Marie R. Kennedy. Informing Content- and Concept-Based Image Indexing and Retrieval through a Study of Image Description. A Master's Paper for the M.S. in I.S. degree. February, 2005. 40 pages. Advisor: Jane Greenberg

This research reports on a Web survey of visual resource experts. The research uses the data supplied by the experts to discuss if content-and concept-based image indexing and retrieval approaches may be improved by examining features relevant to both. Image descriptions were gathered from image professionals via a Web survey. The descriptions were analyzed to test whether color images evoke a denser textual description than grayscale images, and whether there is any significant variation in the words used to describe them. Analyses of data resulting from the survey of image description indicate that image professionals describe images with high-level concepts rather than low-level features, and that color does not affect the number or type of words used in their descriptions. This work may prompt future research with other building blocks of the systems in order to better integrate the research done by computer and library scientists in this area.

Headings:

Pictures/Databases

Indexing/Pictures

Information Retrieval

Subject headings/Art

INFORMING CONTENT- AND CONCEPT-BASED IMAGE INDEXING AND RETRIEVAL THROUGH A STUDY OF IMAGE DESCRIPTION

by
Marie R. Kennedy

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

February 2005

Approved by

_____

Jane Greenberg

Introduction

The most powerful approach for image indexing and retrieval may be to integrate methods that automatically capture features with methods to manually assign high-level concepts. The first method allows searchers to locate images in digital libraries based on their physical features (color, shape, and texture) and the latter allows searchers to locate images by their meanings (concepts). Research following these two lines of practice is, however, routed in two different disciplines--the former method is primarily based in computer science with the latter in library science. As a result, advances with these approaches fail to sufficiently inform each other (Chu, 2001; Rasmussen, 1997).

As demonstrated in the literature, researchers have not arrived at consensus about how to index and retrieve the primitive features (color, shape and texture) that are the basis for content-based image retrieval (CBIR) systems. For differing approaches to color identification, for example, see Squire, Müller, and Müller (1999), Mehtre, Kankanhalli, and Lee (1998), and Stan and Sethi (2003). Concept-based approaches have also had limited success, with no overarching framework to guide indexers in the use of controlled vocabularies due to differing views about the very construction and use of those vocabularies and schemas (Gilchrest, 2003; Jörgensen, 1999). These fundamental flaws render both approaches weak.

This study addresses the features of both content- and concept-based retrieval approaches, examines how the two might overlap, and discusses the research design that examines the potential for integration.

To explore if the integration of these two approaches provides a better system, we must first investigate if humans do indeed respond to images in a more sophisticated way than with simple primitive feature identification. This research will focus on the primitive feature of color. The questions that guide this study:

- Do image professionals describe images with high-level concepts rather than low-level features?
- Do color images evoke a denser textual description than grayscale images?
- Is there variation in the terms used to describe either color or grayscale images? If so, is it significant?

It should be noted here that the term "grayscale" refers to images that do not contain color, but are instead measured on a scale of 256 shades of gray--as digitally constructed--with white and black at either end of the scale. The term "grayscale" will be used throughout the description and analysis of this research project. "Black & white," a more commonly used description of grayscale images in colloquial language, will be used only in the survey that is part of this research project. Though "black & white" is a less accurate term it will be used in the survey because the participants are likely to refer to grayscale images as black & white.

Features of Current Systems, as Reflected in the Literature

Content-Based Image Retrieval Systems

Content-based image retrieval (CBIR) systems automatically index and retrieve images based on one or more primitive features, such as color, shape, and texture (Eakins & Graham, 1999). This indexing approach is based on a design that extracts basic information from digital images that can then be mapped, quantified, and standardized.

The information that is automatically derived is termed "extrinsic" by Gudivada and Raghavan (1997, p. 438), or characteristics about an image attribute that can be externally obtained (e.g., name of image creator, date of image creation) in addition to primitive features.

CBIR systems are queried by example, with users providing a drawing or selection of colors/shapes, with the result being an image of similar color/shape being returned, ranked according to similarity (McDonald & Tait, 2003). For example, a searcher submits a query of a yellow circle and the results are pictures of the sun, a yellow ball, and unripe tomatoes. Retrieval by similarity of primitive features is successful when identification by a specific name or keyword is difficult to formulate (Eakins & Graham, 1999). Successful retrieval results are also gained by images that are best described by primitive features, such as line drawings, symbols and other easily identifiable image structures. A strength of CBIR systems is that they can be domain-independent; the indexing attributes of color, shape, and texture are not inherently tied to a particular image domain, such as art history, fine art, or commercial imagery. In addition to domain independence, the automatic indexing provided by these systems is efficient because it is completed within the system; there is no human interaction to slow the process. Notable CBIR systems are IBM's Query by Image Content (QBIC), Virage, and VisualSEEk (Rasmussen, 1997).

Color in content-based image retrieval systems.

A major weakness of existing CBIR design is that researchers disagree about their building blocks--the primitive features--and how they should be described in a system. The color spaces employed across system designs remain non-standardized, resulting in a

variety of approaches across the systems. Stan and Sethi (2003), for example, choose to model color hierarchically, based on hue, saturation and value (HSV), whereas Mehtre et al. (1998) focus instead on centroid clusters of the values for red, green and blue (RGB). Both approaches model color but from fundamentally different points of view and color spaces, one using a similarity metric, another a distance measurement. In addition to this example of varying approaches to color, Tieu and Viola suggest that color may not even be an effective component on which to index, especially in large image warehouses, because "different objects may share similar colors, and objects within a class may be colored differently" (2004, p. 34).

In addition to the construction of the system, a user's expectations of the system may impact failure/success outcomes. A user's concept of color, shape, or texture may differ from its representation in the system, providing computationally correct results that a user may find irrelevant for his purposes. In a survey of image database query formulation, Eakins and Graham (1999) note that some of the searchers were not able to construct queries in text that matched their expectation of the image response. This high-level expectation, met with low-level results is termed a *semantic gap* (Town & Sinclair, 2004).

Despite the current limitations of CBIR systems, there is encouraging work by researchers to establish a benchmark for color appearance for color management systems, which is a step toward an international color standard (Commission Internationale de l'Eclairage, 2004; Barnard, Martin, Funt, and Coath, 2002; Trant, 2004). If there could be agreement on a "best" use of a standardized approach to color, CBIR systems might succeed in establishing a common "language" for this aspect of image indexing and

retrieval. As is discussed in Smith's work, however, it is evident that agreement on a "best" approach has not yet been established (2001).

