
Introduction

BETH SANDORE

THE DEVELOPMENT OF INNOVATIVE RETRIEVAL APPROACHES for access to visual information is among the most significant of technological, conceptual, and institutional challenges for the library and information science community. On a technological level, Gupta and Jain (1997) suggest that visual information retrieval involves a hybrid of older computer science disciplines, including the analytical component of computer vision and the query component of database systems. On a conceptual level, it is clear that humans employ a variety of socio-cognitive processes, as well as sensory skills, in the retrieval and evaluation of visual information. From an institutional standpoint, libraries, archives, and museums have entered into digitization projects, varying in scope and scale, the results of which are beginning to challenge the manner in which information is described, stored, and delivered. While visual resources have long been part of the slide library that supports use in the arts and humanities communities, Charles Rhyne (1996) indicates that technology has enabled a rapid increase in the use of pictures in other disciplines as well as by the general public (p. 4).

The primary goal of this issue of *Library Trends* is to present practitioners, researchers, and educators in the areas of library and information science, archives, and museums, as well as “imagists” working with visual resources in any setting, with a current perspective on the development of visual information retrieval and access tools. The issue’s scope is limited to the analysis and retrieval of bit-mapped or raster images and video (images that are comprised of pixels of varying color information values) and does not include work with vector graphics (images encoded as numeric formulas that represent lines and curves—e.g., Geographic

Beth Sandore, Digital Imaging Initiative, 452 Grainger Engineering Library, University of Illinois, Urbana, IL 61801

LIBRARY TRENDS, Vol. 48, No. 2, Fall 1999, pp. 283-288

© 1999 The Board of Trustees, University of Illinois

Information Systems [GIS]). The contributions provide perspectives from researchers and practitioners—specialists in the areas of library and information science and computer science. In planning this issue, a conscious effort was made to include a perspective on the developing foundation of visual information retrieval, as well as work representing current and experimental systems. The issue is divided into three sections—I. Foundations of Intellectual Access to Visual Information, II. Implementation and Evaluation, and III. Experimentation.

Since 1988, two issues of *Library Trends* have been devoted to various aspects of image and multimedia information retrieval. In each issue, the editors call for a synergy across the disciplines that develop image retrieval systems and those that utilize these systems. Stam and Giral (1988), in the issue of *Library Trends* titled "Linking Art Objects and Art Information," emphasize the need for a thorough understanding of the visual information-seeking behaviors of image database users. Writing in a 1990 issue of *Library Trends* devoted to graphical information retrieval, Mark Rorvig (1990) takes up the fundamental issue that "what can be listed cannot always be found" and uses that statement as a framework for examining progress in intellectual access to visual information. In the ensuing decade, several critical events have unfolded that have brought about some of the needed collaboration across disciplines and have enhanced the potential for advancements in the area of visual information retrieval.

First, the field of computer vision has grown exponentially within the past decade, producing tools that enable the retrieval of visual information, especially for objects with no accompanying structural, administrative, or descriptive text information. Second, the Internet, more specifically the Web, has become a common channel for the transmission of graphical information, thus moving visual information retrieval rapidly from stand-alone workstations and databases into a networked environment. Third, the use of the Web to provide access to the search and retrieval mechanisms for visual and other forms of information has spawned the development of emerging standards for metadata about these objects as well as the creation of commonly employed methods to achieve interoperability across the searching of visual, textual, and other multimedia repositories. Practicality has begun to dictate that the indexing of huge collections of images by hand is a task that is both labor intensive and expensive—in many cases more than can be afforded to provide some method of intellectual access to digital image collections. In the world of text retrieval, text "speaks for itself" whereas image analysis requires a combination of high-level concept creation as well as the processing and interpretation of inherent visual features. In the area of intellectual access to visual information, the interplay between human and machine image indexing methods has begun to influence the development of visual information retrieval systems. Research and application by the visual

information retrieval (VIR) community suggests that the most fruitful approaches to VIR involve analysis of the type of information being sought, the domain in which it will be used, and systematic testing to identify optimal retrieval methods.

