Dynamic Overview Techniques for Image Retrieval [*]

Pearl Pu and Zoran Pečenović

Ergonomics of Intelligent Systems, DMT/ISR
AudioVisual Communications Lab, LCAV/DSC
Swiss Federal Institute of Technology Lausanne
EPFL, CH-1015 Lausanne, Switzerland
(pearl.pu,zoran.pecenovic)@epfl.ch

Abstract. One difficulty often overlooked in information retrieval systems is that search criteria themselves are often poorly defined. People describe their information needs in many different ways and frequently change their goals depending on the current results of their search. We have investigated the hypothesis that overviews of the space of available solutions are a good way to remedy this situation. Our overview techniques allow users to get a feel for the meaning of categories through randomly chosen examples, find similar images using content search, and to inspect the global distribution of images according to certain criteria. Users thus organize the retrieval task into an iterative browsing process that makes them specify their queries more accurately. As a result they are more satisfied with what the system retrieves.
Keywords: Overview techniques, image retrieval, active query, direct search, exploratory search.

1 Introduction

Users have very diverse and complex ways to express their needs for visual information. Some describe a goal such as "I am looking for a horse for my web page because I am an animal lover." Some use a more pictorial expression such as "I want photographs of bluish landscapes." Yet others look for images of an abstract concept, "I want images to represent intelligence." Existing image retrieval systems assume that users can formulate their needs either in terms of keywords, visual content, or color composition. As pointed out in [7], average users need more exploratory search techniques to browse image collections before a precise query can be formulated. We thus treat image retrieval as an information need clarification process consisting of task description, needs expression both in verbal and visual terms, and exploratory search for targets and their alternatives. In addition, we offer overviews of the entire search space to allow comparison and evaluation of alternatives in combination with content search. This active query approach is in contrary to current systems which are more suited to professional users whose needs for images can be more precisely formulated in their queries.

In this paper, we make use of CIRCUS, a content-based image retrieval system (see Sect. 3.1) using visual features such as color, layout, texture, or any combination of them. Furthermore, to support all three stages of active query, we have added a set of dynamic overview techniques that allow users to become familiar with the collection of the images before specifying the query, either using keywords, exploratory, or direct search methods.

2 Current Methods for Image Retrieval

To understand how non-professionals look for images, we conducted a user study on a set of 40 computer literate subjects. We asked them several questions among which to tell us the type of images they are looking for, describe how they usually do the searching. The following give an idea of the different answers recorded:

- I look for images for presentations. I have a Corel CD of images. I browse the book that comes with the CD until I find one that I like.
- I looked for images to represent intelligence as a logo for my research group on intelligent systems. I used the WWW based search engines. I finally discovered Corbis and was able to type the word intelligence.
- I was looking for images of paintings of nature.
- I like to find images of nature with lakes in a bluish background, definitely not dark reddish background such as those in a sunset.
- I'd like to find red images because I like the color red.

To analyze how some of the existing systems accommodate users' needs for images, we searched for images according to the users' task descriptions at two web sites where images can be retrieved by keywords, by browsing a category taxonomy, or by content.

2.1 Picture experience at Corbis

This site (http://www.corbis.com) offers 1,000,000 images online and supports both keyword based search and category browsing. Users can choose to search images in any of the categories: all, digital images, e-cards, and prints. Keywords search supports single word, or composition of words using connectors and, or, and not. However, keyword based search cannot be used in addition to category browsing, or vise versa.

Typing the word "intelligence," we obtained 690 images. Some of the examples are shown in Fig. 1. Our user was quite happy with the results. Typing the phrase "paintings of nature" resulted in only one image. Not satisfied with the image, we typed in the words "painting and nature" (word order is not important), 5 images were found. Still unsatisfied with these few possibilities, we then typed in "painting," and obtained 720 images. But surprisingly, a large number of the initial results are images about painting, and not of paintings as shown in Fig. 2.

