
    - <Shelf>
    - <Door>
      - Doorknob
    - <Window>
    - <Vault>
      - Dome
    - <Roof>
      - Shingle
    - <*Chimney*>
      - Smokestack
    - <Jail-bar>
- <*Environment-component*>
  - <Walkway>
  - <Fence>
    - Picket
- <Traffic> [SN: Refers to land-based travel]
  - <Driveway>
  - <Road>
    - Street
    - Bypass
  - <Highway>
    - <Multilane>
    - <*Interstate*>
      - <Beltway>
        - Beltway 465 [SN: Transliterated from I-465.]
  - <*Traffic-junction*> [SN: Related to intersections of roads.]
    - Roundabout
    - Cloverleaf
    - Exit [SN: Indicates highway exit]
  - <Bridge>
  - <Tunnel>
- <Recreational-facility>
  - <Park>
    - Big-Bend
  - <Golf-Course>
  - <Swimming-pool>
  - <Track>
  - <Stadium>
    - Bleacher
  - <Plaza>
  - <Courtyard>
  - <Yard>

<*Locale*>

- <Rural>
  - Farm
- <Urban>
  - City
  - Town

Suburb
Subdivision
Trailer-park
Neighborhood
City-block

<Natural-place>

<Sky>
- Horizon

<Body-of-water>
- *<Unenclosed-body>*
  - <Ocean>
    - Atlantic-Ocean
  - <Gulf>
    - Gulf-of-Mexico
  - <Bay>
  - <Harbor> [SN: Refers to a natural harbor.]
  - <Inlet>
  - <River>
  - <Stream>
- *<Enclosed-body>*
  - <Lake>
    - Lake-Monroe
  - <Lagoon>
- *<Within-water-body>*
  - Channel

<Shore>
- Bank
- Beach
- Coastline

<Land-water-formation>
- Island
- Archipelago
- Peninsula
- Spit
- Reef

<Terrain>
- <Land>
- <Elevation>
  - <Valley>
    - San-Fernando-Valley
  - <Hill>
    - Foothill
    - Hilltop
  - <Mountain>
  - <Volcano>
- *<Land-formation>*
  - Stalagmite
  - Cave

Crater
Trench
Ravine
Cliff
Terrace
Plateau
Ridge
Peak
Mountain-range

<*Eco-system*>

Desert
Field
Forest

<*Sociopolitical-location*> [SN: Proper nouns.]

<Continent>

Africa

<*Country*>

Afghanistan
Egypt
Greece
India
Italy
Japan
Mexico
United-States

<State>

Florida
New-Mexico
Texas

<*Municipality*>

Chicago
Cincinnati
Giza
Indianapolis
Los-Angeles
Paris
Rome
Washington-DC

<*Generic-place*>

<Area>

Region
Zone
Plot

<*Opening*>

Space
Crack
Slot

<Joint>

- Seam
- Miter
- Hub

<Nowhere>

PROPERTY

<General-concept>

<Existence>

- Real [SN: In the sense of not artificial.]
- Imaginary [SN: In the sense of visual structure and not interpretation.]

<Domain>

- Art
- Biology
- Botany
- Chemistry
- Geometry
- Geography
- Geology
- Topography

<Validation>

- *Affirmation*
- Negation

<Attribute>

<Visibility>

- <Visible>
  - Exposed
- <Disappearing>
- <Invisible>
  - Hidden

<Gestalt> [SN: Attributes that contribute to the theory that the form of an image as a whole is different than and greater than the sum of its parts.]

- *<Completeness>*
  - <Whole>
  - <Part>
    - <Division>
      - <Section>
        - Bisection
        - Panel
        - Quadrant
        - Compartment
      - <Piece>
        - Strip
        - Stick
        - Shard
        - Chip
        - Drop

          - Bit
      - <Fraction> [SN: Specific values from the study are enumerated in Appendix D.]
  - <*Closure*>
    - Open
    - Closed
  - <Equilibrium> [SN: Refers to balance and order (the condition in which every part is in its right place).]
    - <*Simplicity*>
      - Simple
      - Complex
    - <*Regularity*>
      - Regular [SN: Typical form.]
      - Symmetrical
      - Asymmetrical
      - Irregular
  - <*Arrangement*> [SN: Indicates proximity in gestalt theory.]
      - <Ordered>
        - Series
        - Consecutive-order [SN: Transliterated from consecutive.]
      - <Grouped>
      - <Random>
  - <*Continuation*>
    - <*Alignment*>
      - <Aligned>
        - Row
        - Column
      - <Unaligned>
    - <*Frequency*>
      - Frequent
      - Alternating
      - Skipping
      - Intermittent
      - Scattered
      - Infrequent
  - <*Similarity*> [SN: Describes relationship.]
    - Same
    - Similar
    - Different
    - Opposite

- <Quantity>
  - <*Linguistic-quantity*> [SN: Non-specific.]
    - A [SN: The article.]
    - All
    - Enough
    - Many
    - Some
    - Rest
    - Few
    - None
  - <Number> [SN: Isolates consist of the numerical values of whole numbers. Specific

values that occurred in the natural language are enumerated in Appendix A. Words for numbers are represented by their numeric sign (e.g., eight is 8, half is ½).]

- <Dimension>
  - <*Specific-dimension*>
    - Radius
    - Diameter
    - Hypotenuse
    - Perimeter
    - Circumference
  - <*Dimensionality*>
    - Two-dimensional
    - Three-dimensional [SN: Refers to representation of a z-axis.]
- <*Measurement-unit*> [SN: Specific distance intervals.]
  - Millimeter
  - Centimeter
  - Inch
  - Foot-measure [SN: Unit of linear measure.]
  - Mile
  - Latitude
  - *Longitude*
  - Pixel
- <*Comparative-measure*>
  - <Size>
    - Tiny
    - Small
    - *Average*
    - Large
    - Huge
  - <*Width*>
    - Thin
    - Wide
    - Thick
  - <Length> [SN: Refers to horizontal axis.]
    - Short
    - Long
  - <Height> [SN: Refers to vertical axis.]
    - Low
    - High
  - <Depth>
    - Shallow
    - Deep

<Comparison>

- <*Kind-of-change*> [SN: Not a quantity but an indication of comparison]
  - <Increase>
    - Stretched
  - <*Decrease*>
    - Shrunken
- <*Rate-of-change*>
  - Gradual
  - Steady
- <Degree> [SN: Adverbs must be included due to prevalence in raw data.]

  - Barely
  - Almost
  - Very
  - Extremely
- <*Certitude*>
  - Approximate
  - Exact
- <*Probability*>
  - *Improbable*
  - Probable
  - Possible
- <*Equality*> [SN: Refers to a comparison and not a quantity.]
  - <Equal>
  - <Unequal>
    - Less
    - More
- <Emphasis>
  - Dominant
  - Highlight
  - Lowlight
  - *Inconsequential*

<*Condition*> [SN: State, or, figural characteristics.]

- <Consistency>
  - Goopy
  - Dusty
- <Concentration>
  - Filled
  - Dense
  - *Sparse*
  - Empty
- <*Newness*>
  - New
  - Old
- <*Delicacy*>
  - Delicate
  - Subtle
  - Coarse
- <*Clarity*>
  - Blurry
  - Distinct
- <*Dryness*>
  - Dry
  - *Wet*
- <*Hardness*>
  - *Hard*
  - *Soft*
- <*Sharpness*>
  - Sharp
  - Flared
  - Tapered
  - Spikey

    - Blunt
  - <*Smoothness*>
    - Smooth
    - Wrinkled
    - Rough
  - <*Edge-condition*>
    - <*Extension*>
      - Straight
      - Crooked
    - <Deviation>
      - <Protrusion>
        - Bump
      - <Indentation>
        - Hole [SN: Indents completely through.]

- *<Judgment> [SN: Inherent conditions, not transitory.]*
  - <*Goodness*>
    - Good
    - *Bad*
  - <*Neatness*>
    - Neat
    - Messy
  - <*Perfection*>
    - Perfect
    - Imperfect
  - <*Correctness*>
    - *Correct*
    - Incorrect
  - <*Loveliness*>
    - Beautiful
    - Ugly
  - <*Gaiety*>
    - Festive
    - *Everyday*
    - *Serious*
  - <*Strength*>
    - Strong
    - *Weak*
  - <*Normalcy*>
    - <*Normal*>
    - <*Abnormal*>
      - Strange
      - Crazy
      - Wild

- <*Changes-in-condition*> [SN: Due to process.]
  - <*Unchanged*>
  - <Folded>
  - <Peeled>
  - <*Damaged*>
    - Broken
    - Frayed

Mashed
Ravaged
<Destroyed>
<Decayed>
Rotted
Rusted
<Burnt>

<Action>

<Movement-reflecting-growth>
Emerging
Developing
Growing
Encroaching
Radiating
<Random-movement>
Meandering
Floating
Blowing
<Violent-movement—non-fluid>
Shooting
Exploding
Flying
Crashing
<Fluid-movement>
Sweeping
Flowing
Pouring
Spraying

<Art-and-craft-process>

<Creative-process>
<Manmade>
Handmade
Digital
<Natural>
<Art-style> [SN: Images in this context are inherently two-dimensional though they may appear or contain parts that appear three-dimensional.]
<Primitive>
<Classic>
<Modern-art>
Abstract-art
Pop-art
Psychedelic
<Science-fiction>
<Cartoon>
<Children's-art [SN: Transliterated from kid-making-pictures and like-a-kid-drawing.]
<Production-media> [SN: Refers to media applied to a surface (image-foundation).]
Ink
Pencil
Charcoal

