

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

Classifying Images: Criteria for Grouping as Revealed in a Sorting Task

Corinne Jorgensen
School of Information and Library Studies
381 Baldy Hall, Box 601020
University at Buffalo
Buffalo NY 14260-1020
CJORGENS@UBVMS.CC.BUFFALO.EDU

Introduction

In the last few decades, technologies such as video-recording and computer-based graphics packages have made images and tools for their creation widely accessible. Advances in technology have made digitization and transmission of images over long distances possible using both network and satellite technology, and these factors have produced a growing number of collections of images stored in digital format. However, providing means of access to the visual content of images has not kept pace with the advances in their production, storage, and transmission.

Access to the visual content of images is provided through indexing and classification systems. However, many images in collections are indexed only minimally, and even something seemingly as simple as assigning a "creator" for an image can be complex, as in a photograph or slide of a painting or an object. Access points often provide bibliographic information (perhaps the name of the photographer), and frequently do not capture a variety of information relating to the visual content of an image. Many authors concur that the major intellectual problem involved in access to images is the question of how to index them (Lynch, 1991) and there is, at present, a "dismal state of visual indexing in the United States" (Snow, 1990, p. 227).

Currently, human intuition and intelligence is used to develop methods for the analysis and indexing of images, and text-based systems are still primarily used for image retrieval. The museum and library communities have developed a number of systems for image indexing, each addressing particular aspects of an image pictorial matter or iconographical meaning), the needs of a particular collection (Christian art, academic library), or the type of user (art historian, architect), and some researchers have proposed theoretical bases for image indexing (Shatford, 1986). However, an indexing system can also limit the types of access available for images to that provided by the particular indexing system. Choosing an appropriate and useful system is often difficult, and little is actually known about how people or specific groups of users search for images or what attributes or features of an image are most likely to be searched upon.

Image indexing systems

Historically, classification systems for images arose from the need to establish an easily understood and predictable place for each slide, photograph, or other visual item within a manual filing system. The motivation behind many of these cataloging schemes was purely pragmatic, and there was little interest in a theoretical approach to classification; these classification systems are an eclectic mix of subject, iconography, time period, and various other access points.

Many of the more recent classification systems for images have had their foundations in the broader concerns of art information, and several of these draw from theories of aesthetics and interpretation (Couprie, 1983; Garnier, 1984). One concept which is expressed in many of the classification systems in one way or another is that of levels of iconographical meaning (Panofsky, 1955; 1962). Markey (1986, 3-5) explains that iconographical descriptors succinctly carry additional information about possible content. For instance, a subject heading of "Agony in the Garden" implies a praying Christ, an outdoor scene, and at least three sleeping disciples, whereas identification of lower-level primary subject matter would require that many entities be listed. However, these levels of interpretation require a high level of sophistication upon the part of both the indexer and the image searcher, and problems occur when a novice to these schemes tries to search an image within them. Iconographical classification systems require searchers to possess a specific and extensive knowledge of symbolic, literary, and cultural conventions, interpretations, and meanings, and while traditional iconographical systems may be useful for the art historical community, they are not useful for a more diverse community of searchers, because the levels of meaning addressed are frequently quite narrow.

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

Another trend in image indexing has been towards primary subject matter. The Library of Congress revised the list of controlled subject headings from the Prints and Photographic Division of the Library of Congress for use with automated cataloging and indexing systems, and in 1987 the *Library of Congress Thesaurus for Graphic Materials: Topical Terms for Subject Access* (LCTGM) was published. This work "provides a substantial body of terms for subject indexing of pictures, particularly large general collections of historical images" and "offers catalogers a controlled vocabulary for describing a broad range of subjects, including activities, objects, and types of people, events, and places depicted in still pictures" (Parker, 1987, p. v.). It does not include art historical and iconographical concepts but does supply terms for some abstract ideas. It intends to supplement another major image indexing effort, the Art & Architecture Thesaurus, (AAT) by providing terms for a broader range of people, events and activities which occur as picture elements along with specific terms for the built environment found within the AAT (p. viii.).

The AAT, the major thesaurus in the field of art and architecture, set out to build a comprehensive, and controlled vocabulary that could be used by database developers but that would not conflict with the working language of the users of those databases, the scholars and researchers. The broad goal is to provide a means of communication among the various constituents concerned with art object information and to provide a means of access for a variety of resources. However, Rorvig (1988, 51) observes that in the case of images, the use of thesauri to control inconsistency by guiding the choice of terms is not likely on its own to eliminate the individual responses prevalent in human reactions to visual materials.

Indexing of pictures is thus a complex task, as many aspects need to be dealt with to provide full access. Beyond the specific "objective" contents of an image, multiple levels of specific and/or generic class membership of individuals, animals, objects or events may need to be represented (i.e. Lawrence Olivier and/or "ACTORS"), as well as symbolic meanings or emotive aspects. More mundane aspects such as place or time also may also need to be represented both specifically or generically. Markey (1986, 51-56) notes that indexers asked to identify primary subject matter sometimes included other abstract elements as well.

Image searching

Current indexing and retrieval systems are created based on a number of assumptions about what users need and how users search for images. Indexing systems have been implemented based on pragmatic or theoretical concerns, yet, as of yet, there is little empirical evidence to support these approaches. The few published studies have primarily focused on users' queries to particular image collections or directed towards understanding users of a particular collection. Studies focusing on queries analyze and categorize the queries put to image collections.

