

Images and Visualization at the 9th ASIS SIG/CR Classification Research Workshop

Corinne Jorgensen

At the 9th ASIS SIG/CR Classification Research Workshop, held in conjunction with ASIS '98, five of the eight presentations were very relevant to the interests of SIGVIS members. This overview will briefly summarize the key points in these presentations and comment upon some of the interesting issues raised by them (please note that the summaries of these papers are my own particular take on them, filtered through my own experiences and research, and may inadvertently not fully represent the authors' intentions). The full papers will be published as an ASIS Monograph, ***Advances in Classification Research***, Vol. 9, by Information Today, Inc.

A broad theme across these five papers is the interplay of classification structures and the visual representation of information organized according to some principle. The search for effective visualization and display of document surrogates to facilitate information retrieval is a complex enterprise, involving many variables, including that most unpredictable and least-known of all, the human end-user. Knowledge in this area is built slowly, and, most importantly, must be shared. It is hoped that this overview will inform others of these research efforts.

The first paper presents a theoretical discussion of "graphic language documents" and reviews pertinent background theories specifically in relationship to representations of graphic language documents in an information retrieval system.

Graphic Language Documents: Structures and Functions

Carolyn Beebe and Elin Jacob

"Graphic Language Documents" are defined as documents subsuming traditional alphanumeric texts as well as any combination of text, graphics, and/or pictures. This paper presents a theoretical exploration of these documents in a digital environment as spatially oriented objects. That is, in these documents there is a relationship between form (structure) and function and an understanding of this relationship may provide insights into ways of representing and classifying these documents. Any automated logical representation must be informed by real world relationships. The goal of his paper is to suggest categories of physical structure that can be identified by a computer system and subsequently used for representing non-text-based objects.

The paper reviews pertinent background theories specifically in relationship to representations of graphic language documents in an information retrieval system. Areas covered include Gestalt theory, Moles's theory of super signs or holistic elements of information processing, elements of semiotics, theories of cognitive categorization, and visual communication. Difficulties in representing images with verbal surrogates are touched upon. The components of traditional document representation are discussed and the frequent lack of distinction between the data and the document or object for which it serves as a surrogate is noted as problematic.

Other recent approaches reviewed cover the use of techniques to match, sort, and order pictures on the basis of visual characteristics (texture, color, shape, form, tonality, spatial organization) derived from the digitization process, such as Morelli, QBIC, Virage, and VisualSEEK, and WebSEEK. The work of another small group of researchers who have attempted to categorize graphic language documents is also reviewed. The authors discuss the work of Twyman in this section and note that the basic question of the relationship between conventions of configuration and domain-level knowledge remains a key question for the representation of graphic language documents. The domination of a textual approach to document representation, to the exclusion of other useful approaches, runs as a thread throughout the paper. In their discussion of concepts of form and function, the authors argue that "current attempts at the development of standards for representing collections of pictures are

driven by the verbally-oriented ethnocentrism of the current classificatory workforce" and call for incorporation of representations of a picture's internal spatial attributes or geometric structures as a fundamental necessity for indexing pictures for successful retrieval. They further derive their argument for this approach in a discussion of the Bauhaus school, which suggests a unity of theory, material and idea, but they also suggest that this unity cannot be captured by "merely verbalizing descriptors that would represent each of these aspects." Interestingly, the incorporation of form or structural elements as information bearing aspects to be treated in the retrieval process would also benefit text-only documents, as Liddy's work with text structures demonstrated. While focusing on what appear to be unique elements of various document types (e.g. text, images, or a combination of both) can be useful for delineation of new approaches to indexing and retrieval, we should also look for broader applications of whatever methods we may discover.

Two papers are concerned with thesaurus display and construction and the applicability of visualization techniques to thesauri. The first paper is a practical look at software requirements for end-user thesaurus display and the need for a uniform retrieval interface when multiple domains are combined in one system. The second paper establishes a theoretical basis for thesaurus construction and proposes a graphical interface to assist in the process.