    - Pastel
    - Paint
    - Watercolor
  - <*Technique*>
    - <Design> [SN: A scheme of line or shape.]
    - <Pattern> [SN: A repeated design, contributes to the visual illusion of tactile texture.]
      - Dotted
      - Striped
      - Crisscross
      - Grid
      - Checkered
      - Honeycomb
      - Web
      - Openwork
      - Mottled
    - <Cut>
      - Cutout
      - Matte [SN: Use of a material to mask or frame part of an image.]
    - <Carved>
    - <*Drawing-technique*>
      - <Cursive-writing>
      - <Sketch>
        - Scribble
      - <Contour>
      - <Shadow>
        - Drop-shadow
        - *No-shadow*
    - <*Textile-method*>
      - Beading
      - Stitching
      - Hooking
      - Crochet
      - Weaving
  - <View>
    - <*By-position*>
      - Aerial-view
      - Cross-section
      - Magnification
    - <*By-object*>
      - Moonscape
      - Landscape [SN: Refers to natural scenery.]
      - Cityscape
    - <Perspective>
      - Vanishing-point
- <Color>
  - <Palette> [SN: Indicates the specific selection of colors used in an image.]
    - <Hue> [SN: Individual colors (hues) occurring in the raw data are grouped according to the eleven basic colors delineated by Berlin and Kay as typical of developed color vocabularies.]
      - Black

- Gray
- White
- Pink
- Red
- Purple
- Blue
- Green
- Yellow
- Orange
- Brown

<Multicolor> [SN: Indicates a selection of colors.]
- <Spectrum> [SN: Indicates a rainbow palette of seven colors.]
- *<Seasonal-color>*
  - Summer-color
  - Fall-color
  - *Winter-color*
  - Spring-color

<Color-quality>
- *<Purity>* [SN: Saturation.]
  - *Pure*
  - Blend
- <Tone> [SN: Refers to intensity, value, or chiaroscuro (treatment of light and dark in an image.)]
  - Pale
  - Dark
- *<Brightness>*
  - Bright
  - Glowing
  - *Dull*
- *<Transparency>*
  - Transparent
  - Opaque
- <Shaded> [SN: Refers to darkened or graded tones.]

SPATIAL-LOCATION

<Format> [SN: Directional or referential position of the image frame.]

- Landscape-format
- Portrait-format

<Position> [SN: Position of objects with reference to the frame and each other.]

- *<Indexical>* [SN: Refers to the position of the object relative to itself within the context of the image.]
  - *<Presentation>*
    - Front
    - Side
    - Back [SN: Refers to the rear side.]
  - <Orientation>
    - *<Rotated>*
      - Upright

            - Slanting
            - Sideways [SN: Over to the side, as in 90-degrees.]
            - Upside-down
        - <Reversed>
    - <*General-part*>
        - Top
        - Center
        - Bottom
    - <*Extremity*>
        - Start [SN: Refers to the beginning of a shape.]
        - End [SN: Refers to the end of a shape versus beginning.]
        - Edge
        - Surface
- <*Relational*> [SN: Perspectives between objects regardless of frame orientation.]
    - <Distance>
        - Far
        - Near
    - <Interconnection>
        - <Intersecting>
        - <Through>
        - <Cross> [SN: From one side to the other.]
        - <Joined>
            - Intertwined
            - Blocking
    - <Containment>
        - <Inside-of>
            - Concentric
        - <Holding>
        - <Surrounding>
        - <Outside-of>
    - <Placement>
        - <Off>
        - <On>
        - <Facing>
        - <Beside>
            - Parallel-to
        - <Between>
        - <*Horizontal-perspective*>
            - Behind [SN: Refers to a place to the right of an object.]
            - In-front-of [SN: Refers to a place to the left of an object.]
        - <*Vertical-perspective*>
            - Above
            - Below
    - <Layer>
        - Foreground
        - Middle-ground
        - Background
    - <Cover>
        - <*Whole-coverage*>
            - Superimposed [SN: Refers to one thing covering another.]
            - Underneath [SN: Refers to the location of a thing being covered.]

<Partial-coverage>

Overlap [SN: Refers to one thing partially covering another.]

Cap [SN: Refers to covering the top or end of a thing.]

<Direction> [SN: Directional location from viewer perspective oriented to the frame or an object within the frame.]

<Pointing-to>

<Vertical>

Up

Down

<Horizontal>

Left [SN: Analogous to the west position on a person's body when facing north.]

Right [SN: Analogous to the east position on a person's body when facing north.]

<Diagonal>

<Perpendicular-to>

<Compass-orientation> [SN: Bearing.]

East

Southeast

South

Southwest

West

Northwest

North

Northeast

<Clock-orientation> [SN: Refers to o'clock, and is combined with Number to indicate a clock hand location.]

Outline of Facet Hierarchy

For reference purposes, the facet outline presents a compressed version of the facets, without isolates. For readability, special formatting for facet names is not included and after three sub-levels facet names are grouped in bracket forms as follows, indicating successively subordinate status: [({ < > })] .

For example:

Animal-body-part [torso, appendage (arm {wrist, hand <fist, finger>}, leg)]

Facet outline

1. Object
    a. Image
        i. Kinds-of-image [drawing, painting, mosaic, collage, photograph, movie]
        ii. Image-foundation [paper, film, screen]
    b. Non-living-thing
        i. Figure
            1. Sign [alphabet, numeral, punctuation, symbol, text-string]
            2. Shape [two-dimensional-shape (dot, line, angle, curve, polygon, free-form), three-dimensional-shape (block, sphere, pyramid, prism, cone, cylinder)]
        ii. Artifact
            1. Gun
            2. Tool [fastening, sewing, constructing, blacksmithing, cutting, cooking, marking, measuring, viewing, fire-fighting]
            3. Electronics [device, electrical-component]
            4. Vehicle [aircraft, boat, car, train]
            5. Recreation [sports, game, puzzle, playground-equipment, toy]
            6. Container [dish (cup, pitcher, bottle, bowl, pan, plate), bag, gift]
            7. Food [sweet (doughnut, cupcake, pie, fruit-cocktail, jam, candy), beverage, cheese, bread, prepared-food]
            8. Textile [cloth, fiber]
            9. Clothing [item, part-of-clothing (edging, sleeve, zipper), accessory (badge, jewelry, neckwear, headdress)]
            10. Furniture [bed, table, shelf, lamp]
            11. Burial-artifact
            12. Money
            13. Publication [book, newspaper, article, map, postcard, postage-stamp]

iii. Mechanical-part [handle, rung, wheel]

iv. Substance

1. Pure-substance [solid, liquid]
2. Mixture

v. Naturally-occurring-phenomena

1. Outer-space [sun, moon, planet, meteor]
2. Atmosphere
3. Weather [wind, cloud, precipitation (rain, snow), storm]
4. Earth [land-based-substance (rock, soil, mineral, organic), water-event (wave, tide, flood), fire]
5. Electricity

c. Living-organism

i. Animal-life

1. Mammal [human, non-human (domesticated, non-domesticated)]
2. Bird
3. Reptile
4. Insect [kinds-of-insect, stages-of-insect]
5. Worm
6. Fish
7. Crustacean
8. Mollusk
9. Echinoderm
10. Cnidarian
11. Micro-organism
12. Mythical

ii. Plant

1. Plant-form
2. Fruit
3. Vegetable
4. Grain
5. Non-edible

iii. Body

1. Animal-body-part [head (face {eye, nose, jaw, chin, mouth, teeth, facial-expression}, ear), neck, torso, appendage (arm{elbow, wrist, hand , <fist, finger>}, leg {ankle, foot}, non-human-animal-appendage), covering (skin, membrane, hair {follicle, hair-style}, feather, scale, shell, wart), tissue (cell, flesh), blood (blood-vessel), skeleton(skull, bone), reproduction, organ]
2. Plant-body-part [root, trunk, stem, leaf, flower, seed, pod, rind]

iv. Aspects-of-living-thing [gender, maturity, role (religious-figure, leader, military, farmer, architect, artist, dancer]

2. Place
    a. Constructed-environment
        i. Water-based-environment
        ii. Land-based-environment
            1. Air-travel-environment [helipad, airport]
            2. Building [residential, cultural (church, educational-institution, museum, amphitheater, aquarium, planetarium, observatory, monument), political, business (industrial, business-structure, specific business]
            3. Construction-site
            4. Building-element [construction-material (wood, glass, brick, tile, cement, wire, plastic), room, building component (stair, floor, platform, pillar, wall, shelf, door, window, vault, roof, chimney, jail-bar)]
            5. Environment-component [walkway, fence]
            6. Traffic [driveway, road, highway (multilane, interstate {beltway}, traffic-junction, bridge, tunnel]
            7. Recreational-facility [park, golf-course, swimming-pool, track, stadium, plaza, courtyard, yard]
        iii. Locale [rural, urban]
    b. Natural-place
        i. Sky
        ii. Body-of-water [unenclosed-body (ocean, gulf, bay, harbor, inlet, river, stream), enclosed-body (lake, lagoon), within-water-body]
        iii. Shore
        iv. Land-water-formation
        v. Terrain [elevation (valley, hill, mountain, volcano), land-formation]
        vi. Eco-system
    c. Sociopolitical-location [continent, country, state, municipality]
    d. Generic-place [area, opening, joint, nowhere]
3. Property
    a. General-concept [existence, domain, validation]
    b. Attribute
        i. Visibility [visible, disappearing, invisible]
        ii. Gestalt
            1. Completeness [whole, part (division{section, piece}, fraction)]
            2. Closure
            3. Equilibrium [simplicity, regularity]
            4. Arrangement [ordered, grouped, random]
            5. Continuation [alignment (aligned, unaligned), frequency]