Librarians at the National Library of Medicine (NLM) analyzed user groups and query logs and found that "image construct queries," in which images are constructed with words, compose one-third to one-half of users' queries (Keister, 1994). Queries reveal a number of "prototypical" searches, such as "neighborhood drug stores" and "pictures of cows." The NLM adopted "visual element cataloging," in addition to topical cataloging, to meet the needs of diverse groups with diverse types of queries.

Enser (1994) conducted a study of almost 3000 requests submitted by book publishers, advertising and design companies, magazine publishers, newspaper publishers, television/ audio-visual companies, and others to the Hulton Deutsch Collection. The majority of requests submitted (69%) fell into the "unique" category, defined as a specific instance of a general category (George III, as opposed to "kings"). Both unique and non-unique requests were subject to refinement by date (34%), location, action, event, or technical specification (the specification of image orientation or type of image). Requests were recorded by assistants and did not necessarily represent the language of the requester, thus it is difficult to evaluate the observation that users represented their requests at a greater level of specificity than users of online catalogs.

Enser concluded that general classification systems (such as the Dewey Decimal Classification Scheme) are extremely sparse indexing languages for visual information and suggests that a significant proportion of requests could be satisfied by automatic matching operations on picture captions. He concludes that subject indexing, of commercial photographic material, because of its expense, is of low utility and that reliance upon intermediaries will continue to be the norm.

Hibler et al. (1992), evaluated search failures in the Birkbeck system and found that they occurred mostly because of indexing omissions and problems in formulating search statements and assessing relevance. The two most frequent categories of index term omission were of what appeared to be a relatively minor detail ("necktie")

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

and failure to index frequently occurring objects ("walls," "roof"). Formulation of information need is related to the question of specific-generic levels, mentioned above. For instance, a search on the word "eating" failed to retrieve images of children looking at a birthday cake or standing in a lunch line, both judged to be relevant. The highly specific indexing language increased precision but decreased recall, so that searches conducted on general terms (hats, landscapes) were less successful.

Bakewell (1988), researching art historians, found that they also reported the various image access problems discussed above, such as minimal indexing (i. e. architecture by location) or difficult access to materials. Art historians desire access to "visual traditions" by means of a variety of indices.

In these user studies, two very general types of searches appear:

- 1) A "prototypical" search, a search for a representative instance of a person, place, item, or event; and
- 2) A specific item search, in which matching of the item is exact.

Given this multiplicity of approaches to both indexing and searching for visual materials, providing adequate indexing and retrieval mechanisms for images seems to be an insurmountable task with the time constraints bounding such activities. Authors such as Shatford and Turner (1993) suggest limiting the types of indexing done for images based on current knowledge of user needs within specific collections, but this implies a high level of knowledge about the user and the user's needs, a knowledge which most information retrieval researchers would agree is currently lacking. Shatford and others recognize that more user studies are needed to determine the most useful parameters for indexing. Until these studies are done, it seems unwise to limit in an a priori manner the retrieval needs of users and the search keys available to them.

Thus, the purpose of the research reported here was to conduct a study of human image interaction in order to address a fundamental and heretofore unanswered question:

•What attributes of images are typically described by humans in performing various tasks with images?

Several types of tasks were used to explore answers to this question; this information should inform future design of access systems for images, whether they are intellectual approaches such as indexing systems or machine-driven devices such as object parsing and recognition algorithms.

Research Methodology

The methodology for pursuing the above question was qualitative and exploratory, given the lack of knowledge about human interactions with images. Several themes emerge in discussions of qualitative research: that empirical reality is complex, intertwined, best understood as a contextual whole, and inseparable from the individuals who know that reality (Bradley and Sutton, 1993, p. 406). These themes also emerge from what is known about human perception of and interaction with images. For instance, the Gestalt principles act in as yet little understood ways to organize discrete graphic elements of images into larger wholes which are perceived as integrated units. Color perception is dependent upon the context in which the color is presented, as is the perception and understanding of objects themselves. While there have been attempts to specify the individual units or "morphemes" of images (Bertin, 1983), such attempts, while interesting, do not seem able to either capture the essential nature of visual materials nor to translate into useful precepts for image searching and retrieval.

This research accepts the assumption made in the literature on aesthetics that viewing an image is to some extent an interpretive act and that the interpretation of both images and texts is to some extent socially mediated (Solso, 1994). Rosch (1978) notes that formation of categories occurs in a particular culture at a particular point in time. Keil (1987, 1993) has proposed that individuals may be constrained by a communicative principle, in which the need to communicate an idiosyncratic representation of the world forces adoption of category structures which are more analytical, and therefore more stable, or consistent across individuals. It is thus the interaction between the individual perceptions of images and the socially constructed contexts for categorizing and communicating about these perceptions that is fertile ground for investigation and for seeking evidence concerning image attributes. Therefore, while the nature of the task imposes some constraints, there is also the social background which an individual brings to the viewing of the image and the need to communicate.

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

Definitions:

Image

For the purposes of this research, the definition of an image will follow its Latin origin, *imago*, as a representation of a person, scene, or object. Therefore the images with which this research is concerned are pictorial, are drawn from fine arts and graphic arts, and are realistic or semi-realistic representations produced by a number of methods. Abstract representations, which are subject to different content interpretations by different people, will not be dealt with.

Image attribute

An image feature for which verbal evidence is presented by the subject. An image attribute is not limited to purely perceptual characteristics, but includes other cognitive, affective, or interpretive responses to the image such as semantic, symbolic, or emotional characteristics.