Generalized Software Requirements to Access Thesauri and Classification Schemes for User-based Image Collections

Barbara Barnes, Eric H. Johnson, Jennifer B. Young, and Pauline A. Cochrane

This paper describes indexing requirements for two diverse image collections, editorial cartoons and dance videos. It also proposes software that would give the end-user access to thesaurus terms and a graphical display of the faceted classification structures of indexing records. Editorial cartoons are rarely indexed, and access points are limited, rarely addressing the layers of meaning typical in such a cartoon. The authors describe a set of facets for indexing these cartoons, composed of Subject Access (topic, person, place, event, time period), Visual Access (objects, person, place text, and caption), and Item Information (cartoonist, date, publisher, etc.). In addition, the authors propose a list of processing rules for extracting data for indexing these cartoons. Currentness is a limitation of LCSH for providing appropriate indexing terms; thus controlled vocabulary terms were supplemented with indexer and caption terms. Dance materials share similar problems in terms of access to meaning; typical access points for dance videos include name of dance and people involved such as choreographer, principle dancers, dance company, etc. Three facets are also proposed for dance videos: Dance as Product (creators, dance form, and subject of dance), Process of Dancing (performers, staging, administration, etc.), and Item Information (author, title, format, contents, etc.). While parts of these facets are already well covered, "subject" remains less accessible, as interpretation of gestures and other components is often necessary to determine such points as inspiration, literary characters, behavior, emotions, and issues, among others. A set of processing rules is also presented for dance videos. Thesauri that were useful for describing the meaning of the dance videos were the Thesaurus of Sociological Indexing Terms, the Thesaurus of Psychological Index Terms, and LCSH. Name authority is noted as a problem.

Because of the diversity of facets needed to describe images from multiple domains, a more flexible cataloging scheme is needed. The authors propose a retrieval environment that supports SGML or XML and indexer-defined metadata tags. Requirements for a retrieval interface to diverse collections of images are listed and include display of facets and thesauri used for indexing and both standard and flexible search mechanisms. Several examples of a proposed graphical user interface are given. One concern with a system such as this is whether a uniform retrieval interface can compensate for a user's possible disorientation in searching among the multiple domains (and therefore non-uniform facets) across a set of images. Another concern is whether users will be able to understand these different facets and the differences between visual image description and subject description. Having a prototype system to explore these issues is a worthy next step. Two significant points are the need for utilization of multiple thesauri in fully describing a variety of visual materials and the need for access to "interpretive" attributes of images such as emotions and themes.

A Graphical Interface for Faceted Thesaurus Design

Uta Priss and Elin K. Jacob

The authors present the theory of formal concept analysis and use it to develop a formalization of faceted thesauri which can facilitate a graphical display of a thesaurus as a line diagram of mathematical lattices. In their introduction, they note the extensibility of this technique to knowledge bases and lexical databases, which, with thesauri, group words into synonym sets and define the semantic relations between the synonym sets, thus providing a means for coding semantic information into a machine readable and computer processible format. Faceted thesauri offer the advantage of flexibility over standard thesauri; however this flexibility is not easily represented with such standard tools as linear displays or tree hierarchies. Graphically, mathematical lattices can be visualized by line diagrams which represent a concept by a small circle; these lattices are a graphic aid in understanding the concepts by which similar terms are related and can assist in thesaurus construction. The authors suggest that their approach can provide a means for decomposition of facets into domain specific subfacets and provide a design environment specific to the construction of faceted thesauri.

