Research on Human-centered Design and Evaluation of Multimode Image Retrieval

C. Olivia Frost, School of Information
Roberta Johnson, Space Physics Research laboratory
Sang Wook Lee, Electrical Engineering and Computer Science
Yi Lu Murphy, Electrical and Computer Engineering, UM-Dearborn

Position

Image retrieval systems should be multimodal - and exploit the retrieval potential of both image and text.

Image retrieval systems should be designed for the casual user, as well as for users in research domains or with high technical skills. There is a particular need for systems which can be used by younger audiences who are growing up with much more exposure to graphic expression.

Systems for image retrieval, as with other systems designed for users, should incorporate user input in the design and development.

Project Description

We describe a project underway at the University of Michigan to explore the development and evaluation of multimode retrieval schemes which employ both image and text and are multiple path, iterative, and user-directed. The project is conducting grounded research about the creation and use of digital libraries which will support general and non-specialist users in finding information. The testbed contains digital image collections for earth and space sciences. The research is focused by the design, construction, deployment, and evaluation of a digital image library testbed for students at the middle and high school level.

The project is organized around the synergistic intersection of three sub-activities: relevant basic research in user interface design for an image library to serve middle and high school users; image classification and retrieval using text and image contents; design and construction of an evolving testbed system; and its deployment, use, and assessment. These activities are built upon the complementary strengths of a team of faculty members from information retrieval, computer engineering, and space science.

We feel that effective image retrieval systems for non-specialist users will allow users to perform searches which are multiple path, iterative, and user-directed, and part of our research focuses on the integration of text and image content search strategies. While there is a substantial amount of completed and ongoing research in both textual searching as well as content-based retrieval of images, much remains to be done to see how effectively these approaches can complement each other, and how the "casual" or nonspecialist searcher would go about using them to retrieve images.

We have developed a prototype image library currently running on the WWW (http://www.si.umich.edu/Space/). The system currently has a collection of over 1,200 images in the area of earth and space science. The system provides users three different image seeking paths: text-based browse, text-based search, and image content-based search. In designing browsing strategies in which the user's question is framed in textual terms, we drew upon a prototype developed in an earlier project for art images. The text tag-based browse system allows the user to browse by classification categories, and the browsing path has multiple levels. Each node at the second level is associated with a set of thumbnail images. Once the user has selected a category to search, and then chooses from the items enumerated under that category, she can then choose either to see more thumbnail images or select a particular thumbnail image. The text-based search path provides a more direct textual search access, allowing the user to type in keywords within the classification categories.

Our image library also provides mechanisms for the user to retrieve images based on image content. The current image database software allows users to visually search and sort a collection of images based on intrinsic visual attributes such as color and texture and thus enables the user to search using visual information. The user selects an image and asks the system to "Give me more pictures that look like this". The system returns the search results by displaying the images on the screen in the order of similarity to the submitted image.

We are developing a search engine which provides multiple search paths which use integrated information of text and image contents. Users can browse guided by a hierarchical classification tree and search by queries, in order to narrow the textual search to the point where the system can efficiently employ content-based analysis. We are currently augmenting the existing image retrieval software with textual retrieval strategies. Under this path, the user is guided via a classification framework from broad to progressively narrower subtopics. After the user has identified the topic areas of interest, she can then identify candidate images which can be used to perform a similarity search. This approach benefits the non-domain user who is not familiar with terminology or taxonomy or topography of the discipline, and the hierarchical structure allows the user to start out from a broad, general category and work her way down to a more specific one. If a user chooses "comet" by browsing the classification menu, she retrieves a set of thumbnail images, and can then submit the image of her choice to the image search engine to retrieve, for example, a set of comet images containing fading tails, or comets with pinwheel patterns.

In developing both manual and automatic indexing procedures, we intend to utilize methods for deriving as many image-context descriptors as possible based on the user's input and interaction. Human users can play a critical role: (1) in building image-contextual information before an image is inserted into a database, and (2) in developing semi-automatic methods for image context-based query. Our classification categories will be derived in part from our user group. In a pilot study to be implemented this summer, students and teachers from middle and high school who participate in the study will be presented with a set of images, and be asked to specify the content of each image. We will compile an initial list of classification terms based on the outcome of the user study as well as national standards in use for science education.

Our research will focus on the development of computer vision algorithms that extract high level image contents and derive similarity functions to measure the closeness of a matching between a high level image content query to an image in the image collections. For each high level image content descriptor such as "debris", "pinwheels", etc., we will present the potential users a set of primitives extracted from an image that contains the descriptor, and user's input will help us to develop a computer vision algorithm to extract the image features and match the image features with the query.

Our focus is on the usability of image retrieval systems for casual users, and on the way in which users combine image and text retrieval strategies. We are designing a retrieval system for the domain of earth and space science, and with non-specialist users in mind, specifically middle and high school students. However, our goal is also to provide a scheme which is generic and will eventually have application to many domains, including the humanities and the sciences, and to many types of users, including the general adult user.

The University of Michigan Digital Image Library (UMDIL) project will be linked to the architecture of the University of Michigan Digital Library (UMDL), sponsored by NSF/NASA/ARPA, and the Windows to the Universe project, sponsored by NASA. Windows to the Universe is an effort to develop, implement, deploy and test a World Wide Web site designed for the general public on the earth and space sciences. The UMDIL image collection testbed will be integrated into the Windows to the Universe interface, allowing immediate access to the existing audience of the project, and the outreach institutions currently working with the Windows to the Universe project (primary schools, museums, and libraries) will be involved in the evaluation and evolution of the testbed.