Percept

The mental result or product of perceiving, as distinguished from the act of perceiving; an impression or sensation of something perceived (Random House Unabridged Dictionary, 2d ed., 1993). This term is contrasted to "concept," a purely mental idea with does not require an external physical stimulus.

Sorting Task

The task described in this paper is a sorting task. A method was needed that would allow image attributes and categories to be described by participants without the external constraints imposed by a predefined system. A sorting task was used to provide this environment. Sorting tasks have been used in a number of studies, primarily within the field of art education, to determine which image attributes are the determining factors in categorizing pictures. Other studies using both text (Canter and others, 1985), and images, (Rosch and others, 1976) have been conducted to explore the conceptual frameworks by which humans appear to categorize both objects and terms. Bornstein defines categorization as "treating a set of indiscriminable or discriminable stimuli as equivalent in some way (for example, by giving all of them the same name)," and notes that this is the many-to-one characteristic of categorization (Bornstein, 1987, 287). The sorting process is used to reveal the attributes by which a person distinguishes dissimilar items or groups similar items together, and the evidence provided by sorting experiments reveals the parameters by which people categorize or classify items. While searching for images and sorting are different tasks, they both involve distinguishing similarities and differences among the attributes of the images. Therefore the data provided by the sorting task is relevant to the exploration of image attributes.

Kwasnik (1989) demonstrated that "natural," individually-generated classifications are situated within a specific, person-oriented context, which, in addition to intrinsic features, often incorporates task or other non-object attributes. Most classification systems, on the other hand, adhere to abstract principles of "likeness" based on features intrinsic to the object and attempt to place items within categories apart from a person-centered viewpoint. Indexing and classification systems for images have been developed primarily from an external structural framework or need. Indeed, it has been considered nearly impossible to include other more contextual attributes in such classification schemes.

While there has been much research on how people categorize individual objects, there has been little research on how people classify pictorial images. There are several reasons for this: the complexity of the visual stimuli; the anticipated wide variations deriving from individual differences, expectations, and needs; and the challenges inherent in conducting such experimental research. While one could set up a number of hypotheses concerning the attributes of objects that people will take note of, deriving these from extant classification systems, these hypotheses would only emphasize the external nature of the classification systems and would not explore the internal nature of "person-centered" classifications.

Therefore, the purpose of the sorting task was to allow categorizations of images and their attributes to emerge apart from any predefined classificatory system. Some individual variability was expected in the data, as each participant would bring with them a person-oriented context within which they would be operating (e.g., what happened at breakfast this morning). Balancing this individual context is a broader social context within which classificatory behaviors are learned. The possibility that useful classifications deriving from this broader context may to some extent ameliorate the effects of individual variability needed to be explored.

Thus, the investigation of image attributes within a Sorting Task provides data to answer the following

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

question:

Which image attributes are most typically considered in the formation of self-generated categories for sorting images?

This question may be broken down into a number of sub-questions:

What are the terms that humans typically use to describe images when sorting images?

What image attributes are represented by these terms and what are the relationships among these attributes?

What types of categories do humans generate when asked to sort images for their own organizational and retrieval purposes?

What are the relationships between the spontaneously generated categories and the attributes associated with these categories?

Procedures

Eighteen participants for the Sorting Task were drawn from all levels of an academic setting in a variety of disciplines at Syracuse University, and the participants represent a wide variety of backgrounds, skills, and expertise, both in subject areas and life experiences. Previous research which has attempted to evaluate the effects of "visual orientation" (through training or experience) upon image description has shown no significant effects in terms of ability to describe images (Wilson, 1966; Markey, 1981; Turner, 1993). Therefore, in this study no attempt was made to measure "visual orientation" or to correlate this with the type of information generated.

The main resource for this study was a set of appropriate images for the tasks; previous work in this area (Markey, 1981; Turner, 1993) has used both art images or "ordinary" images. A variety of images were necessary, including both "artistic" and photo-realistic images, black and white and color images, and images depicting a wide variety of subject matter. Illustrations were chosen as the image group from which to select images; illustrations are designed with a wide variety of purposes in mind, yet at the same time they are intended to be a communicative device. Therefore, while not necessarily being realistic, illustrations must achieve a certain level of representational clarity, and this clarity was deemed useful in assisting the perception of a range of image attributes by participants. A total of seventy-seven images were randomly chosen by selecting every sixth picture from the illustrations in *The Society of Illustrators Twenty-Fifth Annual of American Illustration*. Each picture was mounted on black poster-board with a small black border showing around each image, and the order of presentation of the pictures was randomly varied among the subjects.

The participants were given the set of seventy-seven images and were told to sort them into groups for their own use, in a way that would enable them to find the images at a later time. The task was videotaped, and think-aloud protocols were used to gather data from participants. The combination of videotaping and verbal protocols ensured that discrete physical moves relating to individual images were recorded during the sorting task (e.g. Wright and Monk, 1991) and provided data as to which attributes are noted by participants in a categorizing task.

The products of this task, the transcripts of participants' commentary and the participant's sorted groups, were analyzed, with respect to both the categories created and the attributes associated with the categories. This provided evidence from which to determine those image attributes typically described in the formation of groups of images, and whether these attributes in fact form the basis upon which groups are built.