Conceptual Approaches for Image Indexing and Retrieval

Conceptual approaches to image indexing and retrieval are based on the assumption that searchers prefer to search for images with terms related to ideas/concepts/meanings rather than with the lower visual feature such as color, shape, and texture. This conceptual approach, characterized as "high semantic," is generally expressed through the use of domain-specific terminology (Smith, 2001; Sánchez, Chamorro-Martinez, and Vila, 2003).

This manual approach indexes based on keywords and other descriptors that categorize images into classes, according to a pre-determined schema, such as the *Library of Congress Thesaurus for Graphic Materials*.[1] Other notable image classification schemas and thesauri are the *Art and Architecture Thesaurus* and ICONCLASS.[2] Manual indexing, with the use of a classification scheme or thesaurus, is costly due to the fact that an indexer must spend time with each image and classification tool in order to describe its content. As noted by Eakins and Graham, average indexing times per image range from 7 to 40 minutes (1999). The average cataloging times per monograph at the University of North Carolina ranges from 11 to 33 minutes (1998). These average times demonstrate that regardless of medium, significant human intervention is required to catalog/index an item properly.

Aside from describing the content of an image, manual indexing can provide a contextual vocabulary, which may be useful for searchers. For example, providing words that describe what the image is "about" may be as useful as describing what the image is

"of" (Shatford, 1986). Empirical research by Jörgensen confirms that people describe images based on classes of images such as abstract concepts and story, in addition to content-related information (2003, p. 213). This suggests that images should be indexed according to both their content and context.

There are some inherent weaknesses with a conceptual approach to indexing images. The obvious issue is that choosing keywords--even from a controlled vocabulary such as listed above--is subjective, and an indexer may respond differently to an image over time (Markey, 1984). Images have historically been indexed only once by one cataloger, however, and the indexer's choice of keyword assignment will determine the future retrieval possibilities of that image. In addition, a single image may serve a variety of purposes (art, visual resource, archive), though not one descriptive schema exists to address all possible purposes (Greenberg, 2001; Shatford, 1986). Regardless of ambiguity in how to apply the given indexing tools, Shatford argues that "subjective judgments and analysis can provide valuable access to information," and should not be dismissed as an insurmountable issue in image classification (1986, p. 57).

Conceptual approaches to indexing color.

A limitation with indexing from a thesaurus or controlled vocabulary is that the thesauri and schemas may not be constructed with relevant terminology. The *Library of Congress Thesaurus of Graphic Materials,* for example, does not widely address the aspect of color, though empirical studies show that image searchers expect to use terms from this class to formulate queries and descriptions (Jörgensen, 1999). Indexers themselves may even reject the tools that impose difficult or archaic terminology, preferring instead to construct in-house vocabularies (Gilchrest, 2003).

The most notable classification tools, the *Library of Congress Thesaurus of Graphic Materials* (TGM), the *Art & Architecture Thesaurus* (AAT), and ICONCLASS offer varying approaches to color. Since the TGM does not widely address the aspect of color, an image cataloger using the tool as a guide for indexing may simply drop any notable color information from the index term assignment even if its color could be thought of as the "subject" of the image. The AAT, however, offers significant guidance in assigning color terms, from preferred descriptions to variants.[3] Jörgensen notes, however, that only a few of the terms found in the color hierarchy of the AAT can be assigned without user training (1999). ICONCLASS places *color* in the hierarchy of *nature--natural phenomena--light--colour, pigments, and paints*. These varying approaches to description of color leave the image indexer without an overarching focus for how to apply terms that will be useful for resource discovery.

Possible Interaction of the Existing Systems: Evidence for an Integrated Approach to Image Indexing and Retrieval

A key issue in the development of improved indexing and retrieval tools will be communication between researchers in the fields of library science and computer science. Examination of the work being done in the area of image retrieval, as reflected in the literature, finds that the computer scientists that focus on automatic indexing systems do not appear to publish with the library scientists that focus on manual indexing systems (Persson, 2000; Chu, 2001). The result of this gap is that the separate lines of research continue without influence from each other instead of drawing on each other's strengths.

Addressing the gap problem between content- and concept-based system designs is particularly important for the field of information science, the focus of which is to

provide patrons/searchers with the information they seek in an efficient manner. If this gap is not addressed in cross-discipline research, image users will continue to pursue ineffective searches for materials, thus limiting their own research results. Without cross-discipline research in the area of image retrieval users will continue to be presented with only low-level perceptual search options, when it has already been agreed upon that they prefer to search at higher levels (Smith, 2001).

This research project suggests that employing the strengths from both content- and concept-based systems can create an integrated approach to image retrieval, effectively bridging the gap between the two lines of research. One way to advance an integrated approach is to identify a feature used by both systems and then analyze how it used and how it could be improved. This research identifies the problematic primitive feature of *color*, which has already been acknowledged as a building block of content-based systems, and examines it using a concept-based approach. The questions guiding this study are:

- Do image professionals describe images with high-level concepts rather than low-level features?
- Do color images evoke a denser textual description than grayscale images?
- Is there variation in the terms used to describe either color or grayscale images? If so, is it significant?

Methodology

This study examines the primitive feature of color using a conceptual approach. A Web survey was created in which image professionals were asked to spontaneously "describe" four images. One may expect that the survey participants describe these four

images with high-level terms rather than with low-level terms such as related to color, shape and texture. One may also expect that color images will evoke a denser textural description than grayscale images. An evaluation of the survey data can provide evidence of whether or not image professionals use high-level terms to describe images. The results of this research will also comment on whether or not the respondents do, in fact, describe color or grayscale images with a significant difference in number of words. In addition, the frequency of repeated words will be examined in order to discuss whether or not the repeated terms are related to color or grayscale concepts.

A Web survey was constructed using the professional version of survey tool, SurveyMonkey (http://www.surveymonkey.com), in order to gather descriptive statements about color or grayscale images from participants. The participants are from groups of image experts--people who use images in their work on a daily basis. In particular, the groups surveyed are from the professional groups, Society of Photographic Education (SPE), the Visual Resources Association (VRA), the Art Libraries of North America (ARLIS), and a local student group, Art & Museum Libraries and Information Student Society (AMLISS). The request for survey participants was sent through the groups' email lists. Email list members that identify themselves as image professionals were invited to participate. The total number of email list members at the time of the call for participation are as follows: ARLIS—1587; VRA—635; AMLISS—25. The request for participation from the SPE was posted on the "opportunities" page of their website, due to restrictions about what kind of information is disseminated through their email list. The SPE "opportunities" web page was accessed 2296 times during the open survey period.

The Web Survey Design

The Web survey is constructed of an instruction page, four images to examine, four image description forms to complete, brief demographic information to supply, and a summary page (see Appendix A for survey instrument). The two portions of the survey contain the same content and format. The images used in the two portions are the same, with the color information dropped for the grayscale version. The URL of the survey was distributed to the professional groups noted above via email list to several groups of image experts that use images on a daily basis. At random, participants engage in either the color portion or the grayscale portion of the survey. The image description forms were submitted via Web form and retained on SurveyMonkey's secure server.