Realizing that recently a category search method has been added to the site, we drilled down following Art and Architecture → Fine Art → Paintings and found 1748 images of paintings. Browsing this collection in a linear fashion, we found several images of paintings of nature. Unfortunately, we were unable to find more images of the same kind using content search.
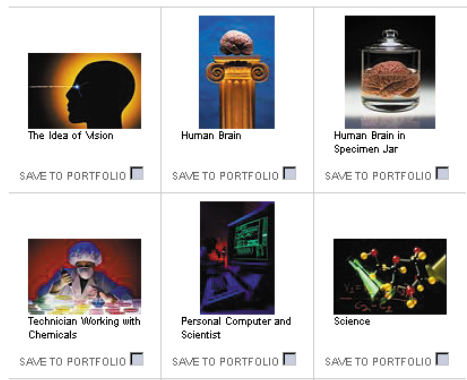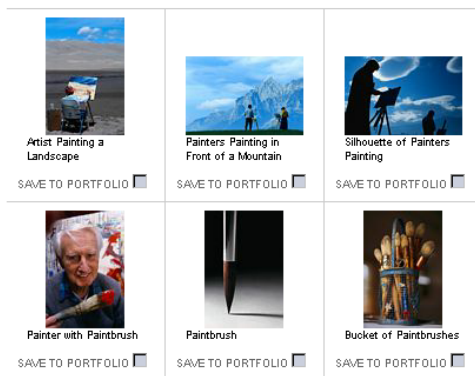
Fig. 1. Images representing "intelligence"

Fig. 2. Images about painting

Typing in "red," we obtained 720 out of 13,000 images, of which many images are indeed red in color, but few of the initial sets contain images of the Red Square Fig. 3). The results of red images improved when we selected the digital picture category.
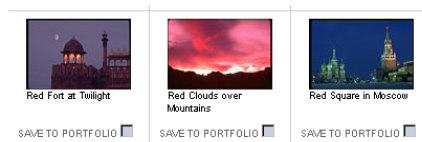

Fig. 3. Some results for "red" query from Corbis (only the second actually red).

2.2 QBIC system

QBIC [5] is a content-based image retrieval system. The online demo[1] offers three collections: U.S. stamps, stock photography and trademarks. Users can search an image collection by randomly sampling through it or by typing keywords. Content search can be used to find more images of similar visual qualities of a sample image by clicking on color, layout, or texture. It is not possible, however, to use combinations of these visual features.

Figure 4 shows the images that we found by typing lake and then clicking on one image that was our near target, shown as the first one on the upper-left corner. This is a significant improvement from the search method that uses keywords only. Here users have the option to either describe needs in words, in visual aspects, or a combination of the two. Unfortunately QBIC's keyword search is not very powerful since no results

Fig. 4. Images similar to the top-left lake with a bluish background.

were returned for "intelligence," only four images returned for "red," and three returned for "painting".

When users describe their image needs abstractly, and once they have a good idea of the database contents, the keyword based method is natural and effective. The images retrieved corresponded to the subject's previous notion of intelligence: brains, eyes, images of scientific experiments. Furthermore, serendipitous findings are thought provoking and educational. For example, the user has never thought of using an image of Einstein to represent intelligence. However, quite often users got false hits because the image collection's taxonomy is not externalized into visual structures to allow accurate formulation of the keyword query.

Furthermore, keyword and category based methods supported poorly query formulation when visual aspects of images are important to a user's need. As indicated in our study, both paintings of nature and red images took a while to find satisfactory results.

---

[1] http://wwwqbic.almaden.ibm.com/stage

3 Overview Techniques For Image Retrieval

Overview techniques have been used in some document retrieval systems (Starfield [1] and FilmFinder [2]), text retrieval (Tilebar [6]), software visualization (SeeSoft [9]), visualization of web navigation (WebBook and WebForager [4]), and visualization of a collection of photographs (DynamicTimeline [8]). These techniques have shown considerable successes to help users navigate in a complex information space, orient themselves, zoom in on interesting details, and select the targeted items. Furthermore, overview techniques allow users to easily locate, compare and evaluate alternative data items. This feature is especially useful when users have formulated a query that results in a futile search.

The key in designing powerful overview techniques lies in providing users with the following set of characteristics:

1. A compact and dynamic display technique that allows users to see the structure, organization, and content of the entire information space;
2. An easy and real-time navigation method to get to any data of the information space to examine details;
3. A multi-attribute scatter plot to dynamically organize the information space to tradeoff space so that users are able to compare and evaluate alternatives in order to select optimal solutions.