Data analysis

The attributes were determined directly from the data using the methods of content analysis. Conceptually similar terms were grouped together to become an *Attribute*. For example, the color terms "red," "green," and "blue" became an attribute coded as *Color*. When repetition suggested that no new attributes were likely to emerge, attributes were grouped into conceptually related classes. In the previous example, *Color* became a member of the class COLOR, which also included other general aspects of color (such as hue or tint) grouped as *Color Value*. The terminology used to describe the results of the research follows this hierarchical order: words or phrases from the raw data are referred to as "terms" and are placed in quotation marks ("red," "round"). A term is then assigned the appropriate content analytic code and becomes an *Attribute*; these are in italics (*Color*, *Shape*). The attributes are

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

then grouped into higher-level CLASSES (such as COLOR or VISUAL ELEMENT).

It should be noted that some attributes could still be further subdivided depending upon the specific needs or preferences of a particular group of users (an example would be the *Object* attribute, which includes, in addition to objects, living organisms such as plants and animals). The determination of specific attribute and class inclusiveness was guided by the content of the data. Thus terms representing objects such as clothing were assigned their own separate attribute code, *Clothing*, because of the frequency with which these terms appeared. The codes were not considered to be rigidly mutually exclusive. Some codes (such as *Level* and *Negative*) are consistently double-coded. A participant may also express *Uncertainty* or *Conjecture* about any of the above attributes. The definitions of the classes themselves also assisted in determination of which attribute code was appropriate for a particular term. In some cases a term may be coded differently depending on whether it is used as a noun or an adjective. Codes were also developed for personal reactions and behaviors in relation to the image and the task. Appendix A is a listing of the set of attribute codes and classes.

Results of attribute/class coding

The largest attribute class represented in this set of data was the class ART HISTORICAL INFORMATION, (23.8%) which includes the following interpretive attributes: *Artist, Medium, Representation, Style, Type, Technique, and Time Reference* (see Table 1). Since participants were given no written information about the images, the decisions as to what to "name" these various elements is considered more interpretive than perceptual, as decisions about *Artist* or *Style* could only be reached through a process of gathering visual clues and reasoning from these clues. This peak is in marked contrast to the series of describing tasks which were also part of this research (Jørgensen, 1995), in which LITERAL OBJECT was the most frequently-occurring class, accounting for between 25% and 35% of the classes.

Table 1 Attribute Classes - Sorting Task

Art Historical Information	23.8%
Viewer Response	14.1%
Abstract	13.8%
Literal Object	8.7%
External Relationships	8.7%
People	8.6%
Content /Story	8.5%
People Attributes	4.0%
Visual Elements	3.8%
Description	2.7%
Color	2.6%
Location	0.8%

The next two most frequently mentioned classes of attributes were VIEWER RESPONSE, which included personal reactions and accounted for 14.1% of the total, and ABSTRACT, which accounted for 13.8%. PEOPLE (8.6%) combined with PEOPLE ATTRIBUTES (4.0%) represents 12.6 % of the total. CONTENT/STORY (*Atmosphere, Activity, Category, Setting, Time Aspect, Theme, Event, Format*) represented about 8.5% of the total. OBJECTS and EXTERNAL RELATIONSHIP both measure 8.7%.

There are several interesting aspects to this set of data. The first is the much lower occurrence of the LITERAL OBJECT class when compared to the describing tasks, in which the LITERAL OBJECT class was the most frequently-occurring class (accounting for well over 25% in most cases and for over 30% in one of the describing tasks). The second is the frequent occurrence of ABSTRACT attributes, which describe such things as abstract concepts, themes, and symbolic interpretations. The third is the low frequency of the perceptual attribute classes of COLOR, LOCATION, DESCRIPTION, and VISUAL ELEMENTS. These categories each account for

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

less than five percent of the attribute class distribution. Therefore, for the Sorting Task participants appear to rely more heavily on what are termed "interpretive" attributes rather than perceptual attributes, at least in terms of what they choose to communicate about these images in their process of sorting and placing the images in groups.

A fourth interesting point is that an additional attribute which emerged as more typically noted in the Sorting Task was *Similarity*; in other tasks *Similarity* only accounted for between 0.5% and 1.5% of the attributes, whereas in the Sorting Task this one attribute accounted for 5.0% of the total attributes. This would seem to be only logical, in that in the process of grouping images participants must define some dimensions of similarity upon which they will base their groupings. When statements of similarity between two pictures are analyzed, the dimensions which appear most frequently as decisional as reflected by the attributes noted in association with the similarity term or phrase (i. e. "these two are alike" or "these are similar") mirror the distribution of the most frequently noted interpretive and perceptual attributes.

Image groups

The eighteen participants created a total of 292 different image groups, ranging from six to twenty-nine groups and with an average of about 16 groups per participant. The criteria for grouping images concerned both similarities and differences among the images and provided data on the attributes by which images are grouped or distinguished. Figure 1 contains a sample of group names by participant. Groups contained from one to twenty-seven images, with an overall average of five images per group.

Sorting "style" varied somewhat among participants. Most felt it was desirable to have mutually exclusive and non-overlapping groups, but two sorted more on a "continuum" with groups having "rough" edges. Those that used the first type of group found themselves "challenged," as one participant put it, by their groups as the sorting progressed, and participants would "loosen" their original definition of a group so as to make additional images "fit." Participants would also combine groups, apparently in an attempt to limit the complexity of the task. This behavior seems to fit rather well with theories on cognitive limits, in which humans seem best able to process from five to nine different "chunks" of information.