6. Similarity

iii. Quantity

1. Linguistic-quantity
2. Number
3. Dimension [specific-dimension, dimensionality]
4. Measurement-unit
5. Comparative-measure [size, width, length, height, depth]

iv. Comparison

1. Kind-of-change [increase, decrease]
2. Rate-of-change
3. Degree
4. Certitude
5. Probability
6. Equality [equal, unequal]
7. Emphasis

v. Condition

1. Consistency
2. Concentration
3. Newness
4. Delicacy
5. Clarity
6. Dryness
7. Hardness
8. Sharpness
9. Smoothness
10. Edge-condition [extension, deviation (protrusion, indentation)]

vi. Judgment

1. Goodness
2. Neatness
3. Perfection
4. Correctness
5. Loveliness
6. Gaiety
7. Strength
8. Normalcy [normal, abnormal]

vii. Changes in condition

1. Unchanged
2. Folded
3. Peeled
4. Damaged
5. Destroyed [decayed, burnt]

viii. Action [movement-reflecting-growth, random-movement, violent-movement—non-fluid, fluid-movement]

ix. Art-and-craft-process
  1. Creative source [manmade, natural]
  2. Art-style [primitive, classic, modern-art, science-fiction, cartoon, children's-art]
  3. Production-media
  4. Technique [design, pattern, cut, carve, drawing-technique (cursive-writing, sketch, contour, shadow), textile-method]
  5. View [by-position, by-object, perspective]

x. Color
  1. palette [hue, multicolor (spectrum, seasonal-color)]
  2. color-quality (purity, tone, brightness, transparency, shaded]

4. Spatial-location
   a. Format
   b. Position
      i. Indexical
         1. Presentation
         2. Orientation [rotated, reversed]
         3. General-part
         4. Extremity
      ii. Relational
         1. Distance
         2. Interconnection [intersecting, through, cross, joined]
         3. Containment [inside-of, holding, surrounding, outside-of]
         4. Placement [off, on, facing, beside, between, horizontal perspective, vertical perspective]
         5. Layer
         6. Cover [whole-coverage, partial-coverage]
   c. Direction
      i. Pointing-to
      ii. Vertical
      iii. Horizontal
      iv. Diagonal
      v. Perpendicular-to
   d. Compass-orientation
   e. Clock-orientation

Appendix F
Concept List

There are 465 concepts listed each with new pair value and frequency counts after synonym/near-synonym collapse, domain ratings (i.e., 0, 1, 2, 3), and ordered by position in the hierarchical structure of the faceted vocabulary. Concepts with pair values of 0 have been removed. Concept numbers 1 to 465 have been assigned for ease of reference.

				Shared-ness Ratings				
Concept number	465 Concepts	Pair Count	Freq. Total	0	1	2	3	Sum
1	<OBJECT>	8	20	13	4	3	1	13
2	<Drawing>	7	8	15	5	1	0	7
3	<Painting>	8	14	14	7	0	0	7
4	<Mosaic>	3	11	17	4	0	0	4
5	<Collage>	1	1	20	1	0	0	1
6	<Photograph>	19	55	2	14	5	0	24
7	<Movie>	2	2	19	2	0	0	2
8	<Paper>	2	2	18	3	0	0	3
9	<Film>	6	10	14	5	2	0	9
10	<Screen>	1	1	19	2	0	0	2
11	<Figure>	6	6	15	6	0	0	6
12	<Sign>	8	59	13	4	3	1	13
13	<Alphabet>	1	3	19	2	0	0	2
14	<Letter>	15	233	6	7	6	2	25
15	<Case>	1	1	20	1	0	0	1
16	<Numeral>	2	2	19	2	0	0	2
17	<Punctuation>	8	35	13	4	4	0	12
18	<Symbol>	19	110	2	6	12	1	33
19	<Text-string>	2	3	19	2	0	0	2
20	<Shape>	21	686	0	0	6	15	57
21	<Dot>	21	426	0	2	6	13	53
22	<Line>	21	1647	0	0	1	20	62
23	<Angle>	21	861	0	0	1	20	62
24	<Curve>	19	278	2	5	9	5	38
25	<Closed-curve>	21	776	1	1	10	9	48
26	<Spiral>	4	5	17	4	0	0	4
27	<Loop>	2	2	19	2	0	0	2
28	<Curl>	9	27	11	8	1	1	13
29	<Arc>	8	54	12	5	1	3	16
30	<Hook>	1	1	19	2	0	0	2

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
31	<Polygon>	1	1	20	1	0	0	1
32	<Triangle>	20	345	1	5	10	5	40
33	<Diamond>	3	17	19	2	0	0	2
34	<Rectangle>	21	1227	0	0	2	19	61
35	<Hexagon>	1	1	20	1	0	0	1
36	<Octagon>	3	21	19	2	0	0	2
37	<Free-form>	21	222	0	4	10	7	45
38	<Three-dimensional-shape>	1	1	20	1	0	0	1
39	<Block>	2	2	12	6	2	1	13
40	<Sphere>	7	55	14	3	4	0	11
41	<Pyramid>	12	92	9	12	0	0	12
42	<Prism>	3	33	17	3	1	0	5
43	<Cone>	1	1	20	1	0	0	1
44	<Cylinder>	9	39	13	7	1	0	9
45	<Gun>	3	7	18	3	0	0	3
46	<Fastening>	2	30	18	3	0	0	3
47	<Sewing>	2	5	19	2	0	0	2
48	<Constructing>	1	1	18	3	0	0	3
49	<Blacksmithing>	1	3	20	1	0	0	1
50	<Blade>	5	12	16	5	0	0	5
51	<Axe>	2	6	19	2	0	0	2
52	<Cooking>	1	7	20	1	0	0	1
53	<Marking>	6	38	14	6	1	0	8
54	<Measuring>	2	3	19	2	0	0	2
55	<Viewing>	16	73	5	12	4	0	20
56	<Fire-fighting>	1	2	20	1	0	0	1
57	<Device>	3	6	18	3	0	0	3
58	<Electrical-component>	3	10	0	19	1	1	24
59	<Aircraft>	9	21	13	7	1	0	9
60	<Boat>	2	12	19	2	0	0	2
61	<Car>	5	7	15	6	0	0	6
62	<Train>	6	12	15	6	0	0	6
63	<Sports-equipment>	6	25	13	8	0	0	8
64	<Type-of-sport>	7	26	15	5	1	0	7
65	<Game>	1	1	20	1	0	0	1
66	<Play>	1	1	20	1	0	0	1
67	<Board-game>	1	1	20	1	0	0	1

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
68	<Computer-game>	3	4	18	3	0	0	3
69	<Puzzle>	2	6	19	1	1	0	3
70	<Playground-equipment>	1	1	20	1	0	0	1
71	<Toy>	1	8	20	0	1	0	2
72	<Container>	1	2	20	0	1	0	2
73	<Cup>	2	2	19	2	0	0	2
74	<Pitcher>	1	9	20	1	0	0	1
75	<Bowl>	3	19	18	2	1	0	4
76	<Pan >	1	1	20	1	0	0	1
77	<Plate>	1	1	19	2	0	0	2
78	<Bag>	1	6	20	1	0	0	1
79	<Gift>	1	2	20	1	0	0	1
80	<Food>	1	1	20	1	0	0	1
81	<Doughnut>	1	1	20	1	0	0	1
82	<Cupcake>	1	3	20	1	0	0	1
83	<Pie>	1	4	20	1	0	0	1
84	<Fruit-cocktail>	1	1	20	1	0	0	1
85	<Jam>	1	4	20	1	0	0	1
86	<Candy>	2	3	19	2	0	0	2
87	<Beverage>	2	6	20	1	0	0	1
88	<Cheese>	1	1	20	1	0	0	1
89	<Bread>	1	3	20	1	0	0	1
90	<Prepared-food>	5	20	14	6	1	0	8
91	<Textile>	1	1	19	2	0	0	2
92	<Cloth>	7	20	13	7	0	1	10
93	<Fiber>	9	20	10	9	1	1	14
94	<Item-of-clothing>	9	33	12	9	0	0	9
95	<Edging>	4	29	17	3	1	0	5
96	<Sleeve>	1	2	20	1	0	0	1
97	<Zipper>	1	2	19	2	0	0	2
98	<Badge>	1	1	20	1	0	0	1
99	<Jewelry>	1	1	19	2	0	0	2
100	<Neckwear>	2	3	18	3	0	0	3
101	<Headdress>	4	11	17	4	0	0	4
102	<Bed>	1	1	20	1	0	0	1
103	<Table>	1	1	20	1	0	0	1
104	<Shelf>	1	1	20	1	0	0	1