In this paper we will present three overview techniques for combining exploratory and direct search methods into one system for image retrieval. We first briefly describe the underlying content-based image retrieval system, CIRCUS.

3.1 The CIRCUS IR System

The retrieval model used by CIRCUS[2] is an adaptation to image retrieval of the Latent Semantic Indexing [3] method. The core idea is to extract a compact, useful representation of the relations between terms and documents. This is achieved through a lower-rank approximation of the term-document occurrence matrix. LSI shows a slightly higher semantic level and better performance than using term weighted vector space models (*idf-tdf* and extensions). However its translation to image retrieval does raise the fundamental issue, namely the identification of the "terms" images are made of. We opted for a low-level, visual, description of the image including vector quantized color histograms, wavelet based texture and layout descriptors. Since LSI permits the use of any countable descriptor, we include as well higher-level textual annotations. The visual features are extracted both globally and locally.

CIRCUS offers a query construction interface allowing query: by example (Fig. 5a), by color (Fig. 5b) and by sketch (not complete yet). It also represents query results in simple lists (Fig. 5a) and in a 2-D "orbit diagram" (Fig. 6). This 2-D tradeoff space represents resulting images according to their similarity along several user selectable axes (color,texture,layout,keywords). This view gives the user a chance to understand how the system perceives the above mentioned notions of similarity and helps her/him to come up with more accurate and meaningful queries.

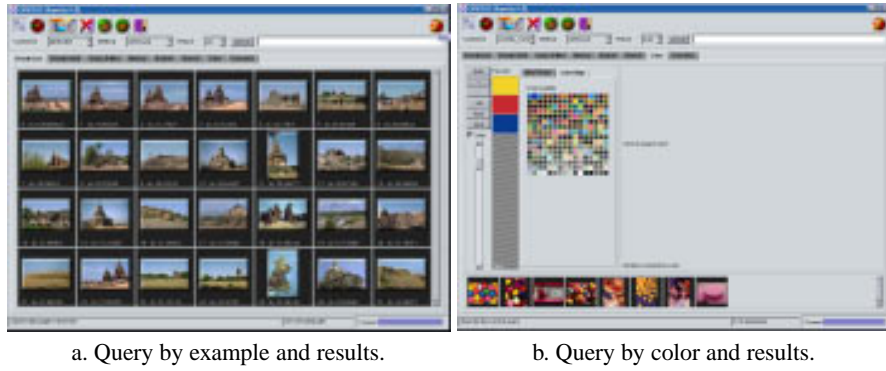[2] A simple demo and further information is available at http://lcavwww.epfl.ch

| a. Query by example and results. | b. Query by color and results. |

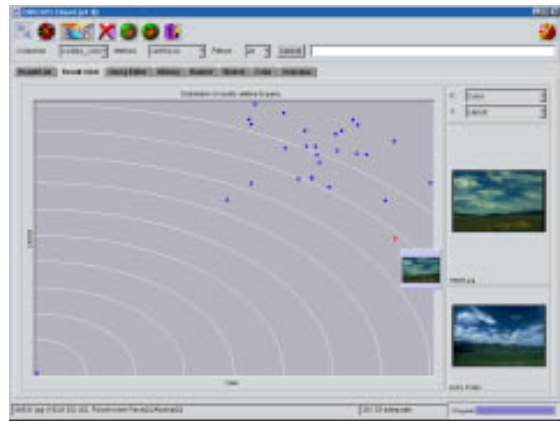**Fig. 5.** Query specification tools and sample results.

Fig. 6. Orbit: color on X; layout on Y axis.