The behavior of the participants in the Sorting Task was more akin to a generalized browsing behavior, in which, as Buckland (1979) notes, informal searching takes place which is low on document specificity but may be on any level of information specificity. As participants began their sorting, a number of attributes were described and group names were generated representing concepts on a variety of levels of information specificity, ranging from, for instance, the very specific (*Cars*) to much more general concepts (*Landscapes*). Initially, some modifications were made to the group names in order to accommodate more images and generally keep the number of group names low (i.e. *Boats* and *Fishing* become a more general category such as *Water Recreation*). As sorting progressed, however, document specificity remained at a lower level of importance than the information specificity of the attribute terms used as group names, as participants compromised and put images (documents) in categories in which they did not fit "well," rather than changing the names or number of the groups once a "manageable" set had been determined.

For those participants who created a large number of groups, there appear to be in many cases several "supergroups" constituting of other groups. For instance, one participant created twenty-four groups, yet many of these groups could be subsumed under several larger group names which carry a common element among the groups, such as type of representation (*Landscape*). These larger groupings were often subdivided by such facets as style (*Realistic*) or medium (*Watercolor*). For instance, for one participant, four subgroups fall under the type *Landscape* which carries the common element "outdoor;" the subgroups are subdivided primarily on the basis of style (*Outdoors/Precise*; *Abstract Outdoor*; *Surreal/Abstract Outdoor*), with an additional element of included object: *Surreal Outdoor with Animals*. Similarly, when two *Animals* groups are analyzed, one finds that *Animal* is again subdivided by style into *Artsy* and *Cartoon*. When all the groups are analyzed in this way, there are an average of seven larger groups per participant; only one participant had more than nine groups, and in this case the tenth group was labeled the *Not sure* group.

Therefore it would appear that participants handled the cognitive complexity of the task of sorting seventy-seven pictures by creating conceptual "supergroups" which were then subdivided based on other facets. These supergroups, while not actually named, were implied by the physical arrangement of the groups on the work surface. For instance, all *Landscape* groups would be placed together in one area, but were named as separate groups by the addition of other specific facets. Thus there may be a "basic level" which applies to named groupings

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

of images, and this is an interesting idea for further investigation.

The group names given by participants to their sorted groups were coded with the same coding scheme as the individual attributes, and this proved to be a very adequate way of organizing and representing the participants' group names. There were a total of 442 attribute codes assigned to 292 group name terms for 2,337 images. Approximately one-third (97) of the groups were composed of multiple terms and thus received multiple codes. The analysis of group names thus took place at the same three levels of abstraction as the analysis and coding of attributes, moving from the actual terms which the participants used to name groups ("artsy animals," "nostalgia") to attribute code assigned to these terms (*Style, Object, Atmosphere*) to classes of attributes (ART HISTORICAL INFORMATION, LITERAL OBJECT, ABSTRACT CONCEPT).

There was wide variation in the names assigned to groups, and there were no group names which were used by all participants. Even when group terms are assigned attribute codes, there is wide variation (for instance, not all participants had group names including terms for people or animals, even though these were quite common visual elements). Because of this wide variation in terminology, coded attributes representing group terms are organized into related classes so that useful comparisons can be made. Participants' groups are thus described in terms of the higher level classes of attributes, with mention of the specific attributes which played a major role in each class. The distribution of attribute classes for the group names follows somewhat closely the distribution of the individual attribute classes over all images (see Table 2). Thus there appears to be a close relationship between the image attributes described by participants and the groups into which they placed the images.

Table 2 Attribute Classes of Group Names

Art Historical Information	35.8%
Abstract	17.0%
Content /Story	12.2%
People	11.5%
People Attributes	5.7%
Viewer Response	4.1%
Color	2.3%
Literal Object	1.9%
Visual Elements	1.1%
Description	0.5%

For instance, the Class which accounted for the majority of the terms used as group names was ART HISTORICAL INFORMATION (which included terms relating to type of image, such as "landscape," and to style of image, including terms relating to method of depiction, such as "realistic"). Thus the predominance of image attributes belonging to the class ART HISTORICAL INFORMATION in the participant's image descriptions appears to be reflected in the names which participants chose for their groups of images (e.g. "Cartoon Animals," "Fantasy Landscape", or "Portrait of Things").

The two interpretive classes ABSTRACT and CONTENT/STORY account for 29.2% of the terms in the group names. As in the overall class distribution, the class LITERAL OBJECT appears to be of much less importance, accounting for only approximately 2.0% of the total number of codes assigned to terms within the group names and 8.0% of the total attributes. PEOPLE and PEOPLE ATTRIBUTES terms maintain about the same level as in the other tasks.

A slightly different picture emerges when group names composed of multiple terms are analyzed. There were 97 multiply-coded group names and 110 pairs of attributes within these groups. When group names are analyzed in terms of the attribute pairings, certain attributes occur frequently and appear to be the "lead" attribute by which the grouping is made, with other attributes modifying the lead attribute. For example, the term "Animal" in a group name is classified as *Object*; it appears in a number of group names across participants with other attribute terms acting as modifiers: Animals With Human Attributes; Animal - Cynical; Animals - Cartoonish, Animals - Realism, and so on. Similarly, the term "Landscape" (*Type*) occurs frequently with modifiers belonging to several different classes: Empty Landscape, Fantasy Landscape, City Landscape.

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

Those individual attributes occurring most frequently in paired group names are *People*, *Object*, *Type*, *Style*, and *Abstract*. Among attribute classes within groups, the pairwise distribution of attributes occurs most frequently in the classes ART HISTORICAL INFORMATION (75 pairs), PEOPLE (37 pairs), ABSTRACT (36 pairs), and CONTENT/STORY and LITERAL OBJECT (both 26 pairs). Thus LITERAL OBJECT regains some importance in terms of group names composed of multiple terms. See List 1 for a listing of the major classes and attributes found in multiple word group names.