				Shared-ness Ratings				
Concept number	465 Concepts	Pair Count	Freq. Total	0	1	2	3	Sum
105	<Lamp>	1	1	19	2	0	0	2
106	<Burial-artifact>	2	3	20	1	0	0	1
107	<Money>	2	8	19	2	0	0	2
108	<Book>	2	6	19	2	0	0	2
109	<Newspaper>	2	4	19	2	0	0	2
110	<Article>	1	1	20	1	0	0	1
111	<Map>	6	20	14	7	0	0	7
112	<Postcard>	1	4	20	0	1	0	2
113	<Postage-stamp>	1	1	20	1	0	0	1
114	<Handle>	4	7	17	3	1	0	5
115	<Rung>	1	1	20	1	0	0	1
116	<Wheel>	4	13	17	4	0	0	4
117	<Substance>	9	27	11	6	3	1	15
118	<Solid>	5	13	15	4	1	1	9
119	<Liquid>	15	120	5	9	7	0	23
120	<Mixture>	5	9	17	3	1	0	5
121	<Outer-space>	2	2	19	2	0	0	2
122	<Sun>	4	10	16	5	0	0	5
123	<Moon>	7	17	14	6	1	0	8
124	<Planet>	1	1	20	1	0	0	1
125	<Meteor>	1	5	20	1	0	0	1
126	<Atmosphere>	11	31	10	5	5	1	18
127	<Weather>	2	3	19	2	0	0	2
128	<Wind>	3	5	18	3	0	0	3
129	<Cloud>	6	18	15	5	1	0	7
130	<Rain>	6	11	16	5	0	0	5
131	<Snow>	3	11	18	3	0	0	3
132	<Ice>	3	6	18	3	0	0	3
133	<Storm>	1	3	20	1	0	0	1
134	<Earth>	3	5	17	4	0	0	4
135	<Rock>	11	31	9	11	0	1	14
136	<Soil>	6	19	14	6	1	0	8
137	<Mineral>	9	45	13	7	1	0	9
138	<Organic>	3	7	18	3	0	0	3
139	<Wave>	11	76	9	7	4	1	18
140	<Tide>	1	1	20	1	0	0	1
141	<Flood>	1	1	20	1	0	0	1

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
142	<Fire>	3	17	18	2	1	0	4
143	<Electricity>	2	10	19	2	0	0	2
144	<Living-organism>	2	4	19	2	0	0	2
145	<Animal-life>	2	2	19	2	0	0	2
146	<Human>	5	12	16	5	0	0	5
147	<Domesticated>	13	32	10	11	0	0	11
148	<Non-domesticated>	6	22	15	6	0	0	6
149	<Bird>	6	17	15	5	1	0	7
150	<Reptile>	10	31	10	10	1	0	12
151	<Insect>	3	6	20	1	0	0	1
152	<Kinds-of-insect>	9	35	12	9	0	0	9
153	<Stages-of-insect>	1	2	20	1	0	0	1
154	<Worm>	4	8	15	6	0	0	6
155	<Fish>	6	22	15	5	1	0	7
156	<Crustacean>	2	7	19	2	0	0	2
157	<Mollusk>	3	9	18	3	0	0	3
158	<Echinoderm>	1	2	20	1	0	0	1
159	<Cnidarian>	2	2	19	2	0	0	2
160	<Micro-organism>	7	24	13	7	1	0	9
161	<Mythical>	1	2	20	1	0	0	1
162	<Plant>	4	6	17	4	0	0	4
163	<Plant-form>	13	44	8	12	1	0	14
164	<Fruit>	6	19	14	6	1	0	8
165	<Vegetable>	9	45	12	6	3	0	12
166	<Grain>	2	6	20	1	0	0	1
167	<Non-edible>	5	9	16	4	1	0	6
168	<Body>	7	16	16	4	1	0	6
169	<Head>	17	80	4	14	2	1	21
170	<Face>	8	19	13	8	0	0	8
171	<Eye>	6	23	15	6	0	0	6
172	<Nose>	9	20	11	9	1	0	11
173	<Jaw>	1	3	20	1	0	0	1
174	<Chin>	1	2	20	1	0	0	1
175	<Mouth>	5	15	16	4	1	0	6
176	<Teeth>	3	10	18	3	0	0	3
177	<Facial-expression>	2	4	18	3	0	0	3
178	<Ear>	6	9	14	7	0	0	7

				Shared-ness Ratings				
Concept number	465 Concepts	Pair Count	Freq. Total	0	1	2	3	Sum
179	<Neck>	8	23	13	7	1	0	9
180	<Torso>	2	10	19	2	0	0	2
181	<Appendage>	1	1	20	1	0	0	1
182	<Arm>	9	55	12	7	1	1	12
183	<Elbow>	1	2	20	1	0	0	1
184	<Wrist>	1	2	20	1	0	0	1
185	<Hand>	6	15	15	5	1	0	7
186	<Fist>	1	1	20	1	0	0	1
187	<Finger>	9	37	13	4	3	1	13
188	<Leg>	11	27	10	10	1	0	12
189	<Ankle>	1	2	20	1	0	0	1
190	<Foot>	6	10	15	6	0	0	6
191	<Non-human-animal-appendage>	14	40	7	9	4	1	20
192	<Skin>	3	9	18	2	1	0	4
193	<Membrane>	2	4	19	2	0	0	2
194	<Hair>	6	28	15	4	1	1	9
195	<Follicle>	2	2	19	2	0	0	2
196	<Hair-style>	3	5	18	3	0	0	3
197	<Feather>	5	9	16	5	0	0	5
198	<Scale>	1	3	20	1	0	0	1
199	<Shell>	4	14	16	5	0	0	5
200	<Wart>	2	3	20	1	0	0	1
201	<Tissue>	5	6	20	1	0	0	1
202	<Cell>	16	147	6	14	1	0	16
203	<Flesh>	2	5	18	2	1	0	4
204	<Blood>	5	9	16	5	0	0	5
205	<Blood-vessel>	8	28	13	6	2	0	10
206	<Skeleton>	2	2	19	2	0	0	2
207	<Skull>	2	2	19	2	0	0	2
208	<Bone>	4	13	17	4	0	0	4
209	<Reproduction>	4	8	16	4	1	0	6
210	<Organ>	5	13	15	6	0	0	6
211	<Root>	2	8	18	3	0	0	3
212	<Trunk>	2	2	19	2	0	0	2
213	<Stem>	3	3	19	2	0	0	2
214	<Leaf>	6	10	15	6	0	0	6

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
215	<Flower>	4	6	17	4	0	0	4
216	<Seed>	5	8	17	3	1	0	5
217	<Rind>	1	1	20	1	0	0	1
218	<Gender>	5	9	18	3	0	0	3
219	<Maturity>	2	3	19	1	1	0	3
220	<Religious-figure>	1	1	20	1	0	0	1
221	<Leader>	2	3	19	2	0	0	2
222	<Military>	1	2	20	1	0	0	1
223	<Farmer>	1	1	20	1	0	0	1
224	<Architect>	2	2	19	2	0	0	2
225	<Artist>	7	19	14	5	2	0	9
226	<Dancer>	1	3	20	1	0	0	1
			9697					
227	<PLACE>	12	27	9	9	2	1	16
228	<Water-based-environment>	5	22	16	5	0	0	5
229	<Helipad>	1	2	20	1	0	0	1
230	<Airport>	4	21	17	4	0	0	4
231	<Building>	19	86	2	16	3	0	22
232	<Residential-building>	18	59	2	14	5	0	24
233	<Church>	2	2	19	2	0	0	2
234	<High-school>	1	1	19	2	0	0	2
235	<University>	2	4	18	3	0	0	3
236	<Museum>	4	4	17	3	1	0	5
237	<Amphitheater>	1	4	20	1	0	0	1
238	<Aquarium>	1	2	20	1	0	0	1
239	<Planetarium>	1	1	20	1	0	0	1
240	<Observatory>	2	5	19	2	0	0	2
241	<Monument>	7	16	13	7	1	0	9
242	<Political-building>	4	7	16	4	1	0	6
243	<Industrial>	5	8	14	7	0	0	7
244	<Business-structure>	5	11	14	7	0	0	7
245	<Specific-business>	3	5	18	3	0	0	3
246	<Wood>	9	21	11	10	0	0	10
247	<Glass>	6	23	15	6	0	0	6
248	<Brick>	1	2	20	1	0	0	1
249	<Tile>	2	4	19	2	0	0	2
250	<Cement>	2	2	19	2	0	0	2