Section I—"Foundations of Access to Visual Information"—is intended to provide a background in the familiar concept-based approach to describing and retrieving images, as well as the more recently developed content-based approach to visual information retrieval using inherent features such as color, shape, and texture. The importance of the articles in this section cannot be over-emphasized. In their own way, each clarifies the inevitable need to consider the interaction between high-level semantic concepts and inherent content in VIR. Content retrieval, the area which is newest to the library and information science community, will demand increased understanding and analysis in order to determine its value to users as we build more robust and lasting visual information retrieval systems. The authors in section I emphasize the need for a greater understanding of the interplay between concept-based indexing (performed by humans) and the automatic or semi-automatic process of indexing an image or a video sequence (using software) based on inherent image attributes. In "Intellectual Access to Images," Hsin-liang Chen and Edie M. Rasmussen explore current image retrieval systems and analyze the methods that have been employed to provide intellectual access to the various image collections. Throughout the article, the authors focus on the problems that are inherent in image description and access with the objective of identifying traditional and new solutions to these challenges. The second contribution, by P. Bryan Heidorn, presents a framework for understanding image retrieval from the standpoint of the user's cognitive models for seeking visual information. Heidorn examines the process by which models, based on linguistic and inherent visual attributes, are constructed and employed by users in seeking visual information that answers particular queries. He also discusses the types of information that can be found to have common values in the socio-cognitive sense. In "Computer Vision Tools for Finding Images," David Forsyth describes and discusses the use of two types of methods that have been developed in the computer vision field to facilitate the searching of images by inherent content. Forsyth groups these methods into two categories—"appearance methods" that compare images based on their overall content (e.g., color histograms, texture histograms, spatial layout), and "finding methods" that focus on matching subparts of images with the goal of identifying and finding specific objects. Forsyth explains, in terms understandable to a broad range of readers, the complexities involved in identifying and matching inherent features within whole images and corresponding objects within segments of images. He is careful to explain that computer vision tools cannot be used in monolithic ways to resolve user queries of large

image collections but rather describes areas in which they have been found to be most useful and promising for future development.

Section II—"Implementation and Evaluation"—focuses more specifically on the implementation and evaluation of visual information retrieval systems with cultural heritage information since this is a primary interest of libraries, museums, and archives. In the first article in the section, Teresa Grose Beamsley addresses the challenge of securing and ensuring image integrity—an issue that is integral to the quality of VIR system search results yet often goes unremarked in discussions beyond the initial point of digital capture. Beamsley's article, "Securing Digital Image Assets," examines the issues involved in the process of securing the image content in a VIR system. Beamsley indicates that images delivered across the Web are usually low-quality compressed derivatives of higher quality archival digital images. Often, the only tenuous link between the low-quality derivative image and its original digital image is contained in the textual metadata that accompanies the image. Beamsley examines various approaches that institutions can use to secure the integrity of this representative information while pointing out the concomitant challenges in doing so. Throughout her article, Beamsley focuses on the need to achieve a balance between the desire for ownership and authenticity of images and the provision of open access to cultural heritage materials, particularly in public institutions.

In the second article in this section, "Getting the Picture," Caroline Arms provides a thorough description and analysis of Library of Congress efforts to provide access to visual information from their Prints and Photographs Division collections. The Library of Congress' work in this area is of international significance because they are one of the few institutions that makes images in their collections and experimental projects publicly accessible through the Web (except in cases of copyright and ownership restrictions), as well as information regarding the technical underpinnings of their efforts. Arms's account gives important insights into the institutional challenges of providing unified public access to disparate digital collections while addressing the special issues associated with VIR and other programmatic concerns. The article will be useful for any institution considering the development of a digital library or facing the organizational challenges of identifying unified aspects of collections that are otherwise disparate in physical location and diverse in content.

Christie Stephenson, in "Recent Developments in Cultural Heritage Image Databases," uses the Museum Educational Site Licensing (MESL) Project as a point of departure in her exploration of developments in the broad area of cultural heritage image databases. Stephenson draws on her own experience in the management of the MESL project as well as on the work of others to identify the various factors that are known to affect VIR including metadata quality, image quality, display and manipulation

features, and the diverse retrieval results due to a variety of methods employed in search engine indexing and retrieval. She also reviews examples of recent work in the development of federated image repositories at various institutions, as well as advances in user interface design for VIR. Stephenson's writing reiterates the point made by Forsyth, albeit from a different perspective, that it is critical to identify the information-seeking needs of specific user groups in order to tailor the most effective retrieval methods and interfaces to specific domains of use.