List 1

CLASSES and Attributes occurring more than once in multiple word group names.

ART HISTORICAL INFORMATION

Type (25); Style (20); Representation (16); Medium (7); Format (5); Time Reference (2)

ABSTRACT

Abstract (20); Theme; (10); Atmosphere (6)

PEOPLE

People (32); Social Status (5)

OBJECT

Object (26)

CONTENT/STORY

Activity (10); Setting (10); Category (6)

COLOR

Color (5)

It should be noted that the three major classes from which terms were drawn to name the image groups (accounting for 66.7%) are all "interpretive" rather than "perceptual" in nature. ART HISTORICAL INFORMATION is considered to be primarily interpretive in this case, as participants were given no written information about the images, and any terms which were applied to the images by the participant were based on interpretation of perceptual clues and application of a general level of art historical knowledge on the part of the participant. Thus, it would appear that in a sorting task where participants are asked to arrange images in groups which would enable their personal retrieval of the images in the future, participants draw more upon interpretive aspects of images, as revealed by the terms chosen to communicate about these images and the groups into which the images were placed, than upon more directly perceptual image attribute classes such as COLOR, VISUAL ELEMENTS (*Shape*, *Orientation*), LOCATION, or LITERAL OBJECT.

Although VIEWER RESPONSE accounted for about 14% of the attributes coded in this data set, it was only used as a group term in about 8% of the groups named. It would appear that perhaps the nature of the task prompted participants to respond to the images more emotionally or within personal constructs than in the describing tasks, yet it also appears that these responses did not heavily influence the naming of the groups. However, it should also be noted that when participants were asked if they wished to do a second sort, a small number proceeded to sort the images into two large groups, *Those I Like* and *Those I Don't Like*. During these second sorts, very little descriptive attribute information was given by participants, and these second sorts were thus not analyzed beyond the group level.

Some attributes used as group names are more heavily represented within the classes than others. In the class ART HISTORICAL INFORMATION, the most frequently occurring group name is *Type* ("landscape," "portrait"), followed by *Style* ("realistic," "abstract"); these two also lead in the numbers of pictures sorted into those groups. The second most frequently occurring group overall is *Object*, followed by *Abstract* in the ABSTRACT CONCEPTS class. The fifth (*Format*) is also in the ART HISTORICAL INFORMATION class.

Thus there appears to be a close relationship between the attributes named by participants in describing the images and the groups into which these images were placed, with the classes ART HISTORICAL INFORMATION, ABSTRACT, PEOPLE, CONTENT/STORY, and OBJECT occurring most frequently, and within those classes, the attributes *Type*, *Style*, *Abstract*, *People*, *Object* and *Similarity* being the ones most typically described. For the Sorting Task, participants described interpretive attributes of images much more heavily than perceptual attributes, and appeared to use the interpretive attributes more heavily as the basis for sorting images into groups which would

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

be useful to them at a later time in finding the images.

As mentioned above, this is in contrast to several other tasks which were performed as part of the larger research into image attributes. In three descriptive tasks, OBJECT was the most typically mentioned class of attributes, and other perceptual attribute classes such as COLOR were of much more importance. Thus, the type of task (descriptive or sorting) seems to have been a major determinant in the types of attributes most typically mentioned by participants, as both participants and images were similar across the several tasks.

Implications

A Sorting Task was used for this research based on the assumption that sorting images into groups involves a decision-making process which is representative of some aspects of "natural" human classification of images. Another assumption of this research is that evidence such as that provided by the Sorting Task may point to useful strategies for the indexing and classification of images. By studying how the participants grouped the images into categories to assist in their future personal retrieval, dimensions which are meaningful and useful in this type of task are revealed. The dimensions which appear to be most typically used to categorize this set of images closely mirror the image attributes which were described by participants. These dimensions were of a primarily interpretive nature and include *Style*, *Type*, and *Abstract*. Those perceptual dimensions which were most typically noted were *Object* and *People*.

One important question to be asked is to what extent these categories and the implicit classifications which they represent are stable. Recent research in cognitive science (in both lexical priming and retention) has demonstrated that in many instances category representations are unstable (Rey, 1983; Medin and Barsalou, 1987); other research demonstrates that mechanisms exist to establish stability of concepts (Smith, Medin, and Rips, 1984). Barsalou (1988) argues that both stability and flexibility are characteristic of concepts, and that concepts are dynamically constructed in working memory from category knowledge which involves three types of information: context-independent information, context-dependent information, and a third type which he characterizes as recent context-dependent information. Thus Barsalou argues that a concept is a representation in working memory only, and that rather than being definitional, "concepts simply provide an individual with useful expectations about a category based on long-term past experience, recent experience, and current context" (Barsalou & Medin, 1986). He notes that other models suggest that conceptual cores contain intuitive theories and idealized cognitive models, which stress the importance of category competence (Barsalou, 1988), whereas he suggests that at the core of any category is what he calls Context-independent (CI) information, a view which stresses category performance, focusing on those properties most frequently central to category use: properties which have been incorporated frequently; which are highly discriminative; which are functionally important; and which are highly accessible and low in context specificity. He proposes that CI information provides a certain amount of stability for concepts, and that because the same CI information is incorporated into all of a person's concepts for a particular category, some stability is maintained across contexts, and also because some CI information is shared by members of a population, some stability is maintained across individuals. He notes that averaging data across individuals or contexts may reveal sources which are constant across the concepts observed, but cautions that such data should not be construed as completely representing a person's concept in a particular context.