3.2 Visual taxonomy

The first approach to creating an overview of an image collection summarizes the collection into a visual taxonomy using a single screen space. In one collection, we visualize 9 categories with a total of 650 starting images. An example image from each category is displayed along with the name and the number of images contained in each category. Users can browse through the collection by using the random sampling button so that the surface images change accordingly (Figure 7). If an image somewhat satisfies the user' needs (the near-target image), s/he can use content search to get similar ones in the same or in any combination of categories. Figure 8a shows the first 25 matches after we selected "painting" and "nature". The results are ordered in concentric circles with the example image either in the upper-left corner or more adequately in the center of the display, and the closest ones immediately around it. In the nature category, we are able to single out landscape photos with bluish/gray background (Fig. 8b).
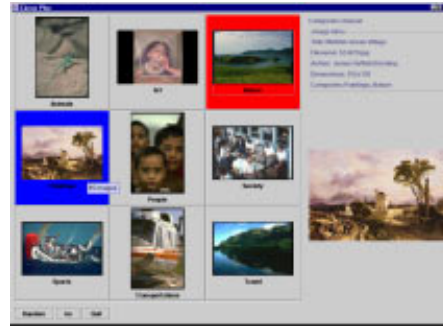
Fig. 7. Browsing the visual taxonomy

a)paintings or photos of nature   b)images of lake scenery in a bluish background

Fig. 8. Visual taxonomy and similarity retrieval

3.3   Real-time display and Galaxy overview

The second approach displays an overview of an image collection as a galaxy of thumb-nails (shown in Fig. 9). It is a real-time display system offering zoom and pan capabili-ties. The axes meanings can be defined by the user among a choice of visual and alpha-numeric attributes. The characteristics of an image that map to numerical values (such as hues, lightness or "texturedness") are displayed with a continuous axis. Other dis-crete characteristics like annotation, class or color correspondences to a fixed color-map are displayed on a discrete axis with additional either random jitter or jitter according to a third attribute (usually name of image) for spacing. For instance when the average hue (vertical) and category (horizontal) axes are chosen, users can easily spot images of particular color properties in a given category. The horizontal displacement of the stars within the same category is in this case given by the image name.

For example, natural, bluish photographs of landscapes are located in the lower-left corner. Users can navigate by zooming in on a region of interest and/or panning through the constellations, either by direct manipulation or by clicking on an thumbnail. Once the images become large enough, some details are displayed next to them (author, title, etc). The image at the center of the view becomes the current target and is automatically
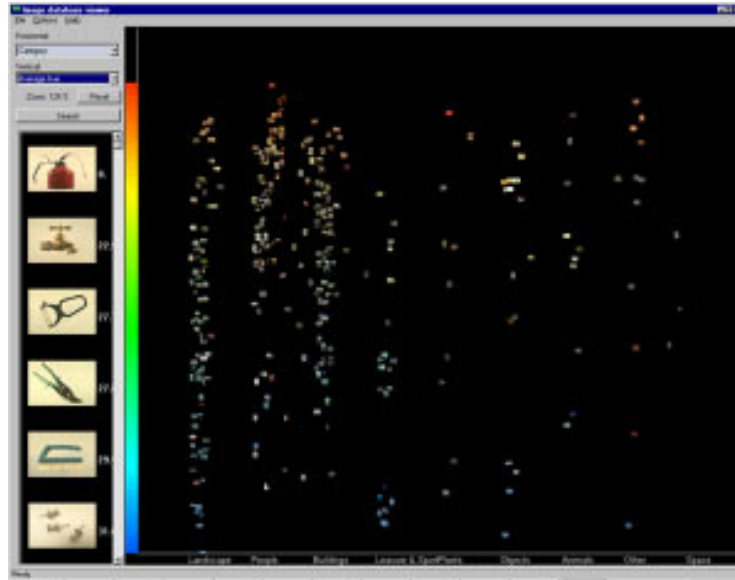
Fig. 9. A galaxy of images spread by Category and Name on X axis, and by Hue on Y axis. (Reproduced in color at the end of the volume.)

centered into the viewing area. The user can then execute a similarity based search to get more similar images.

This overview also combines the navigation and tradeoff space into a single display. That is, for users who had vague search criteria, the tradeoff visualization helps them decide what is the optimal space to explore. For example, consider a user who is looking for images of tools on uniform backgrounds. Browsing in the galaxy where the horizontal axis is category and the vertical axis is hue, s/he will soon realize that there are more photographs of tools on bright then on dark backgrounds and will narrow down the field of view to the appropriate region. S/he will find a number of near-optimal images, then it is sufficient to track on a good candidate and CIRCUS will display a list of similar ones in the side window.