				Shared-ness Ratings				
Concept number	465 Concepts	Pair Count	Freq. Total	0	1	2	3	Sum
251	<Wire>	2	2	19	2	0	0	2
252	<Plastic>	2	2	19	2	0	0	2
253	<Room>	6	9	14	7	0	0	7
254	<Stair>	1	2	20	1	0	0	1
255	<Floor>	3	6	18	3	0	0	3
256	<Platform>	2	5	19	2	0	0	2
257	<Pillar>	2	2	20	1	0	0	1
258	<Wall>	7	43	14	4	1	2	12
259	<Door>	1	5	20	1	0	0	1
260	<Window>	3	12	18	3	0	0	3
261	<Vault>	8	25	12	8	1	0	10
262	<Roof>	6	14	14	7	0	0	7
263	<Chimney>	1	1	20	1	0	0	1
264	<Jail-bar>	1	1	20	1	0	0	1
265	<Walkway>	1	2	20	1	0	0	1
266	<Fence>	3	37	17	4	0	0	4
267	<Traffic>	1	1	20	1	0	0	1
268	<Driveway>	1	2	18	3	0	0	3
269	<Road>	21	359	0	17	4	0	25
270	<Highway>	16	58	5	16	0	0	16
271	<Multilane>	2	2	19	2	0	0	2
272	<Interstate>	2	5	18	3	0	0	3
273	<Beltway>	4	10	17	4	0	0	4
274	<Traffic-junction>	6	17	14	7	0	0	7
275	<Bridge>	9	23	12	9	0	0	9
276	<Tunnel>	1	1	20	1	0	0	1
277	<Recreational-facility>	1	1	20	1	0	0	1
278	<Park>	5	12	16	4	1	0	6
279	<Golf-course>	1	1	19	2	0	0	2
280	<Swimming-pool>	2	3	18	3	0	0	3
281	<Track>	2	2	19	1	1	0	3
282	<Stadium>	4	15	16	5	0	0	5
283	<Plaza>	1	1	20	1	0	0	1
284	<Courtyard>	1	1	20	1	0	0	1
285	<Yard>	1	1	20	1	0	0	1
286	<Rural>	5	7	15	5	1	0	7
287	<Urban>	19	162	2	15	4	0	23

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
288	<Sky>	8	16	14	6	0	1	9
289	<Ocean>	9	14	12	7	2	0	11
290	<Gulf>	2	2	19	2	0	0	2
291	<Bay>	5	27	15	6	0	0	6
292	<Harbor>	1	6	20	1	0	0	1
293	<Inlet>	1	1	20	1	0	0	1
294	<River>	9	77	12	6	3	0	12
295	<Stream>	2	5	19	2	0	0	2
296	<Lake>	7	23	14	7	0	0	7
297	<Lagoon>	1	1	20	1	0	0	1
298	<Within-water-body>	2	5	19	2	0	0	2
299	<Shore>	12	61	8	12	1	0	14
300	<Land-water-formation>	10	42	11	10	0	0	10
301	<Terrain>	3	4	17	4	0	0	4
302	<Land>	20	100	1	19	1	0	21
303	<Valley>	3	6	18	1	2	0	5
304	<Hill>	9	23	12	9	0	0	9
305	<Mountain>	14	47	8	10	2	1	17
306	<Volcano>	2	5	19	2	0	0	2
307	<Land-formation>	16	92	6	6	5	4	28
308	<Eco-system>	16	66	6	13	1	1	18
309	<Continent>	2	3	18	3	0	0	3
310	<Country>	8	17	12	7	2	0	11
311	<State>	3	22	18	2	1	0	4
312	<Municipality>	8	22	12	8	1	0	10
313	<Area>	20	149	1	11	4	5	34
314	<Opening>	18	139	2	4	8	7	41
315	<Joint>	4	21	17	3	1	0	5
316	<Nowhere>	1	1	20	1	0	0	1
			2188					
317	<PROPERTY>	3	4	18	2	1	0	4
318	<Existence>	19	79	2	7	8	4	35
319	<Domain>	13	61	8	7	4	2	21
320	<Validation>	21	701	0	1	2	18	59
321	<Visible>	14	28	6	9	5	1	22
322	<Invisible>	8	16	16	3	2	0	7
323	<Whole>	21	132	0	4	8	9	47

Concept number	465 Concepts	Pair Count	Freq. Total	Shared-ness Ratings 0	1	2	3	Sum
324	<Part>	21	354	0	1	8	12	53
325	<Division>	19	140	2	7	6	6	37
326	<Section>	19	294	2	3	7	9	44
327	<Piece>	21	315	0	0	8	13	55
328	<Fraction>	20	1053	1	1	2	17	56
329	<Closure>	11	20	10	8	3	0	14
330	<Equilibrium>	1	1	20	1	0	0	1
331	<Simplicity>	9	20	11	7	3	0	13
332	<Regularity>	18	152	3	2	5	11	45
333	<Ordered>	17	44	4	12	3	2	24
334	<Grouped>	11	33	9	8	3	1	17
335	<Random>	8	18	13	4	4	0	12
336	<Alignment>	16	65	5	9	4	3	26
337	<Frequency>	11	30	8	9	4	0	17
338	<Similarity>	21	3059	0	0	0	21	63
339	<Quantity>	9	22	11	6	4	0	14
340	<Linguistic-quantity>	21	2174	0	0	0	21	63
341	<Number>	21	1021	0	0	1	20	62
342	<Dimension>	5	21	17	2	1	1	7
343	<Specific-dimension>	9	24	11	6	4	0	14
344	<Dimensionality>	6	8	16	4	1	0	6
345	<Measurement-unit>	17	784	3	5	1	12	43
346	<Size>	21	1558	0	0	0	21	63
347	<Width>	21	306	0	0	7	14	56
348	<Length>	21	335	0	3	4	14	53
349	<Height>	21	259	0	6	4	11	47
350	<Depth>	6	12	14	6	1	0	8
351	<Comparison>	2	3	17	4	0	0	4
352	<Increase>	16	172	4	6	8	3	31
353	<Decrease>	1	2	20	1	0	0	1
354	<Rate-of-change>	8	23	12	5	4	0	13
355	<Degree>	21	1157	0	0	0	21	63
356	<Certitude>	21	1361	0	0	3	18	60
357	<Probability>	17	143	3	2	10	6	40
358	<Equal>	16	51	5	9	3	4	27
359	<Unequal>	21	336	0	2	5	14	54
360	<Emphasis>	20	146	1	8	4	8	40

				Shared-ness Ratings				
Concept number	465 Concepts	Pair Count	Freq. Total	0	1	2	3	Sum
361	<Consistency>	13	29	8	9	4	0	17
362	<Concentration>	20	235	1	3	4	13	50
363	<Newness>	4	14	16	5	0	0	5
364	<Delicacy>	4	12	17	3	1	0	5
365	<Clarity>	14	55	6	4	10	1	27
366	<Dryness>	4	4	18	2	1	0	4
367	<Hardness>	2	3	18	3	0	0	3
368	<Sharpness>	15	89	6	8	3	4	26
369	<Smoothness>	9	24	10	5	3	3	20
370	<Extension>	21	317	0	3	3	15	54
371	<Deviation>	5	15	15	5	1	0	7
372	<Protrusion>	16	75	5	6	7	3	29
373	<Indentation>	8	28	13	6	1	1	11
374	<Goodness>	7	11	15	4	2	0	8
375	<Neatness>	7	11	13	6	2	0	10
376	<Perfection>	12	28	8	10	2	1	17
377	<Correctness>	7	10	13	7	1	0	9
378	<Loveliness>	2	2	18	3	0	0	3
379	<Gaiety>	1	1	20	1	0	0	1
380	<Strength>	3	5	19	2	0	0	2
381	<Abnormal>	12	52	7	11	2	1	18
382	<Folded>	2	3	19	2	0	0	2
383	<Peeled>	1	1	20	1	0	0	1
384	<Damaged>	11	32	7	11	3	0	17
385	<Decayed>	2	2	19	2	0	0	2
386	<Burnt>	1	1	20	1	0	0	1
387	<Action>	6	9	15	5	0	1	8
388	<Movement-reflecting-growth>	11	61	10	4	5	2	20
389	<Random-movement>	9	27	11	7	3	0	13
390	<Violent-movement--non-fluid>	16	39	5	12	3	1	21
391	<Fluid-movement>	15	49	7	8	5	1	21
392	<Manmade>	4	13	17	3	1	0	5
393	<Natural>	5	6	15	6	0	0	6
394	<Art-Style>	12	31	7	8	4	2	22
395	<Primitive>	2	2	20	1	0	0	1
396	<Classic>	1	1	20	1	0	0	1
397	<Modern-art>	17	72	4	11	5	1	24

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
398	<Science-fiction>	1	1	20	1	0	0	1
399	<Cartoon>	4	5	17	4	0	0	4
400	<Children’s-art>	3	3	19	2	0	0	2
401	<Production-media>	12	70	9	9	2	1	16
402	<Design>	6	16	15	3	2	1	10
403	<Pattern>	17	124	4	6	9	2	30
404	<Cut>	19	85	2	9	7	3	32
405	<Carved>	1	2	20	1	0	0	1
406	<Drawing-technique>	1	7	20	1	0	0	1
407	<Sketch>	4	8	17	4	0	0	4
408	<Contour>	15	59	6	8	5	2	24
409	<Shadow>	3	4	18	3	0	0	3
410	<Textile-method>	5	23	16	5	0	0	5
411	<View>	17	66	3	11	5	2	27
412	<By-position>	20	107	1	12	7	1	29
413	<By-object>	6	15	16	4	1	0	6
414	<Perspective>	5	13	15	6	0	0	6
415	<Color>	20	140	1	5	11	4	39
416	<Palette>	3	3	18	2	1	0	4
417	<Hue>	21	706	0	1	9	11	52
418	<Multicolor>	5	6	15	6	0	0	6
419	<Spectrum>	1	1	20	1	0	0	1
420	<Seasonal-color>	2	2	18	3	0	0	3
421	<Purity>	4	8	17	3	1	0	5
422	<Tone>	21	222	0	3	9	9	48
423	<Brightness>	9	34	12	6	3	0	12
424	<Transparency>	3	4	18	3	0	0	3
425	<Shaded>	13	71	8	6	5	2	22
			19711					
426	<SPATIAL-LOCATION>	2	2	19	2	0	0	2
427	<Format>	3	7	18	3	0	0	3
428	<Position>	3	5	18	1	2	0	5
429	<Presentation>	21	980	0	0	1	20	62
430	<Orientation>	8	25	13	2	4	2	16
431	<Rotated>	19	283	2	2	3	14	50
432	<Reversed>	8	22	12	7	2	0	11
433	<General-part>	21	3167	0	0	1	20	62