For the purposes of this research, Barsalou's theories provide support for both the approach used and the utility of the resulting data in suggesting attributes which may be useful for categorizing and retrieving images. This research suggests that in a Sorting Task attributes which are interpretive in nature such as *Style*, *Type*, and *Abstract* are most typically noted. Those perceptual dimensions which were most typically noted were *Object* and *People*. This research suggests that of a large number of possible attributes, these five operate as the most highly discriminative, functionally important, accessible and low in context specificity for this task. While individual term variability for these attributes was not assessed for the Sorting Task, four of these five attributes exhibit some measure of consistency (all except *Abstract*) in the describing tasks performed as part of the larger research. Thus, in addition to indexing of literal object and the human form, some interpretive aspects of images such as *Style* and *Type* may provide additional beneficial access points in retrieval.

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

The "story" of the image

Another very interesting aspect of this data was that in many cases, descriptions of image content included a "story" connected with the image: "she's getting drunk at a party;" "this guy looks ready for a fight". The structure of a phrase or sentence was frequently composed of an "actor" ("a fairy") and an "action" ("riding a large red animal"). Thus, participants would create a "story construct" with words.

Story elements were represented by a number of different attributes, some within the class CONTENT/STORY (*Activity, Setting, Category*), some within such other classes as ART HISTORICAL ATTRIBUTES (*Time Reference*), PEOPLE ATTRIBUTES (*Social Status, Emotion*), and ABSTRACT (*Atmosphere, Theme*). These eight attributes, in addition to *People* (actors) and *Object* (things acted upon) were the ones most frequently used to create story constructs. This suggests that perhaps a different class grouping of these types of elements would be more useful.

An area which could be drawn upon for insight into providing access to these more abstract, emotive, and interpretive elements of images is that of fiction indexing. Those wishing to find fiction face many of the same problems that image searchers face, and retrieval points such as author and title frequently reveal little content information. However, genre arrangements have proven useful for access to fiction (Baker, 1988; Pejtersen, 1989) and Small (1991) states that access to genre scenes such as banqueting would be useful. Beghtol suggests that fiction classification extract and classify individual instances of various kinds of data elements, and she notes the need to further evaluate such elements as "event" (Beghtol, 1991). Genre, as used for indexing fiction, corresponds to the attribute *Category* in the current research, which presents evidence that people organize their thoughts about images in a similar manner ("romance," "mystery," "science fiction").

The results of this research are being combined with evidence from three describing tasks and a Search Task to specify heuristics for the indexing of images. Much additional information may also be gained by analyzing in more detail the actual process of grouping by the participants, in addition to analysis of the products of this process, which were the focus of this paper. Analysis of the participants' behaviors should reveal key decision points clarifying the exact nature and role which various image attributes play in classifying images for personal retrieval. Further studies of this type are necessary to validate both the range of attributes and the typicality with which they appear in image descriptions, as well as the context within which they appear.

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

APPENDIX A

1. LITERAL OBJECT (perceptual)

Object; Text; Body Part; Clothing

2. PEOPLE (perceptual)

The presence of a human form (People) was remarked upon with very high consistency.

3. PEOPLE-RELATED ATTRIBUTES (interpretive)

Relationship; Social Status; Emotion

4. ART HISTORICAL INFORMATION (interpretive)

Artist; Format; Medium; Representation; Style; Technique; Time Reference; Type

5. COLOR (perceptual)

Color; Color Value

6. VISUAL ELEMENTS (perceptual)

Composition; Focal Point; Motion; Orientation, Perspective; Shape; Texture; Visual Component