3.4 Overview 3: Combining all

While many systems can support different search methods, this technique allows the simultaneous use of keyword, category, and color composition based search methods in any combination. Selecting the category painting and using the sliders to define red color, the system returns red paintings in a linear nearest-first list, thus providing an overview of the red paintings. Once a suitable image is identified a similarity search can be executed and the cluster of red paintings can be visualized in 2D space using radial coordinates Fig. 10) or Cartesian coordinates (Fig. 11). The position of each thumbnail is determined by the level of similarity according to the user defined axes (color, layout, texture), additionally the size of the thumbnail can be assigned a third dimension of

similarity. In the Cartesian layout the similarities are mapped directly on the axes; in the radial layout, the choice of the mappings is constrained by the largest varying attribute being mapped to the angle coordinate. These representations allow easy browsing of alternatives. This approach also makes the machine's measurement on certain qualities apparent for users and thus facilitates subsequent query specifications. A small user
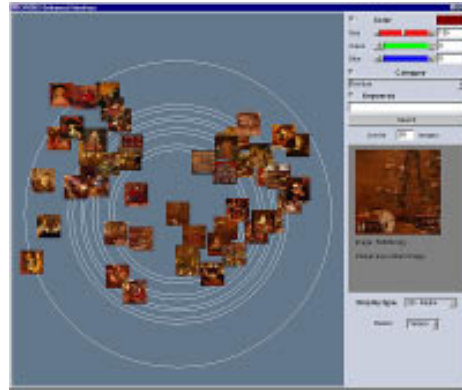

Fig. 10. Results satisfying category Painting, color red and similarity to a given image (center) in 2D radial coordinates.
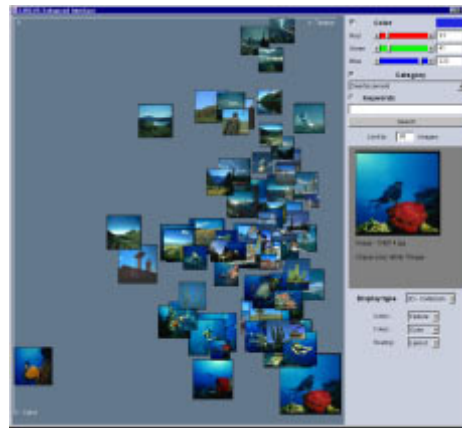

Fig. 11. Blue "sports and entertainment" images in Cartesian coordinates. (Reproduced in color at the end of the volume.)

study showed that satisfaction with the results increased with the use of overviews and comparative result representations. The other major benefits of such approaches was

the ability they confer to the user to understand the systems functioning allowing them to improve future query constructions.

4 Conclusion

We have argued that most available image retrieval systems lack some important functionalities, namely the integration of visual & semantic query capabilities, the ability to present an overview of the available data to the user, and finally a way of communicating to the user the systems conception of match and similarity. We have then shown several overview techniques combining visual taxonomy, real-time display, navigation tools, and multi-attribute scatter plot methods into an exploratory and direct search based image retrieval system. This approach models users' information needs as a process, thus offering an open taxonomy to reduce false hits, a dynamic display of an image collection in Cartesian space to show alternatives and tradeoff analysis, and a hybrid method to match and optimally exploit humans' verbal and visual abilities.

References

[1] C. Ahlberg and B. Shneiderman. Visual information seeking: Tight coupling of dynamic query filters with starfield displays. In *Proceedings of ACM CHI'94, Information Visualization*, volume 1, pages 313–317, 1994.

[2] C. Ahlberg and B. Shneiderman. Visual information seeking using the filmfinder. In *Proceedings of ACM CHI'94, VIDEOS: Part I: Browsing Navigation*, volume 2, page 433, 1994.

[3] M. W. Berry, S. Dumais, and G. O'Brian. Using linear algebra for intelligent information retrieval. *SIAM Review*, 37:573–595, Dec. 1995.

[4] S. K. Card, G. G. Robertson, and W. York. The webbook and the web forager: Video use scenarios for a world-wide web information workspace. In *Proceedings of ACM CHI 96, VIDEOS: World Wide Web*, volume 2, pages 416–417, 1996.

[5] C. Faloutsos et al. Efficient and effective querying by image content. *Journal of Intelligent Information Systems*, 3:231–262, 1994.