				Shared-ness Ratings				
Concept number	**465 Concepts**	**Pair Count**	**Freq. Total**	**0**	**1**	**2**	**3**	**Sum**
434	<Extremity>	21	1345	0	0	1	20	62
435	<Distance>	19	232	2	0	5	14	52
436	<Interconnection>	2	4	20	1	0	0	1
437	<Intersecting>	11	41	10	8	2	1	15
438	<Through>	17	63	2	9	7	3	32
439	<Cross>	21	285	0	4	16	1	39
440	<Joined>	21	359	0	2	7	12	52
441	<Containment>	4	8	17	3	0	1	6
442	<Inside-of>	21	1773	0	0	0	21	63
443	<Holding>	1	1	20	1	0	0	1
444	<Surrounding>	21	218	0	3	8	10	49
445	<Outside-of>	21	355	0	0	7	14	56
446	<Placement>	2	2	20	1	0	0	1
447	<Off>	21	271	0	4	6	11	49
448	<On>	21	913	0	0	1	20	62
449	<Facing>	13	37	8	7	5	1	20
450	<Beside>	20	276	1	1	5	14	53
451	<Between>	20	151	2	2	8	9	45
452	<Horizontal-perspective>	16	46	4	11	6	0	23
453	<Vertical-perspective>	21	292	0	1	5	15	56
454	<Layer>	16	88	4	5	9	3	32
455	<Cover>	18	100	3	6	10	2	32
456	<Whole-coverage>	15	62	8	5	4	4	25
457	<Partial-coverage>	9	23	12	6	0	3	15
458	<Direction>	15	43	6	5	9	1	26
459	<Pointing-to>	16	54	4	9	6	2	27
460	<Vertical>	21	1823	0	0	0	21	63
461	<Horizontal>	21	3200	0	0	1	20	62
462	<Diagonal>	17	222	3	1	4	13	48
463	<Perpendicular-to>	6	13	14	6	1	0	8
464	<Compass-orientation>	9	62	13	6	2	0	10
465	<Clock-orientation>	4	28	18	3	0	0	3

Appendix G
Content-based Image Retrieval Resources

Product vendors:

Virage (and its VIR technology) is at http://www.virage.com/

QBIC at http://wwwqbic.almaden.ibm.com/

Excalibur is now Convera at http://www.convera.com/

AMORE, by NEC, at http://www.ccrl.com/amore/

COMPASS at http://compass.itc.it/

Research and development projects:

WebSEEk (Columbia University) at http://ei.cs.vt.edu/~mm/cache/WebSeek.htm or Http://persia.ee.columbia.edu:8008/

Blobworld (University California, Berkeley) at http://elib.cs.berkeley.edu/photos/blobworld/

COLLAGE (Guildhall Library, London) at http://collage.nhil.com/ (abandoned? 11/22/02)

KIWI at http://telesum.insa-lyon.fr/kiwi

MARS (Multimedia Analysis and Retrieval System Project at http://www-db.ics.uci.edu/pages/demos/index.shtml

ImageRover at http://www.cs.bu.edu/groups/ivc/ImageRover/

PicToSeek at http://somac.wins.uva.nl:5345/ret_user/

PicSOM at http://cis.hut.fi/picsom

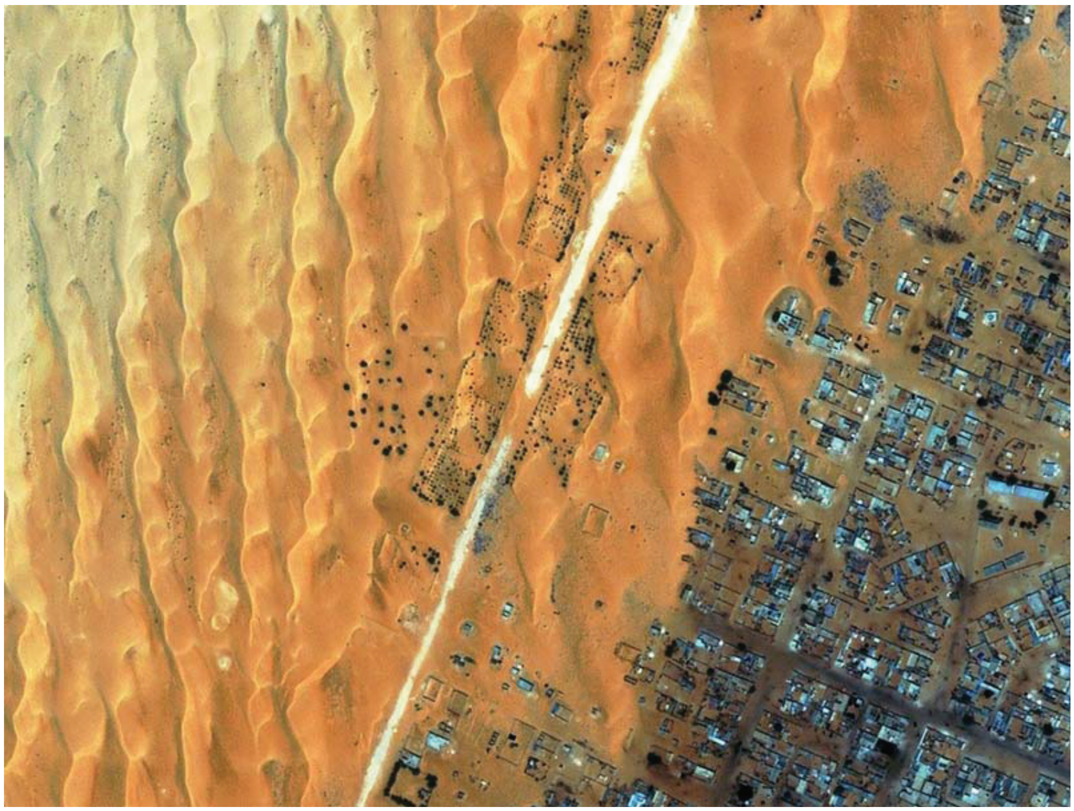
CBIR implementations

Hermitage Museum, using QBIC, at http://www.hermitagemuseum.org

LCPD: the Leiden 19th Century Portrait Database at http://nies.liacs.nl:1860/

Chabot at http://www.cs.berkeley.edu/~ginger/chabot.html (State of California Department of Water Resources)

Appendix H
Materials sample, actual size



Appendix I
Image Evaluation

Sorting instructions for image judges:

Rate the images on the following attributes using this scale:
(Note: the scale is “fuzzy,” not exact, rely on your first impressions)

1. minimal
2. average
3. dominant

RECOGNIZABLE OBJECTS
(Note: Do not be concerned with objects that just “look-like” something, only objects you recognize.)

1. minimal	no objects recognizable
2. average	one or more objects recognizable
3. dominant	image is recognizable, can be named as a place or thing, or creator known

COLOR

1. minimal	limited color
2. average	some colors
3. dominant	many colors

COMPLEXITY

1. minimal	few forms or patterns represented
2. average	more forms or patterns represented
3. dominant	many forms or patterns represented

EXEMPLARS

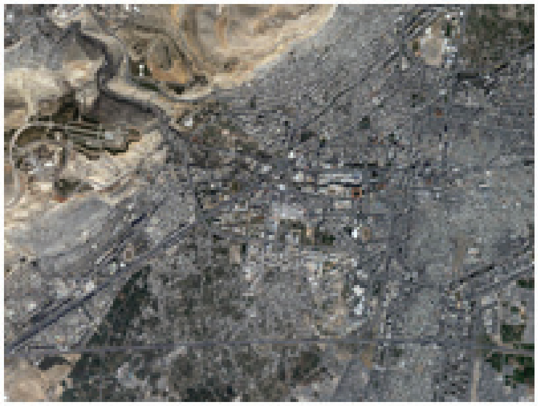
Select several images that are representative or serve as exemplars of each category:

SAT: satellite imagery
ART: abstract art
MIC: photo-microscopy

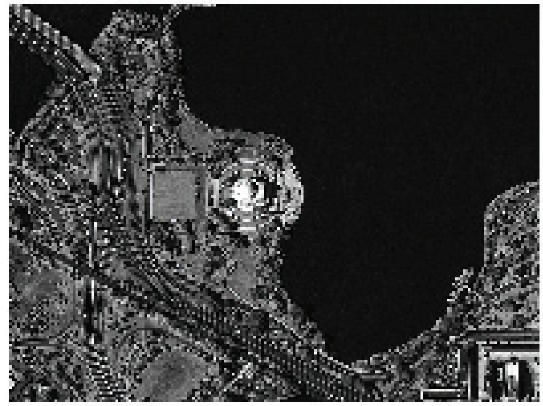
Appendix J
Sample Judge Data

Set/pic order	Accession#	Image Type: ABA, MIC, SAT	Image filename --S = Standard pic	exemplar	Judge CC color	Judge CC complx
13.1.5	246	MIC	diatomslarge.html	CC	2	1
16.1.5	267	MIC	anasatrismouthparts4#968.jpg	CC	2	1
17.2.7	206	MIC	snowflake1.jpg	CC	1	2
2.2.4	68	ABA	Halaby	CC	2	1
6.1.5	226	ABA	Lissitzky1	CC	2	3
7.2.4	118	SAT	dubai_hotel_1m.jpg	CC	2	3
8.2.3	163	SAT	sanfran800.jpg	CC	2	3
8.2.4	152	SAT	Nikumaroro_800.jpg	CC	2	2
12.1.6	101	ABA	Sommer3	CC	1	1
1.1.1	9	ABA	Balla [9A]--S	CC	2	3
1.2.3	133	SAT	JEFF800.GIF--S	CC	1	3
20.1.7	180	SAT	venice_800.jpg	CC	2	2
1.1.5	105	ABA	Leger2		2	3
1.1.6	208	MIC	taenia2-histopathology.jpg		2	1
1.2.4	99	ABA	Sommer1		1	3
1.2.5	12	SAT	CAIRO_C.TIF		2	3
1.2.6	270	MIC	DINO5.jpg		3	2
10.1.1	227	ABA	Lissitzky2		2	3
3.2.6	84	ABA	Morris		3	1
3.2.7	131	SAT	Invesco_Mile_high_800.jpg		2	3
4.1.2	87	ABA	Paik		2	3
4.1.5	138	SAT	l3_sanbay.jpg		2	3
4.1.6	266	MIC	oleanderleaf.arge.jpg		3	2
4.1.7	269	MIC	BUCKY1A.JPG		3	2
4.2.1	171	SAT	sprately_islands_800.jpg		1	1
4.2.3	264	MIC	youngstarfishlarge.jpg		3	2
4.2.6	32	SAT	1arcdem800.gif [32G]		1	2
11.1.2	169	SAT	shiraz800.gif		2	3
12.1.2	6	SAT	beijin_c800.gif		2	3
15.1.3	104	MIC	malaria-histopathology.jpg [13B]		2	3
16.2.4	245	MIC	pectinatellalarge.html		3	1
19.1.2	231	ABA	Rodchenko		2	3
19.2.7	113	SAT	DEN800.GIF		2	2
3.2.2	78	ABA	KrasnerSG		2	3
6.1.1	139	SAT	l3_sanbay_full.jpg		3	2
1.1.2	205	MIC	monomer.jpg--S		3	2
1.2.1	89	ABA	Poliakoff2.jpg--S		3	1
1.1.3	109	SAT	damascus_syria_800.jpg--S		1	3

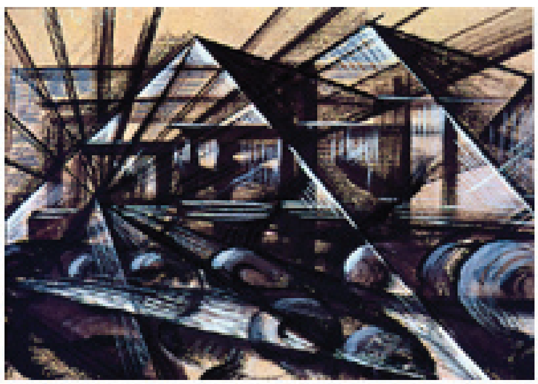
Appendix K
Standard images selected by judges



SAT: Damascas © Space Imaging, Inc.

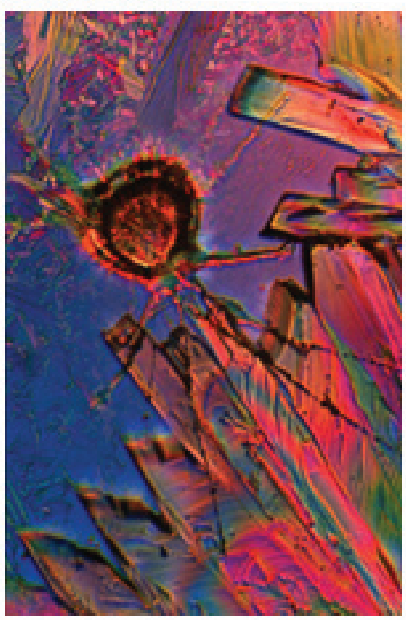


SAT: Jefferson Monument © Space Imaging, Inc.

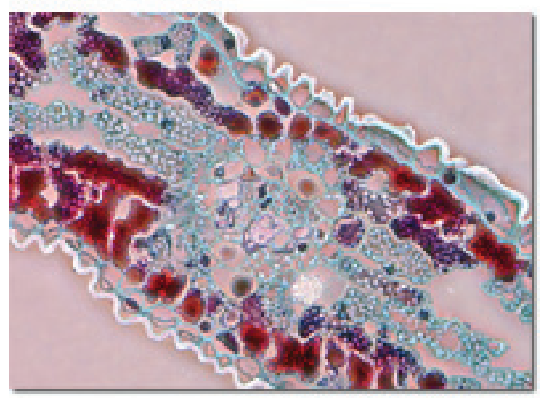


ABA: Balla

ABA: Poliakoff



MIC: Monomer
© Imaging Technology Group



MIC: Gingko Leaf

Appendix L

Subjects Wanted!

Volunteers Needed to Describe Pictures

Volunteers are needed for an experiment to collect the vocabulary that people use to describe pictures. This vocabulary is expected to contribute to the development of more effective image retrieval systems.

Working with a partner, one of you will describe a picture while the other creates a drawing from the description. No drawing skill is required. When each Describe-Draw task is completed, partners will compare the picture that was described with the drawing and discuss similarities and differences. The Describe-Draw task will be repeated for fourteen (14) pictures. Your participation will require no more than 1.5 hours for which you will be compensated $10.

Volunteers must be native English speakers and cannot have taken any college-level courses in art, architecture, photography, or geographic mapping.

For more information and participation scheduling please contact:

Caroline Beebe
Email: beebe@indiana.edu
812-335-0701

School of Library and Information Science
Main Library 011
Indiana University
Bloomington, IN 47405

Appendix M

Study # ___________

INDIANA UNIVERSITY – BLOOMINGTON
INFORMED CONSENT STATEMENT

Verbal Description of Pictures

You are invited to participate in a research study. The purpose of this study is to collect vocabulary that can be used to describe the composition of pictures. This vocabulary is expected to add to the ability to develop more effective image retrieval systems. The pictures you will be asked to describe will generally contain few if any recognizable objects.

INFORMATION

You will be given a study number and asked to complete a brief profile questionnaire. The profile questionnaire should take no more than five (5) minutes to complete. For this study, you will be paired with another subject. After completing one Describe-Draw task for practice, the Describe-Draw task will be repeated 14 times. Each task repetition should take no more than five (5) minutes. In each of the tasks one subject will provide a description; the other will create a drawing from this description. You will switch Describe-Draw roles after every third description until each of you has completed seven (7) descriptions. The total amount of time for your participation in this study should be no more than one and a half-hours (1.5).

If you withdraw from the study prior to completing all Describe-Draw tasks, your data will be destroyed and you will receive no compensation.

Every task will be taped for transcription purposes only. Only your study identification number will be noted on the recordings and transcriptions. While any reports written by research personnel may include quotations from your descriptions or samples of your drawings, you will not be identified by name. The tape recordings will be archived for five (5) years for potential re-analysis of the vocabulary or for future research on the descriptive process.

RISKS

There are no risks associated with participation in this study.

BENEFITS

By participating in this study you will help us collect a general vocabulary that can be used to describe the composition of images.

CONFIDENTIALITY

The information in the study records will be kept confidential. The researcher will retain one list that connects subject names with study identification numbers. This list will be destroyed once the study has been completed. Data will be stored securely and will be made available only to the person(s) conducting the study. No reference will be made in verbal or written reports that could link you to the study.

COMPENSATION

For participating in this study you will receive $10. You must complete all 14 Describe-Draw tasks to receive compensation. subject's initials:

CONTACT

If you have questions at any time about the study or the procedures used, you may contact the researcher, Caroline Beebe, at the School of Library and Information science, Main Library 011, Indiana University, Bloomington IN 47405, or by phone at (812) 855-2531.

If you feel you have not been treated according to the descriptions in this form or that your rights as a participant in this research have been violated during the course of this study, you may contact the office of the Human Subjects Committee, Bryan Hall 110, Indiana University, Bloomington, IN 47405, or by phone at 812/855-3067, or by e-mail at iub_hsc@indiana.edu.

PARTICIPATION

Your participation in this study is voluntary; you may refuse to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of any benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed, your data will be returned to you or destroyed.

CONSENT

I have read this form and received a copy of it. I have had all my questions answered to my satisfaction. I agree to take part in this study.

Subject's signature____ Date _________________

Investigator's signature___________________________________ Date _________________

Appendix N

PROFILE QUESTIONAIRE
Verbal Description of Pictures

ID No. _______________

1. Please indicate whether or not you have taken any college-level courses in art, architecture, photography, or geographic mapping by circling the appropriate response.