7. LOCATION (perceptual)

Location - General; Location - Specific

8. DESCRIPTION (perceptual).

Description; Number

9. ABSTRACT CONCEPTS (interpretive).

Abstract; Atmosphere; Symbol; State; Theme

10. CONTENT/STORY (interpretive)

Activity, Category; Event; Setting; Time Aspect

11. EXTERNAL RELATIONSHIP (interpretive)

Comparison; Similarity; External Reference

12. PERSONAL REACTION (reactive)

Conjecture; Personal Reaction; Uncertainty

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

BIBLIOGRAPHY

- Jørgensen, Corinne. Image attributes: an investigation. SIGIR '95. Proceedings of the 18th Annual International Conference on Research and Development in Information Retrieval, ed. Edward A. Fox, Peter Ingwersen, and Raya Fidel. Seattle, WA, USA, July 9-July 13, 1995, p. 372.
- Bakewell, Elizabeth. 1988 *Object, image, inquiry : the art historian at work*, ed. Marilyn Schmitt. Santa Monica, Calif.: AHIP.
- Barsalou, L. W. 1987 *The instability of graded structure: Implications for the nature of concepts*. In *Concepts and conceptual development: Ecological and intellectual factors in categorization*, ed. U. Neisser. 101-140. Cambridge: Cambridge University Press.
- Beghtol, Clare. 1991 *The classification of fiction*. SIG/CR News (July): 3-4.
- Bertin, Jacques. 1983 *Semiology of graphics*, Madison, Wis.: University of Wisconsin Press.
- Besser, H. and M. Snow. 1990 *Access to diverse collections in university settings: The Berkeley dilemma*. In *Beyond the book: Extending MARC for subject access*, eds. Toni Petersen and Pat Molholt. 203-224. Boston: G.K. Hall.
- Bornstein, M. H. 1987 *Perceptual categories in vision and audition*. In *Categorical Perception*, ed. Stevan Harnad. 287-300. New York: Cambridge University Press
Bradley, Jana and Brett Sutton. 1993 *Reframing the paradigm debate*. *The Library Quarterly* 63 (4): 405-410.
- Buckland, M. K. 1979 *On types of search and the allocation of library resources*. *Journal of the American Society of Information Science* 30 :143-147.
- Canter, D., J. Brown, and L. Groat. 1985 *A multiple sorting procedure for studying conceptual systems*. In *The Research Interview: Uses and Approaches*, eds. M. Brenner, J. Brown, and D. Canter. New York: Academic Press.
- Coupric, Leendert. 1983 *Iconclass, an iconographical classification system*. *Art Libraries Journal* 8 (2): 32-49.
- Drabenstott, Karen Markey. 1986 *Subject Access to Visual Resources Collections : A Model for Computer Construction of Thematic Catalogs*, New York: Greenwood Press.
- Drabenstott, Karen Markey. 1986 *Subject Access to Visual Resources Collections : A Model for Computer Construction of Thematic Catalogs*, New York: Greenwood Press.
- Enser, P. G. B. 1993 *Query analysis in a visual information retrieval context*. *Journal of Document and Text Management* 1 (1): 25-52.
- Garnier, Françoise. 1984 *Thesaurus Iconographiques, Systems Descriptif des Representations*, Paris: le Leopold d'Or.
- Hibler, J. N. David, C. H. Leung, and K. L. Mannock. 1992 *Image Storage and Retrieval Systems*. In *Proceedings - SPIE, the International Society for Optical Engineering in San Jose, CA*, edited by Albert A. Jamberdino and Wayne Niblack, , 80-92.
- Keil, F. C. 1987 *Conceptual development and category structure*. In *Concepts and conceptual development: The ecological and intellectual factors in categorization*, ed. U. Neisser. 175-200. Cambridge, Eng.: Cambridge University Press.
- Keister, Lucinda H. 1994 *User types and queries: impact on image access systems*. In *Challenges in Indexing Electronic Text and Images*, eds. Raya Fidel, Trudi Bellardo Hahn, Edie M. Rasmussen, and Philip J. Smith. Medford NJ: Learned Information, Inc.
- Kwasnik, Barbara. 1989 *The Influence of Context on Classificatory Behavior*. Ph. D., Rutgers, The State University of New Jersey.
- Lynch, Clifford A. 1991 *The technologies of electronic imaging*. JASIS. *Journal of the American Society for*

PROCEEDINGS OF THE 6th ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP

Information Science 42 (8): 578-585.

- Medin, Douglas L. and Lawrence W. Barsalou. 1987 Categorization processes and categorical perception. In *Categorical Perception*, ed. Stevan Harnad. 1-25. New York: Cambridge University Press.
- Panofsky, Erwin. 1955 *Meaning in the Visual Arts: Papers in and on Art History*, Garden City NY: Doubleday Anchor Books.
- Panofsky, Erwin. 1962 *Studies in Iconology*, Reprint ed. ed., New York: Harper & Row.
- Parker, Elizabeth Betz. 1987 *LC Thesaurus for Graphic Materials: Topical Terms for Subject Access*, Translated by with an introduction by Jackie M. Dooley. Washington, D. C.: Library of Congress.
- Pejtersen, A. M. 1989 The book house: modelling user's need and search strategies as a basis for system design. *Risø M* :2794.
- Petersen, Toni, ed. 1990 *Art & Architecture Thesaurus*. New York: Oxford University Press.
- Rey, G. 1983 Concepts and stereotypes. *Cognition* 15 :237-262.
- Rosch, E., C. B. Mervis, W. Gray, D. Johnson, and P. Boyes-Braem. 1976 Basic objects in natural categories. *Cognitive Psychology* 8 :382-439.
- Rosch, Eleanor. 1978 Principles of Categorization. In *Cognition and Categorization*, eds. Eleanor Rosch and Barbara B. Lloyd. 27-47. Hillsdale NJ: Lawrence Erlbaum Associates.
- Shatford, Sara. 1986 Analyzing the subject of a picture: A theoretical approach. *Cataloging & Classification Quarterly* 6 (3): 39 - 62.
- Small, Jocelyn Penny. 1991 Retring images verbally: no more key words and other heresies. *Library Hi Tech* 9 (1): 51-67.
- Smith, Edward E., Douglas L. Medin, and Lance J. Rips. 1984 A psychological approach to concepts: Comments on Rey's "Concepts and stereotypes." *Cognition* 17 (3): 265-274.
- Solso, Robert L. 1994 *Cognition and the Visual Arts*, Cambridge, MA: The MIT Press.
- Turner, James M. 1993 Subject access to pictures: considerations in the surrogation and indexing of visual documents for storage and retrieval. *Visual Resources* 9 (3): 241-271.
- Weithas, Art, ed. 1984 *The Society of Illustrators Twenty-Fifth Annual of American Illustration* Illustrators 25. New York: Society of Illustrators.
- Wilson, Brent Guy. 1966 *The Development and Testing of an Instrument to Measure Aspective Perception of Paintings*. Ph.D., The Ohio State University.
- Wright, P. C. and A. F. Monk. 1991 The use of think-aloud evaluation methods in design. *SIGCHI Bulletin* 1 :55-57.