The images for the survey are taken from the *Tending the Commons: Folklife and Landscape in Southern West Virginia* collection of the American Memory Historical Collections for the National Digital Library (Library of Congress, 2000). The images chosen from this collection exist in the public domain. There are no copyright concerns for their use in an academic setting. See Figure 1 for images used in the survey.

Figure 1: Images used in this Web survey

Image 1

Image 2

Image 3

Image 4

The survey was created of the following pages, and provides an overview of its construction:

- Introduction. To begin the survey the participants were asked to think of the day on which they were born, to decide if this date was odd or even, and to select a corresponding radio button. This question directs half of the participants to the color portion of the survey (with an "odd" response) and half to the grayscale portion (with an "even" response).
- Images. After viewing the first image the participant was directed to a blank form.
- Forms. The participant was asked to simply "Describe the image you just viewed." The text box was set with scroll bars both horizontally and vertically so

that the viewer could respond with whatever type of response he wished, without constraint by length of line or total response. When finished typing the image description, the participant then pressed the "done" button and moves on to the next three images and descriptions.

- Demographic information. After repeating the image-description process four times the participant was asked to identify his sex; whether he works primarily with color or grayscale image, or both equally; the number of years he has been working with images; and the frequency of his interaction with images.
- Completion. The participant was then directed to a page that thanked him for participating, described the purpose of the research, and gave an email address if he desired follow-up information.

Data Analysis

Given two sets of images that are the same except for the lack of color in one set, the descriptions can be assumed to be similar if the presence of color has no effect in how image professionals describe an image. The similarity may be evident in the number of words given and the frequency of terms used. It was hypothesized, however, that the color images would evoke a denser textual description than the grayscale images. Analysis will be done to determine if there is a significant difference in the number of words image professionals use to describe color or grayscale images. Term frequencies will be analyzed to gauge if frequently used words in the image descriptions are related to color or the grayscale. In addition, analysis of this data is expected to confirm previous research that suggests that people prefer to interact with images on a high level rather

than at a low level (Smith, 2001; Sánchez, et al., 2003). The analysis reports on participation statistics, word count, term frequency, and a brief word analysis.

Response Rate & Attrition

To be considered complete for this analysis, one response must have descriptions for all four images and appropriate responses to the four demographic questions. Of 226 participants responding to the survey, 154 were useful for data analysis; 82 color responses, 72 grayscale responses.

Participant Statistics

After viewing and describing four images, participants were requested to provide brief demographic information by answering four questions. The first question was, "Do you work primarily with color or black & white images? (Select both responses if you work with both in equal amounts)." 74 (48%) participants use primarily color images in their work, 17 (11%) use black & white images, and 63 (41%) use both color and black & white. One participant did not respond to this question. The second question was, "How frequent is your interaction with images? (Select one response)." 129 (84%) of the participants marked their interaction as "very frequent (I use them daily)", 25 (16%) as "frequent (I use them about once a week)", 10 as "not too frequent (I use them about once a month)." One participant noted his interaction with images as "never (I do not interact with images)." One participant did not respond to this question. The third question was, "How long have you worked with images? (Select one response)." One participant has been in an image professional position for less than a year, 37 (24%) for 1-5 years, and 116 (75%) for more than five years. One participant did not respond to this question.

The last question was, "What is your sex?" 31 (20%) of the participants were male and 123 (80%) female. One participant did not respond to this question.

Based on the responses to the demographic questions, the participant that marked that he never interacted with images was eliminated from analysis, as the purpose of the survey is to gather descriptions from those that use images in their work. Also eliminated were those that marked their interaction with images as "not too frequent," as the study was designed to seek descriptions from image professionals which interact with images frequently. To ensure the homogeneity of the group of participants, the descriptions from the participant that chose not to answer the demographic questions were eliminated from analysis. Based on the qualifications identified here, of the 226 responses, 154 were useful for analysis; 82 color responses, 72 grayscale responses.

Word Count, Term Frequency

Word count.

Responses from the Web survey were gathered on the secure server of SurveyMonkey. The data was exported via Microsoft Excel. The number of words used in each description was counted, using Microsoft Word's *Word Count* function.

The number of words used to describe color images was compared with those describing the grayscale images. A model of the word count for the images was attempted in SAS using a Poisson regression but the dispersion was too great. The word count was then modeled using PROC GENMOD in SAS to perform a negative binomial regression with standard errors adjusted for multiple observations within subjects. Standard errors were adjusted to acknowledge that the observations within an individual survey participant are not independent of each other. The predictor in the model was

whether the image was color or grayscale. There were 154 clusters (i.e., survey participants) all with four observations, for a total of 616 observations. Table 1 shows that the difference between the parameters (i.e., the number of words used in the color and grayscale image descriptions) is not significant, as demonstrated by the *p*-value of 0.44.

Table 1: Analysis of GEE (Generalized Estimating Equations) Parameter Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	*p*-value
Color or grayscale	-0.1262	0.1660	-0.4515	0.1991	-0.76	0.4469

Term frequency.

Each image description was corrected for spelling, using the Microsoft Word *Spelling* function. No corrections were made for grammar. The terms *b/w* and *b&w* were expanded to *black and white* so that they could be counted as whole terms and not be discarded as nonsensical words when analyzed. The data was then migrated to Atlas.ti content analysis software to compare term frequencies for each color image and its counterpart grayscale image. Eliminating the terms *a, an, and, are, in, is, of, the,* and *this*, Table 2 displays the five most frequently used words in the image descriptions, as well as how many times the words were used.

Table 2 : Five most frequently used terms in the Web survey image descriptions

Rank of Term Frequency	First		Second		Third		Fourth		Fifth	
	TERM	*times used*								
Color image 1	TOMATOES	82	MAN	78	PICKING	55	GARDEN	53	RED	46
Gray image 1	MAN	72	GARDEN	43	TOMATOES	40	WHITE	37	PICKING	36
Color image 2	WOMAN	71	SHELVES	63	ON	57	WITH	52	CANNED	39
Gray image 2	WOMAN	72	WITH	52	ON	46	SHELVES	46	JARS	46
Color image 3	ON	138	WOMAN	84	WITH	80	COUCH	69	HER	58
Gray image 3	ON	127	WOMAN	80	COUCH	59	WITH	55	SOFA	49
Color image 4	RAMPS	75	THREE	56	SHELTER	54	ON	54	FOR	46
Gray image 4	RAMPS	57	TO	42	THREE	41	SELLING	40	FRESH	35

NOTE: Color images, n = 82; Grayscale images, n = 72

One can see from the most frequently repeated terms that the color and grayscale image descriptions are similar, with only slight variation in rank of the highest term frequency. Image 3, for example, has *on* as its most frequently repeated term in both the color and grayscale image (repeated more than once per image description, as noted in the number of times reported, 138 and 127, respectively). This is followed by the second most frequently repeated term, *woman*. The third and fourth terms for image 3 are closely aligned, as well. This demonstrates that the survey participants describe these images similarly, regardless of color/lack of color.