   I have not taken any courses   I have taken courses:
   List courses _____________

2. Please indicate if English is your first language by circling the appropriate response.

   Yes   No

3. Please indicate your sex by circling the appropriate response.

   Female   Male

4. Please indicate your age (to the closest year): ________________

5. Please indicate the highest level of education you have completed by circling the appropriate response.

   Some High School   College Graduate   Major:_______

   High School Graduate   Some Post Graduate

   Some College   Post Graduate Degree __________
   Please indicate degree, topic and year

6. Please indicate your home state and/or country: _______________________

Appendix O
SCRIPT

Describe-Draw task instructions

Thank you for agreeing to participate in this study. The study is intended to identify the vocabulary people use to describe the physical attributes of pictures. This vocabulary is expected to add to our ability to develop more effective image retrieval systems. Your ability to draw or describe images is NOT being tested. Rather, we are looking for the vocabulary that a general population might use to describe these pictures. To aid in analysis of your responses, I will be tape recording the entire session.

You will be working as a team to perform 14 Describe-Draw tasks. You will sit on either side of the partition and then I will give one of you a picture. You will describe the picture and your partner, using a pencil, will attempt to draw the picture from your description in **only enough detail so that the drawing could be used to find the picture in a collection of several hundred.** Pictures have been chosen based on their lack of readily identifiable objects or scenes. When you are describing the picture, you may not use any gestures. You cannot show the picture to your partner until you are done with the verbal description and your partner has completed the drawing. The one who is drawing may not ask any questions during the verbal describing task, **but may say "please repeat" or "please slow down" or "wait" (followed by "OK, ready").** Once you have seen a picture, I will not answer any questions. When the drawing has been completed, both of you will be asked to comment on the similarity of the drawing to the original picture and to describe any difficulties you experienced in describing or drawing the picture. Remember that drawing ability is **NOT** being evaluated. The only consideration is similarity of the drawing to the original picture, based upon the verbal description.

The first describer will be the partner whose last name falls first alphabetically: _______ will be subject 1 and _______ will be subject 2. You will take turns between describing and drawing. [Subject 1 name] will describe one picture for [Subject 2 name], then [Subject 2] will describe one picture for [Subject 1]. Then, [Subject 1] will describe two pictures for [Subject 2], and [Subject 2] will describe two pictures for [Subject 1]. After the first six descriptions (three each) have been completed you may take a 10 minute break after which you will begin the task again, with two descriptions by [Subject 1], and then two descriptions by [Subject 2], and repeat the sequence for two more pictures each.

When you describe a picture, you might begin with an overall orientation and then describe the details.

To complete the task, each of you must describe 7 pictures and draw 7 pictures. You may stop at any time. You will not receive full financial compensation of $10 unless you complete all 14 Describe-Draw tasks. If one partner stops the Describe-Draw task, participation by the other partner is also stopped. If the task is stopped before the completion of 14 Describe Draw tasks, you each will be compensated 50 cents for each drawing and each description that has been completed.

Are there any questions before we begin?

We'll do a trial picture first. [Give Subject 1 Picture 1: (a very simple example)] Try this picture.

Let's begin.

[When the Describer pauses in his/her description, ask:]
Is the description complete?
[When the Drawer pauses in his/her drawing, ask:]
Is the drawing complete?

Show your picture now.
What did you find easy or difficult about drawing from the description?
What did you find easy or difficult about describing the picture?
Is there anything either of you wants to add?
[Repeat for a total of 14 pictures, alternating Describe-Draw tasks between subjects as described above.
Ask if they want to take a break between Round 2-3 and again between Round 3-4. When they take a break tell them: "Please do not discuss the task during your break."]

Do you have any final comments?
Thank you for you participation.

Appendix P
Transcription sample

[P7.1#200] 127 words

Draw the same rectangle we've been using, about 1/4 inch from the bottom make a point on the edge on the left-hand side, then on the other side, the right-hand side, make a point about a 1/2 inch from the bottom, connect these points, it's a line but not completely straight, about an inch from that line in the center, draw an elliptical shape, about the length of your index finger, but, don't close it, it will look like a whale somewhat, the oval shape does, imagine a whale that's spouting water, draw that, then, on either side of the spout, all the way across, there are these shapes which kind of look like these little organisms, these oval shapes, so draw those all the way across the box.

[P5.2# 250] 463 words

This is one of those that's wider than it is tall so it's a little rectangular, the space we're going to be inside-of is rectangular, its wider than it is tall, there's a rectangular space that's centered in the center so put a little dot in the center, then its about, it's a rectangular space that's centered in the center and it is about 1/4, and this rectangular space is about 1/4 of the way out from the center towards the left-side edge and 1/4 of the way out from the center of the right-side edge so its equal-distance right to left and about 1/4 top to bottom, and this rectangular space is filled with big black globs and little tiny dotted black globs so there's a lot-of stuff in it, some of the big black globs are the size-of what would be raindrops on the paper, and on the left side coming from top to bottom is an irregular shape, its like a ribbon that's irregularly shaped at the top, the bottom goes out toward left-side edge but does not quite touch, its got 2 sides to it and its filled-in and goes all the way down top to bottom KO bowing-out towards the outside and the same thing is true on right side, we're not doing colors so its KO near the edge but not quite over there and it bows-out, they're not the same, they are irregularly shaped they have a left side and a right side KOL a strip or a ribbon runs down, goes all the way to the bottom those 2 little strips or ribbons irregularly shaped ones are shaded-in with color and then going from left to right there's another one of these irregularly shaped strips that goes toward the top from left to right across and KO bows-up towards the top a little bit, they're thin, they're not real fat and the same thing on the bottom, there's an irregularly shaped strip that runs from left to right and KO bows-out down toward the bottom and then surrounding the shape in the middle that had all the dots and drops-[rain] and stuff in it, surrounding that there are some little you could make them almost LL little wisps-of-flame that on the top edge up toward the top but they do not go past the strip that you drew up there so pretend its like a little fire on the top with a bunch of little strips and wisps and stuff then down the left side they KO go out way down the right side a little bit, they're fewer and on the bottom there's a few but not many so there just a little like if you were a little kid-making-pictures-of-a-campfire the fire-part-of the campfire

Caroline Beebe
1100 Brummetts Creek Road
Bloomington, Indiana 47408

EDUCATION

Master of Library Science, Indiana University, 1995
Education Specialist, Instructional Systems Technology, Indiana University, 1986
Master of Arts in Teaching, Elementary Education, Depauw University, 1976
Bachelor of Arts in Fine Art, concentration in Commercial Design, Depauw University, 1970

EMPLOYMENT

Indiana University Teaching Appointments
- 2003 Digital Libraries, School of Library and Information Science (SLIS)
- 2002 Digital Libraries, Organization and Representation of Knowledge and Information,
  School of Library and Information Science (SLIS)
- 1996 Computer Based Information Tools (SLIS)
- 1988 Computer Graphics, School of Fine Arts
- 1986 Computers in Interior Design, College of Arts and Sciences (COAS),
  Apparel Merchandising and Interior Design Department
- 1982-84 Associate Instructor, COAS, Mathematics Department

North Carolina State University, Raleigh, North Carolina
- 1998-2001, Head, Digital Library Initiatives Department, NCSU Libraries
- 1998-1999, co-Director, Learning Technology Service

University Computing Services, Indiana University, Bloomington, Indiana
- 1994-1996 Coordinator, Document Image Management Systems
- 1992-1994 Manager, Resource Center for Digital Presentation Technologies
- 1987-1992 Graphics and Multimedia Specialist, Workstation Systems
- 1984-1986 Graphics Consultant and Designer

Harmony Education Center, Bloomington, Indiana
- 1977-1982 Elementary School, founding team member, Math and Art curriculum developer

Selected PUBLICATIONS:

Beebe, C. (2000, Winter) “Integrating Local, Enterprise, and Global Visual Resources in the Academic Environment.” In Visual Resources Association: VRA Bulletin, Vol. 27, No. 4. pp 33-39.

Beebe, C. (2000, Fall), “Indexing Images for Multiple Needs” In Art Documentation, Bulletin of the Art Libraries Society of North America (ARLIS), Vol. 19, No. 2. Pp 16-21.

Beebe, C. and Jacob, E. (1998) “Graphic Language Documents: Structures and Functions.” In Structures and Relations in Knowledge Organization, Proceedings 5th International ISKO Conference, Lille, France, edited by W. Mustafa el Hadi, J. Maniez, and S. Pollitt. P 244-254 Wurzberg: Ergon Verlag.

Beebe, C. (1998). Joining the Club: Issues, Problems and Practices Initiating GIS. In Computer Applications and Quantitative Methods in Archaeology: New Techniques for Old Times, edited by J.A. Barcelo, I. Briz, A. Vila. p. 89-95. Oxford, England: Archaeology Press.

Beebe, C., Nedell, E., Song, M., Sullivan, J., Overfelt, K., Schatz, J. and Davenport, E. (1994, September) “New technology and publishing: six case studies in search of a theory.” Aslib Proceedings, 46:9, 217-224.

Beebe, C. (1993, November) “Diffusion and Adoption of the Digital Future: A case study in bringing multimedia to the campus community.” Proceedings of the ACM SIGUCCS User Services Conference XXI, 102-107.

Beebe, C. (1990, October). “Friendly But Streamlined: Developing an Integrated University Computing Environment.” Proceedings of the ACM SIGUCCS User Services Conference XVIII, 31-34.

Beebe, C. (1988, Spring) “Developing Screen Design Standards.” Proceedings of the Digital Equipment Computer Users Society USA, Cincinnati, OH, 77-84.

Beebe, C. & Spiaggia, C. (1985, April) “Philosophy and Capability Issues in Computer Graphics Software Evaluation: What Do You Mean I Can’t?” Proceedings of the 5th Symposium on Small Computers in the Arts, 29-38.