The authors use formal concept analysis to formally define the constituents of a faceted thesaurus and discuss three types of faceted thesauri, demonstrating the flexibility of this approach, which allows the same set of objects to be associated with different terms in different facets. The authors' approach can be applied in thesaurus construction to the formation of classes and relational structure and to a lesser extent to the collection of terms. This is characterized as combining both bottom-up and top-down strategies by using both analysis of objects (extensions) of classes or concepts (term identification) with utilization of general knowledge about a generic class hierarchy. An example using these techniques for one type of faceted thesaurus is presented using a small collection of objects representing library documents. Typical facets and values of these objects are determined (the bottom-up portion) and are then modified using CODA (Conceptual Design Application), a graphical computer interface the authors are developing to edit classes and their relationships, compose and decompose facets, and determine relationships between the different facets that can be combined in nested line diagrams.

The paper thus moves from defining thesauri and facets in the context of formal concept analysis to developing a graphical tool based on this definition for use in thesaurus construction. The next step, as the authors note, is to perform usability testing of the tool to determine how much training is required for developers and end-users before they can interpret and apply the diagrams appropriately. While the approach seems to offer a powerful flexibility in the design of faceted thesauri, at present, the system appears to require a fairly sophisticated understanding of the principles upon which it is based in order to use it.

The final two papers explore the relationship between classificatory structures and browsing behavior. The researchers test browsing interfaces in quite different domains and using quite different methodologies. The first uses controlled vocabulary as the basis for document representation in the domain of fine arts images, while the second uses statistical and visualization techniques for representing text documents.

Browsing Images Using Broad Classification Categories

C. Olivia Frost and Anna Noakes

The authors of this paper describe a research project at the University of Michigan School of Information which developed and tested a text-based digital image retrieval system (the SI Art Image Browser at http://www.si.umich.edu/Art_History/). The goal of the system was to demonstrate the effectiveness of using classification techniques to retrieve browsing sets within the domain of fine arts images. The rationale for the browsing approach is discussed and, in the case of visual materials, is based on two major premises, that many concepts associated with these materials are difficult to convey using text, and that indexing should focus on recall rather than precision as the most efficient approach, as human indexing of images is highly labor intensive.

An investigation of well-established classification structures for the fine arts and art history resulted in the use of the Art and Architecture Thesaurus as the major source for broad categories for browsing. Initially, the team theorized that two systems would be necessary, one for expert users and one for general users; however, discussions with students and faculty in art history and fine arts led the researchers to conclude that one system could serve the needs of both user groups. The broad categories used included Agents (artist's names), Space (artist nationality), Time (chronological sub-divisions), Materials (medium and process), and Objects (type of work). The researchers found it necessary to add a Subject category (broad categories such as Animals, People,

Places); this set of terms was developed by the non-expert users. The Styles and periods category from the AAT was omitted in the final version because of concerns about assigning terms which are controversial even within the art history community with a high degree of confidence. The authors note that, ultimately, the determination of the classification categories and structure was most heavily influenced by the characteristics of the data provided by the image content providers. The top level categories were finalized as Artist, Nationality, Title, Object Type, Media, Date, and Subject with a total of three levels of indexing.

Research questions for the project included whether a browsing model utilizing navigable hierarchies is a useful model for access to digitized images, evaluation of various performance criteria, the relationship between visual and textual information in browsing, and the adequacy of one system in relation to different user groups. Research methods included an Ease of Use study, a Quasi-controlled study, focus group interviews, and an online survey. Results of the controlled study showed no significant difference between three different search modes (browse, search by keyword, or a combination) across the two user groups of specialist and generalist subjects. The authors suggest that this indicates that a browsing system can be an effective way for generalist users to gain access to images without having an extensive knowledge of terms and concepts in the subject domain. Both experts and non-experts expressed desire for more "contextualizing" information, such as artist biographies, art criticism, and links to other similar objects or other works by the same artist, in the system. Both groups also expressed the desire for additional search keys for image content such as color, shape, or style.