Word Analysis

Analysis of the words provided as descriptions of the images may suggest whether or not image professionals describe images using low- or high-level terminology. A random sample of 25 color and 25 grayscale descriptions was coded in the content analysis program Atlas.ti, with two codes that I created, “low-level” and “high-level.” The “low-level” code refers strictly to identification of primitive features: color/color value identification, shape specification, or texture description. The “high-level” code is

defined for this analysis as any word in the description that refers to an object or concept that is not color, shape specification, or texture description. Words having little semantic value such as *a*, *an*, *and* and *the* were ignored during coding. The numbers of each code were counted and are provided in Table 3 to demonstrate that image professionals do, in fact, describe images using high-level terms. See Figure 2 for a coding example of an image description, using the two codes defined here.

Table 3: Number of words coded as "high-level" and "low-level" in a sample of image descriptions

	Grayscale image (n = 25)	Color image (n = 25)
Number of words coded as "high-level"	537	579
Number of words coded as "low-level"	19 (11 color, 7 shape, 1 texture)	38 (27 color, 6 shape, 5 texture)

Figure 2: Coding example using the codes "high-level" and "low-level"

Description fragments	**Assigned code**
Old	High-level
man	High-level
stooped	High-level
over	High-level
picking	High-level
tomatoes	High-level
wearing	High-level
a	(ignored)
red	Low-level
hat	High-level
in	High-level
a	(ignored)
garden	High-level
with	High-level
weeds	High-level

None of the image descriptions contained words that exclusively represented primitive features. In fact, many of the descriptions did not contain any primitive feature terms at all.

Discussion of Findings

This research addresses three issues related to content-based and concept-based approaches to image indexing and retrieval: 1) do image professionals describe images with high-level terms rather than low-level terms; 2) do color images evoke a denser textual description than grayscale images; and 3) is there a significant variation in the terms used to describe either color or grayscale images.

This research confirms that image professionals overwhelmingly use high-level terms in their image descriptions, rather than low-level terms. Though they did also use some low-level terminology, this was only a small percentage of the overall description. In the random sample that was extracted from the full set of data in this research, color/color value terms account for only 4.4% of the total number of words used in the color image descriptions.

Analysis of both the number of words used for the image descriptions and the variation in terms used in the descriptions finds no significant difference between color and grayscale images, though the assumption was that there would be a difference. This assumption was due to the dramatic results of a brief research project done two years ago, in which a convenient sample of nine library science students described color images with more than twice the number of words than grayscale images. From this result I assumed that a larger sample would simply provide more robust results while again proving that color images evoke a denser textual description. Since the results of the

analysis from this expanded survey do not provide similar results, I turn to some possible explanations.

Naïve user versus expert user. Does a person's expert status change the way in which he responds to an image? The groups selected to participate in this survey work with images on a daily basis and are quite familiar with how to talk about/describe images. As image professionals, they have trained themselves through repetition to provide succinct but meaningful descriptions of images as they index them in their university or museum online catalogs. This training may provide consistency in image description (via number of words and word choice), regardless of the content or concept of the image.

In sharp contrast to this group of image professionals are the people that were initially surveyed two years ago. Library science students come from varying backgrounds and experiences, without necessarily any formal training in image description. It is possible that some of the students surveyed had completed a class in cataloging, or that some of them had previous cataloging experience, but this information was not requested as part of the survey. It is therefore reasonable to assume that there was no homogeneity in the initial survey sample, and their responses to the color or grayscale images were the result of chance. Indeed, the analysis of the initial survey data (via T-test) found that only one of the color images was described with a significant difference to its corresponding grayscale image.

Training of image professionals. Have image professionals been trained to ignore color? If, through repetition, image indexers have trained themselves not to respond to

color because it is not generally indexed using the available tools, it is possible that they responded to the images as their training and tools dictate.

Digital versus analog image viewing. Does viewing an image on a computer monitor evoke a different image description than if it were viewed in analog form? Image professionals historically work with traditional media: slides, photographs, and print material. Perhaps the format of this survey, using the Web, affected how the image professionals described the images. If the Web is an unfamiliar format for image description for this survey group, it is possible that it prompted a different type of description than if the survey had been completed with analog materials.

Image Description Versus Other Task-Related Activities

Cognitive research suggests that how a person responds to a task is dependent on what kind of cognitive load is required to perform it. Pettersson's psychological tests on how individuals describe, index and create images concludes that different parts of the brain are stimulated for the differing tasks (1988). Following similar thoughts, Ingwersen defines a successful information retrieval system design as employing a mix of cognitive structures, understanding that formulating a query is a different cognitive process than browsing or indexing (1996). Previous image indexing and retrieval research agrees with this, as the experiments typically employ more than one cognitive process as part of the research. In Jörgensen's research on image description and indexing, for example, she found a great difference in the number and kinds of words her participants used, depending on the task they were requested to perform, whether it was describing or indexing (1996).

Future Research

This research suggests that image indexing and retrieval may still in a discipline-dependent phase, evidenced by the differing results of the brief research project and this complete survey, as discussed in this paper. It appears that the amount and kind of terminology used to describe images is dependent on the group describing them. More cross-disciplinary research is needed to develop an image indexing and retrieval approach that is suitable for any user, in any situation. This research suggests that an integrated approach, one using a blend of low-level features and high-level concepts, may eventually provide the foundation for this type of successful system.

Summary

This research explores the potential for collaboration of computer scientists and library science by providing a conceptual discussion of the primitive feature of color. This research is suggestive of how scientists may approach the development of future image indexing and retrieval systems, using the strengths of both content- and concept-based approaches. Without efforts to integrate the approaches from both disciplines, the two lines of research will continue along their separate paths without informing each other. This research suggests a framework of integration that can be used by the scientists working in both fields.

Notes

1. *Library of Congress Thesaurus for Graphic Materials* is available online at http://lcweb.loc.gov/rr/print/tgm1/.

2. *Art and Architecture Thesaurus* is available at http://www.getty.edu/research/conducting_research/vocabularies/aat/; ICONCLASS may be found at http://www.iconclass.nl.