[6] M. A. Hearst and J. O. Pedersen. Visualizing information retrieval results: A demonstration of the tilebar interface. In *in Proceedings of ACM CHI'96, VIDEOS: Visualization*, volume 2, pages 394–395, 1996.

[7] A. Kuchinsky, C. Pering, M. Greech, D. Freeze, B. Serra, and J. Gwizdka. Fotofile: A consumer multimedia organization and retrieval system. In *Proceedings of ACM CHI'99*, 1999.

[8] R. Kullberg. Dynamic timelines: Visualizing the history of photograph. In *Proceedings of ACM CHI'96, VIDEOS: Visualization*, volume 2, pages 386–387, 1996.

[9] J. L. Stephen and S. G. Eick. High interaction data visualization using seesoft to visualize program change history. In *Proceedings of ACM INTERCHI'93, Video: Visualisation*, page 517, 1993.
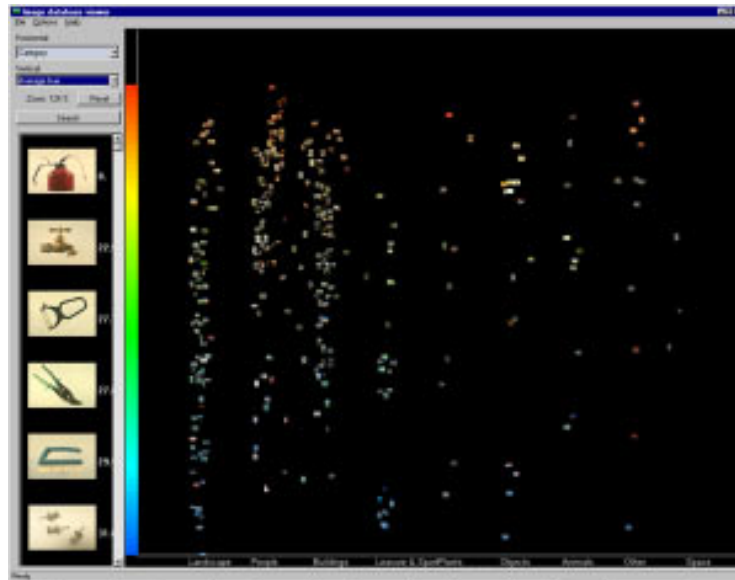
Fig. 9. A galaxy of images spread by Category and Name on X axis, and by Hue on Y axis.
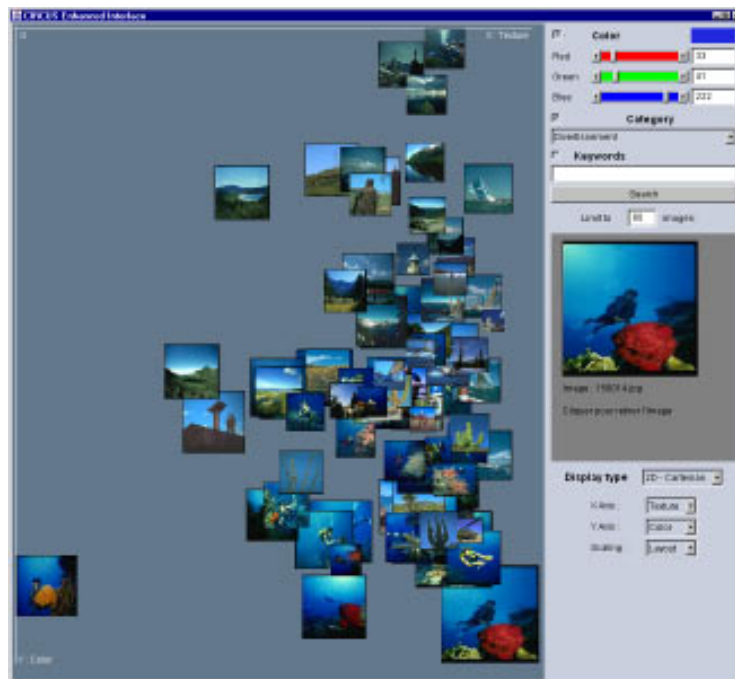

Fig. 11. Blue "sports and entertainment" images in Cartesian coordinates. The query image is in the lower-right corner, we show *similarity* on Color on X and Texture on Y axis. The size represents similarity on Layout.