However, there were some important differences in results between the two user groups. Experts were able to complete tasks quickly and with a high degree of success, relied more heavily on textual descriptions than thumbnails, and found the search interface and top level categories, and search fields straightforward to use. Experts were quite sensitive to perceived inconsistencies in the data, especially in terms of inconsistencies in cataloging. These results are logical given the choice of a familiar, domain specific indexing vocabulary and the experience of the expert users with similar retrieval systems. Experts also noted the need for "dimensions," as the thumbnails and digital format contributed to a loss of scale. In contrast, non-experts took longer, experienced lower success, and had greater difficulty completing tasks, experienced difficulties with the search interface such as assigning search terms to the correct search fields or selecting from long lists of terms, found some categories confusing, relied more heavily on the thumbnails than the textual descriptions and selected images on a purely visual basis. Non-expert users expressed a need for the omitted categories of period and style.

One question remaining unanswered by this study is whether "real" non-expert users in real-world situations would continue to use a system in which they experience the difficulties as noted above. While it is clear from the results that the system appears to serve the needs of expert users and extends the functionality of traditional retrieval systems very nicely with the addition of browsing capabilities, it would be beneficial to investigate this question further. Some of the differences between non-expert and expert users reported here are very interesting, such as the notion that non-experts rely more upon visual strategies, and may point to the need to include broader search heuristics which allow more flexible searching. Also, the tasks which expert and non-expert users bring to an image base are likely to be significantly different and may require different sets of search strategies. While browsing is undoubtedly a useful addition, incorporation of some of the non-experts' suggestions such as "search all fields," exact match, "close match," and "loose match" options should be investigated to overcome the rigid structure imposed upon this browsing system by the incorporation of parts of a hierarchical, highly domain-specific thesaurus. This could hopefully allow the benefits of both a browsing and a controlled vocabulary approach to work in concert with each other.

Cluster-based and Association-based Visualization Systems as Information Exploration Tools

Min Song

The context of this paper is to improve information exploration of unknown data collections, specifically by using visualization tools to partition heterogeneous data collections into manageable sub-spaces which the user can navigate. Because of the large amounts of text available through modern information technologies, tools are needed to represent text content in a succinct way without necessitating traditional reading methods of inspecting, sifting, and synthesis. One way of accomplishing this is to use visual representations based upon content abstraction and spatialization of the original texts. The author previously developed and tested a

visualization system of document space named BiblioMapper; the research presented in this paper investigates the influence of document classification methods on information visualization by comparing a clustering algorithm adapted in BiblioMapper and a neural network algorithm, a Kohonen Self-Organizing Map (SOM). Both of these use descriptors generated by term-frequency-based automatic indexing techniques.

A major research question for information visualization is how effectively documents are represented in terms of these text features. Most visualization systems depend on neural network algorithms and classical clustering algorithms have generally not been used. The study reported here investigates the usability of cluster-based and association-based classification methods in document visualization, using a minimum distance pairs technique (BiblioMapper) and an unsupervised learning method (the Kohonen SOM map algorithm). The two systems had slightly different interfaces and search techniques. BiblioMapper presented a screen with terms representing topic clusters; clicking on a cluster number brings up a list of topics. In the Kohonen SOM interface, the users can search using both an alphabetical subject list or a map display.

The database tested contains 800 document surrogates from the Internet related to computer science and information science. Subjects were ten undergraduate and graduate students in the Humanities with no online searching experience. Evaluation was focused on comparing the two systems with respect to their effectiveness as information exploration tools. The findings found no significant difference in the two systems in terms of selected documents, completion time, accuracy, and navigation or interface style. Therefore, the assumption that the Kohonen SOM would correspond more closely to users' mental models (based as it is upon the theory of the associative neural properties of the brain) was not supported. While both systems appear equally competent in providing a browsing interface for exploratory activities, the author notes that query-based search techniques should also be provided. Subjects relied on the alphabetical subject list rather than upon the map in the Kohonen SOM, and searches involving keywords not appearing in the visual space were difficult and time-consuming. The author notes that the subject domain did not appear to be of interest to the subjects; the search tasks used vocabulary unfamiliar to the participants. It would be enlightening to repeat the experiment with subjects who have more domain knowledge and understanding of the questions and more interest in the topics presented.