3. See http://www.getty.edu/vow/AATFullDisplay?find=&logic=AND&note=&subjectid=3001 31647 for the color hierarchy description.

Appendix A

UNC
SCHOOL OF INFORMATION AND LIBRARY SCIENCE

Welcome to the Image Description Study!

Introduction to the Study:

- We are inviting you to be in a research study of image descriptions. You have been selected for this study because you work with images on a frequent basis.
- Marie Kennedy, under the faculty advisement of Dr. Jane Greenberg, at the School of Information and Library Science, University of North Carolina at Chapel Hill, is doing this study to complete the requirements of a Master's degree in Information and Library Science.

Purpose:

- The purpose of this study is to analyze the types of word choices image experts use in describing images.
- We expect to be able to recommend specific changes to current automatic image indexing and retrieval programs as a result of this study.

What Will Happen During the Study:

1. We will ask you to click on a link to a Web page, where you will view and respond to four images, one at a time.
2. It is expected that this process will take about **five minutes**, but you may choose to take more time as desired.
3. There will be no follow-up contact with you, unless you desire a copy of the results of the research.

Your Privacy is Important:

- We will make every effort to protect your privacy. Any information you provide will remain confidential, due to SurveyMonkey's secure server. Only the Investigator and her faculty advisor will have access to the data. All personal identifiers will be destroyed after data collection.

Your Rights:

- Your participation in this study is completely voluntary and no risks are anticipated for you as a result of participating.
- If you decide to be in the study, you will have the right to stop being in the study at any time.

Institutional Review Board Approval:

- This research is being carried out with the support of the Academic Affairs Institutional Review Board at the University of North Carolina at Chapel Hill, (919) 962-7762, aa-irb@unc.edu.

Thank you for participating in this study. Your input is valuable in strengthening how images are indexed and retrieved in our day-to-day Web use.

Sincerely, Marie Kennedy, under the faculty advisement of Dr. Jane Greenberg.

Please print a copy of this page for your records. By clicking on the following Web link you agree to participate in the study. **To complete the survey, go to: survey website.**

UNC SCHOOL OF INFORMATION AND LIBRARY SCIENCE

Image Description Survey

During this survey you will view and respond to three images, one at a time. After you answer the question below you will be automatically directed to the first image. When you feel that you have viewed the image well enough (give yourself about 15 seconds), click NEXT to be directed to a text box, where you will type in a description of the image. You will repeat this process a total of three times.

To ensure that the images you get will be randomly selected, please answer the following question:

Think about the day of the month that you were born. Is that an odd or even number?

- [x] odd
- [ ] even

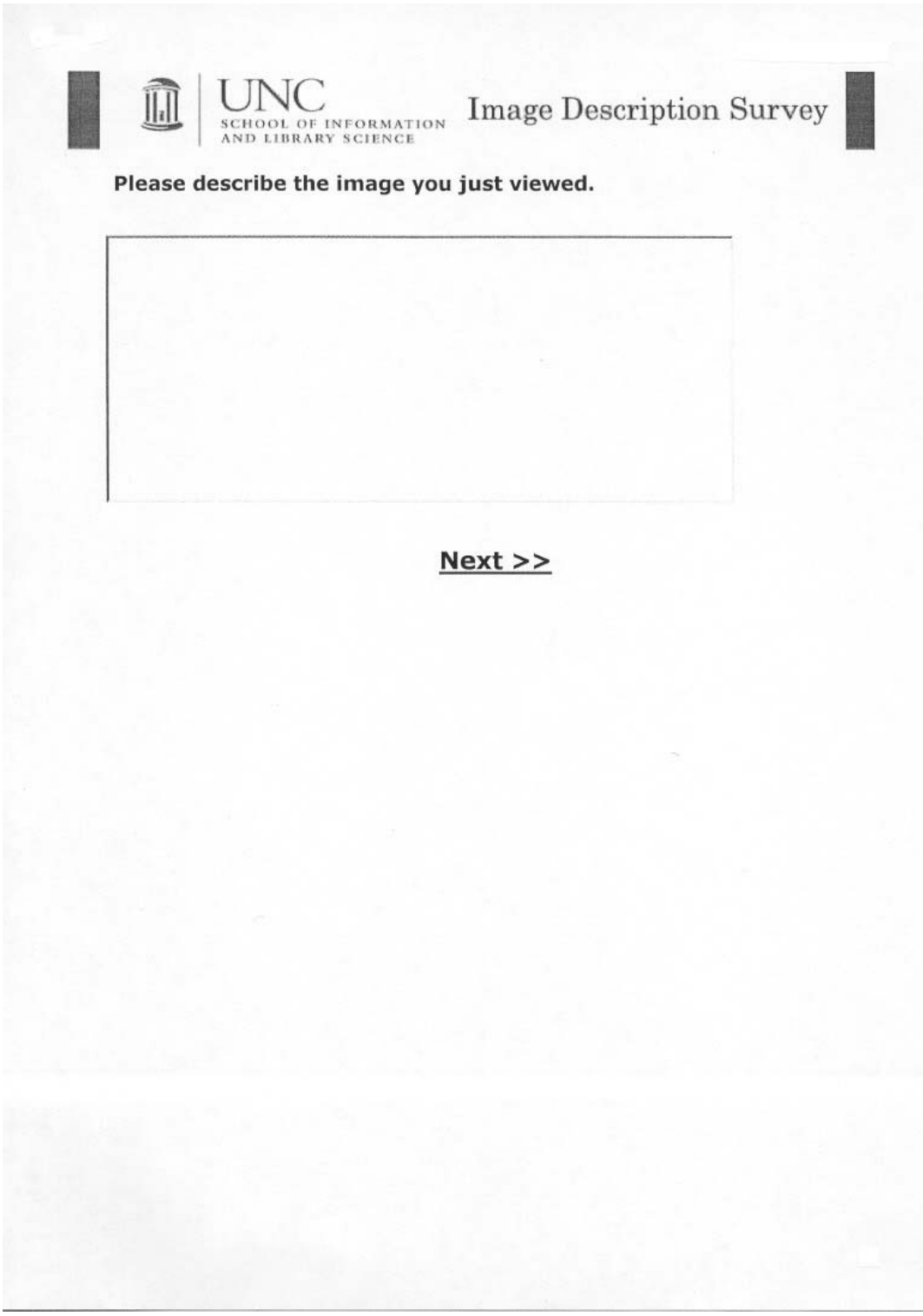
Next >>

UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Image Description Survey

After viewing this image, click NEXT to go to the description form.

Next >>


UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE
Image Description Survey
Please describe the image you just viewed.
Next >>

UNC SCHOOL OF INFORMATION AND LIBRARY SCIENCE

Image Description Survey

After viewing this image, click NEXT to go to the description form.