In her article "Evaluation of Image Retrieval Systems: The Role of User Feedback," Samantha Hastings reviews problems in current VIR evaluation research, presents the preliminary results of a Web-based study of user searching of an art image database, and proposes a framework for user-centered evaluation studies of VIR systems. In particular, Hastings's findings note that over half the user queries submitted in the study were satisfied by the review of thumbnail images. Further, Hastings emphasizes the importance for users to have the capability to construct customized browsing approaches to retrieved image sets and to manipulate images to view more visual detail and to compare more than one image.

Section III—"Experimental Approaches"—presents articles describing three research projects that examine various aspects of image or combined image and text retrieval methods. The articles by Rohini K. Srihari and Zhongfei Zhang and by Neil C. Rowe focus on experimental systems that employ both text and image analysis in the development of effective methods for retrieving highly relevant image sets. Srihari and Zhang define their subject domain specifically. They focus on analyzing faces in pictures and their related captions taken from Web-based newsfeeds such as MSNBC and CNN. From this standpoint, these authors are able to tailor their retrieval algorithms to achieve a fairly high level of precision for this domain of multimedia information. Rowe uses similar methods to analyze databases of images that feature a range of activities at a naval aircraft test facility. Rowe's approach differs from that of Srihari and Zhang in its objective, which he states is to broaden the applicability of linguistic and text processing routines in an attempt to create a more general retrieval process targeted toward increasing precision in the searching of Web-based information in general. The final article in this issue, by Yong Rui et al., focuses entirely on experimental methods using computer vision tools, testing various methods of image analysis (vector analysis, Boolean, fuzzy match) that are commonly used in text processing, and using relevance feedback to "train" the search engine and improve the relevance of the retrieved image result set. Rui et al. use an experimental multimedia system, MARS (Multimedia Analysis and Retrieval System), which they have developed over a period of several years, to test various approaches to the analysis and retrieval of inherent image features. They have tested their methods with various

image sets, including a set of cultural heritage images of artifacts from the UCLA Fowler Museum of Cultural History.

The work represented in this issue suggests that a number of professional communities are contributing different but essential components to the development of useful and innovative image retrieval systems. In spite of the great technology strides in multimedia, image database developers and image content holders continue to grapple with the fluid issues of organization, access, retrieval, delivery, and representation. The words of art historian Barbara Stafford (1996) describe the differential treatment accorded visual imagery over the centuries in Western culture, and they express the hope that computers will be the tool that enables imagery to become a trusted, valued, and rich vehicle (similar to text) for information delivery:

Yet in spite of the arrival of what I have termed the "age of computerism"—rapidly replacing modernism and even postmodernism—a distorted hierarchy ranking the importance of reading above that of seeing remains anachronistically in place. All the while, computers are forcing the recognition that texts are not "higher" durable monuments to civilization compared to "lower" fleeting images. These marvelous machines may eventually rid us of the uninformed assumption that sensory messages are incompatible with reflection. (p. 4)

Despite Stafford's apparent ambivalence, the significant levels of traffic on the Web support the perspective that technology has begun to fuel an important shift in the value that society has previously placed on the written word over things visual. Computers now enable users to incorporate images of art and other works into their own personal information contexts—images which have for centuries been a powerful and efficient medium for conveying landmark concepts, emotions, and events. The concomitant challenge for libraries, museums, and archives also involves a shift—not only in technology and practice but also in focus—i.e., to equip ourselves with an effective understanding of the similarities and differences between text and multimedia information retrieval, and to use this knowledge as a foundation for developing effective access and archiving methods.

REFERENCES

- Gupta, A., & Jain, R. (1997). Visual information retrieval. *Communications of the ACM*, 40(5), 70-79.
- Rhynne, C. S. (1996). *Computer images for research, teaching, and publication in art history and related disciplines*. Washington, DC: The Commission on Preservation & Access.
- Rorvig, M. E. (1990). Introduction. *Library Trends*, 38(4), 639-643.
- Stafford, B. M. (1996). *Good looking: Essays on the virtue of images*. Cambridge, MA: MIT Press.
- Stam, D. C., & Giral, A. (Eds.). (1988). Linking art objects and art information (theme issue). *Library Trends*, 37(2), 117-264.