Next >>

Please describe the image you just viewed.

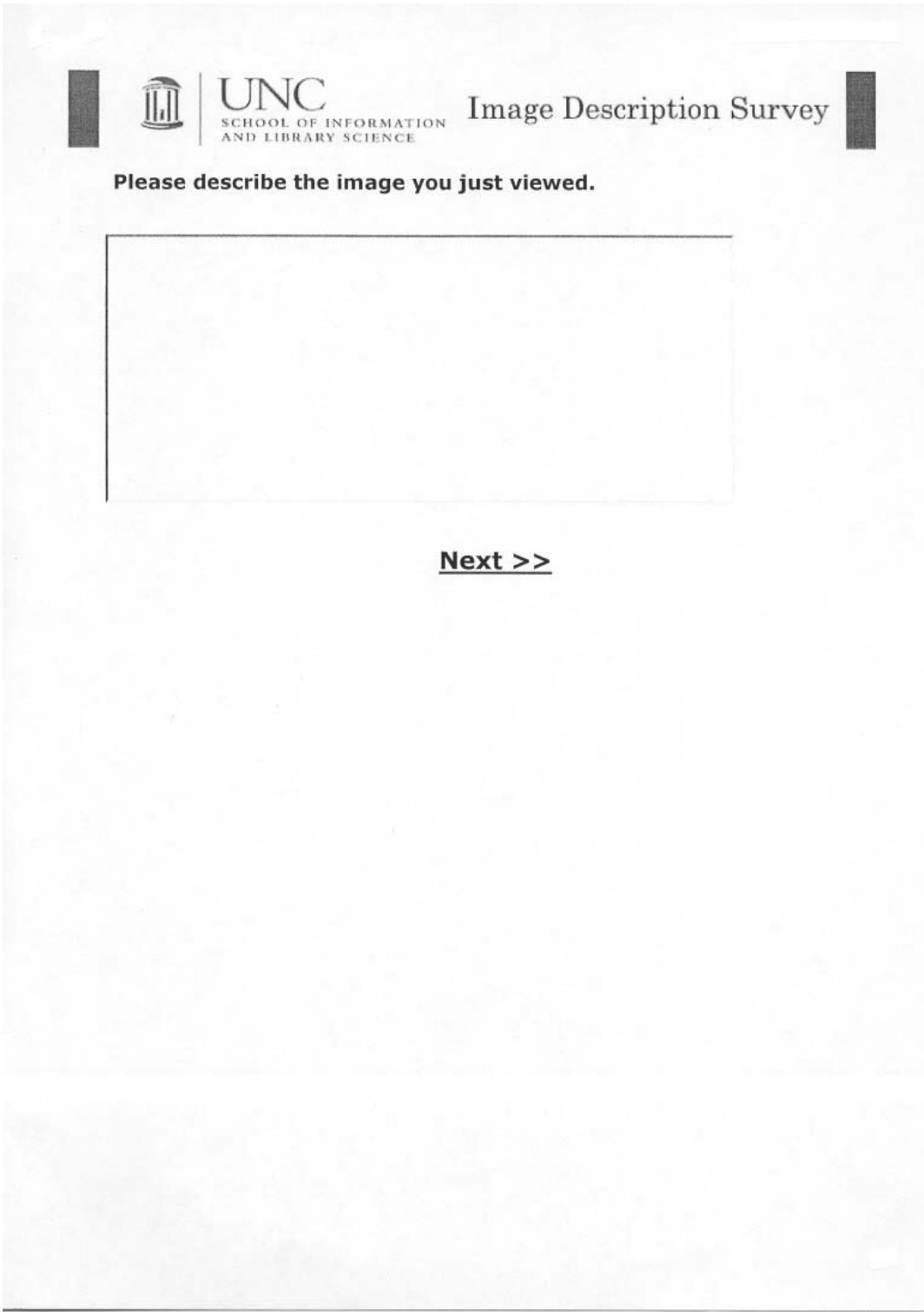
Next >>

UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

Image Description Survey

After viewing this image, click NEXT to go to the description form.

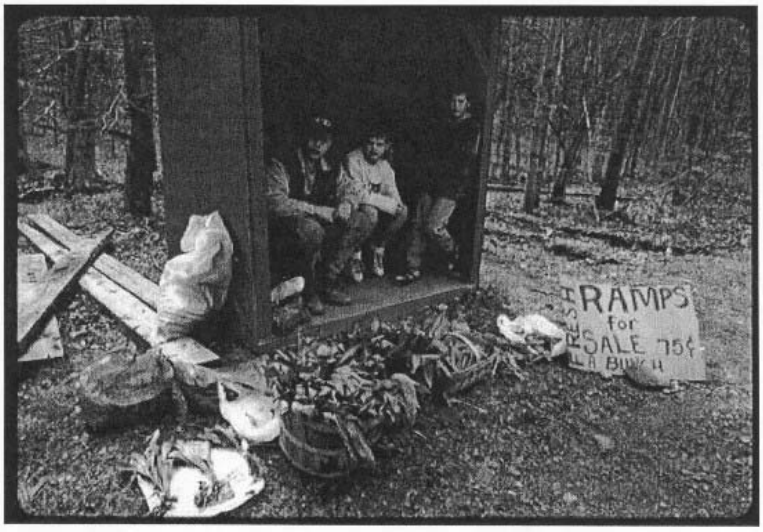
Next >>


UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE
Image Description Survey
Please describe the image you just viewed.
Next >>

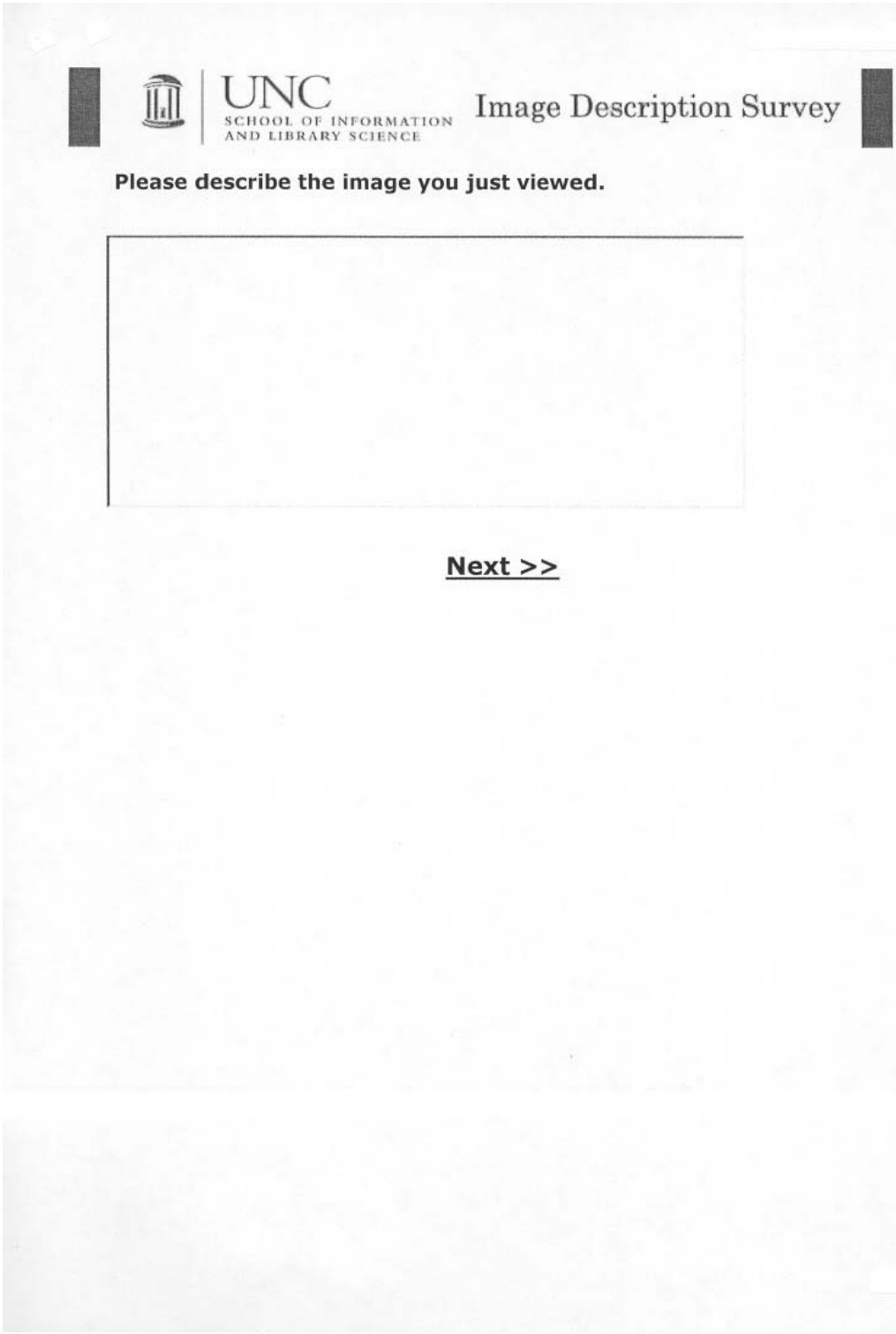
UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE

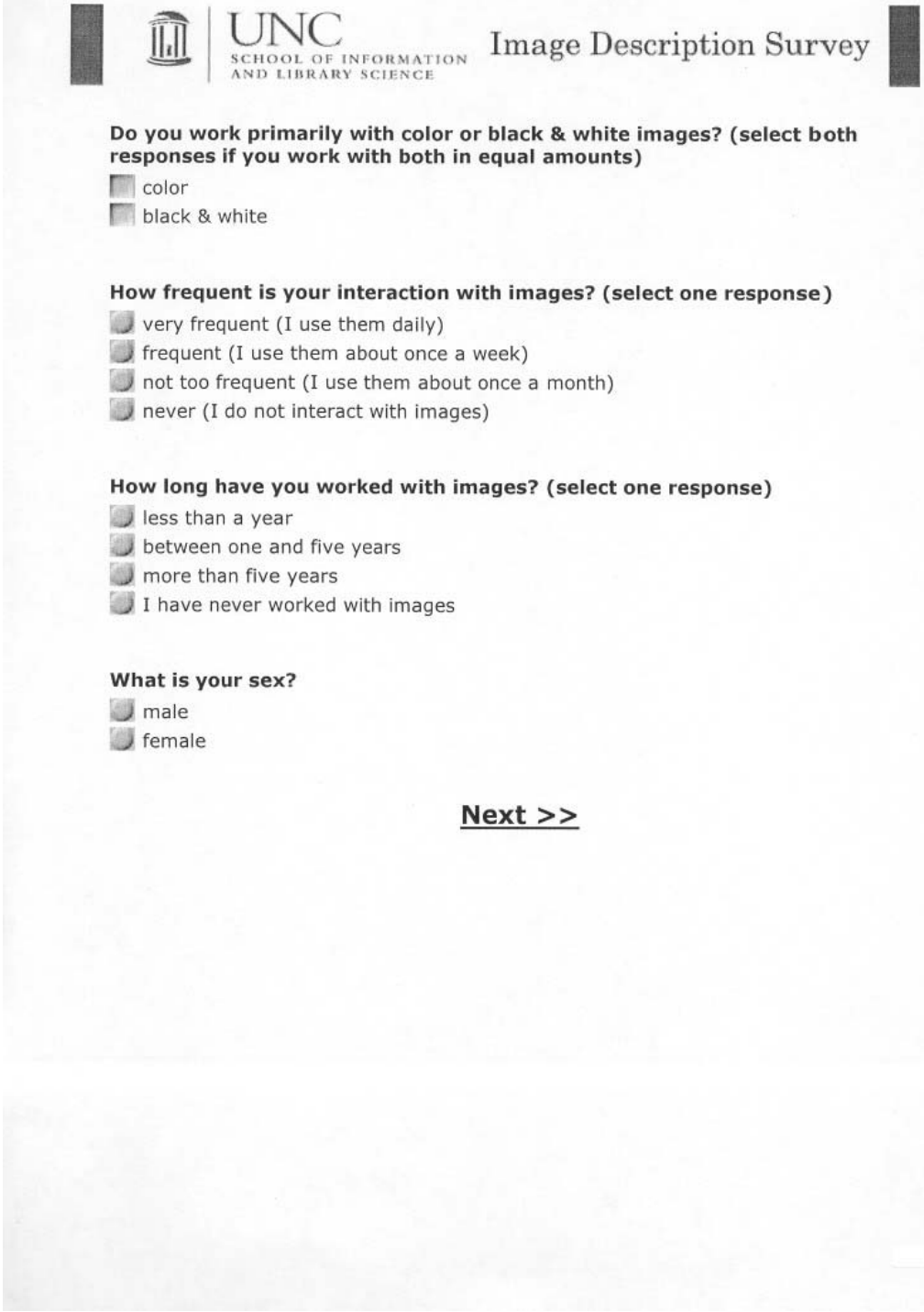
Image Description Survey

After viewing this image, click NEXT to go to the description form.



Next >>


UNC
SCHOOL OF INFORMATION
AND LIBRARY SCIENCE
Image Description Survey
Please describe the image you just viewed.
Next >>



UNC SCHOOL OF INFORMATION AND LIBRARY SCIENCE

Image Description Survey

Do you work primarily with color or black & white images? (select both responses if you work with both in equal amounts)

- color
- black & white

How frequent is your interaction with images? (select one response)

- very frequent (I use them daily)
- frequent (I use them about once a week)
- not too frequent (I use them about once a month)
- never (I do not interact with images)

How long have you worked with images? (select one response)

- less than a year
- between one and five years
- more than five years
- I have never worked with images

What is your sex?

- male
- female

Next >>

Image Description Survey

Thank You!

Summary of the Research Question: As more images are being uploaded to and retrieved from the Web, strong efforts must be made to effectively share these resources. Though significant work is being done in this area, the two lines of research on image retrieval have not yet successfully intersected. The automatic indexing and retrieval that is accomplished through content-based image retrieval (CBIR) systems does not inform the indexing and retrieval accomplished with thesauri and controlled vocabularies through the concept-based retrieval approach, and vice versa. This study examines the concept of color, an attribute inherent to both content- and concept-based approaches. Through this Web survey descriptions of images will be gathered, with the expectation of constructing a "color" vocabulary from the descriptions that can then be employed in future automatic indexing and retrieval systems.

Methodology: At random, you participated in either the color portion or the grayscale portion of this survey. The number of descriptors you and other participants provided will be counted and compared. I will be looking to verify that there is a distinction between the responses to the color and grayscale portions, and trying to tease out what is "color" from the terms you provided in your responses.

Contact me at re@unc.edu if you're interested in seeing the results of this survey!

***All images used in this survey were taken from the Tending the Commons: Folklife and Landscape in Southern West Virginia collection of the American Memory Historical Collections for the National Digital Library* (http://memory.loc.gov/ammem/cmnshtml/cmnshome.html).**

Done >>

Bibliography

Barnard, K., Martin, L., Funt, B., & Coath, A. (2002). A Data Set for Color Research. *Color Research and Application, 27*(3), 147-151.

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.

Commission Internationale de l'Eclairage. (2004). *CIE159:2004: A Colour Appearance Model for Colour Management Systems*. Vienna, Austria: CIE Central Bureau.

Eakins, J. P., & Graham, M. E. (1999). *Content-Based Image Retrieval: A Report to the JISC Technology Applications Programme*. Retrieved January 30, 2005, from http://www.unn.ac.uk/iidr/research/cbir/report.html

Gilchrest, A. (2003). Factors Affecting Controlled Vocabulary Usage in Art Museum Information Systems. *Art Documentation, 22*(1), 13-20.

Greenberg, J. (2001). A Quantitative Categorical Analysis of Metadata Elements in Image-Applicable Metadata Schemas. *Journal of the American Society for Information Science and Technology, 52*(11), 917-924.

Gudivada, V. N., & Raghavan, V. V. (1997). Modeling and Retrieving Images by Content. *Information Processing & Management, 33*(4), 427-452.

Ingwersen, P. (1996). Cognitive Perspectives of Information Retrieval Interaction: Elements of a Cognitive IR Theory. *Journal of Documentation, 52*(1), 3-50.

Jörgensen, C. (1996). Indexing Images: Testing an Image Description Template [Electronic version] *ASIS 1996 Annual Conference Proceedings*. Retrieved January 30, 2005, from http://www.asis.org/annual-96/ElectronicProceedings/jorgenson.html

Jörgensen, C. (1999). Image Indexing: An Analysis of Selected Classification Systems in Relation to Image Attributes Named by Naïve Users. *Annual Review of OCLC Research*, 1999. Retrieved January 30, 2005, from http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000003507

Jörgensen, C. (2003). *Image Retrieval: Theory and Research*. Lanham, MD: Scarecrow Press.

Library of Congress. (2000). *Tending the Commons: Folklife and Landscape in Southern West Virginia*. Retrieved January 30, 2005, from http://memory.loc.gov/ammem/cmnshtml/cmnshome.html

Markey, K. (1988). Interindexer Consistency Tests: A Literature Review and Report of a Test of Consistency in Indexing Visual Materials. *Library & Information Science Research, 6,* 155-177.

McDonald, S., & Tait, J. (2003). Search Strategies in Content-Based Image Retrieval. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Image Retrieval* (pp. 80-87). New York: ACM Press.

Mehtre, B. M., Kankanhalli, M. S., & Lee, W. F. (1998). Content-Based Image Retrieval Using a Composite Color-Shape Approach. *Information Processing & Management, 34*(1), 109-120.

Persson, O. (2000). *Image Indexing: A First Author Co-Citation Map*. Retrieved January 30, 2005, from http://www.umu.se/inforsk/Imageindexing/imageindex.htm

Pettersson, R. (1988). Interpretation of Image Content. *Educational Communication and Technology Journal, 36*, 45-55.

Rasmussen, E. (1997). Indexing Images. *Annual Review of Information Science and Technology, 32*, 169-196.

Sánchez, D., Chamorro-Martinez, J., & Vila, M. A. (2003). Modeling Subjectivity in Visual Perception of Orientation for Image Retrieval. *Information Processing and Management, 39*, 251-266.

Shatford, S. (1986). Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging & Classification Quarterly, 6*(3), 39-62.

Smith, J. R. (2001). Quantitative Assessment of Image Retrieval Effectiveness. *Journal of the American Society for Information Science and Technology, 52*(11), 969-979.

Squire, D., Müller, W., & Müller, H. (1999). Relevance Feedback and Term Weighting Schemes for Content-Based Image Retrieval. In *Third International Conference on Visual Information Systems* (pp. 549-556). Amsterdam, the Netherlands.

Stan, D., & Sethi, I. K. (2003). *e*ID: A System for Exploration of Image Databases. *Information Processing & Management, 39*(3), 335-361.

Tieu, L., & Viola, P. (2004). Boosting Image Retrieval. *International Journal of Computer Vision, 56*(1/2), 17-36.

Town, C., & Sinclair, D. (2004). Language-Based Querying of Image Collections on the Basis of an Extensible Ontology. *Image and Vision Computing, 22*, 251-267.

Trant, J. (2004). *Image Retrieval Benchmark Database Service: A Needs Assessment and Preliminary Development Plan*. Retrieved January 30, 2005, from http://www.clir.org/pubs/reports/trant04.html

University of North Carolina. (1998). *Reducing Delays in the Work Flow of New Monographs through the Technical Services Departments of the UNC-CH Academic Affairs Library*. Retrieved January 30, 2005, from http://www.stat-or.unc.edu/consulting/or/or350/west.htm