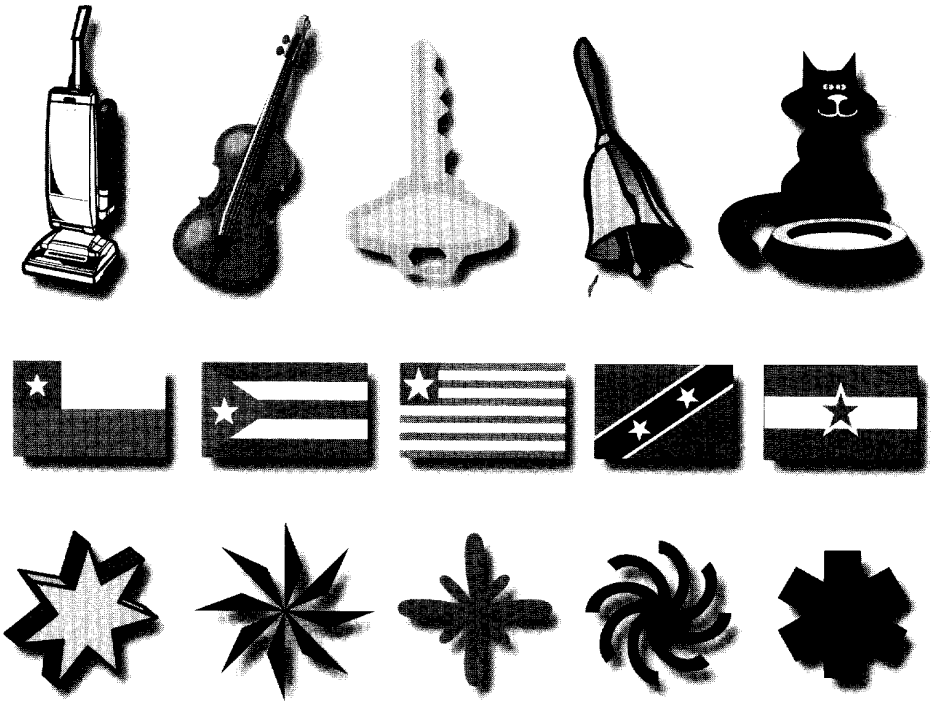
Searching for Images



by Similarity Online

by Péter Jacsó

Searching by similarity, using a known item to find other similar ones, is a common strategy for optimizing the precision of a search. What works for text-based searches should also work for finding images. Unfortunately, description and classification of images lags far behind those of textual documents. Although there are exceptions, images often have only a very short title, or perhaps a more descriptive caption. Collections of artwork may have keywords about the artists, the medium, the period, or the object, but these are not as descriptive or precise as those used for journal articles, conference papers, or dissertations.

→

Searching for Images by Similarity Online

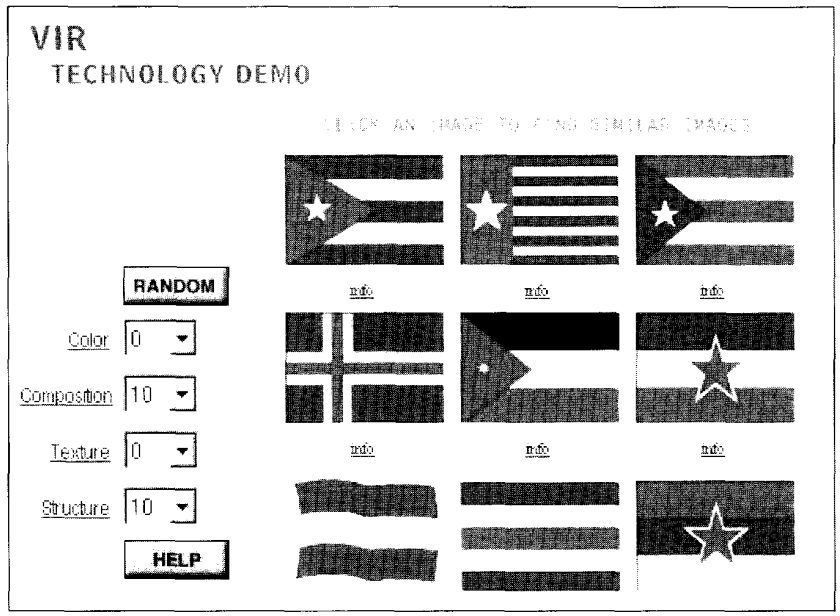
Keywords have been used to describe photographs, drawings, paintings, and other visual works of art by stock photo agencies, clip art collectors, and curators of museum collections, but words are often inadequate to describe such items. Paula Berinstein has written about using keywords and how the popular Web search engines find images in a recent column in *ONLINE* [1].

Shape, texture, color, brightness, and proportions are essential for defining images. While keywords and classification codes can help limit the domain of a search, similar items are best found by using an image's inherent traits.

Some significant developments in search software help users find images by using visual attributes as search criteria. Three companies (Virage, Excalibur, and IBM) have developed image retrieval software that uses an image's traits (shape, texture, brightness, etc.) to locate similar images in a database or on the Web. Obvious applications for these image retrieval offerings are within organizations, such as museums, but the implications and potential extend also into the business realm. IBM has already proven that searching by similarity using software that features this technique has promising applications in the world of trademark retrieval as well.

This article is an overview of how the image retrieval software from these three companies works, based on demo image collections on their Web sites. Their software is usually licensed to third-party vendors for specific applications, unlike the software employed by the major search engines. But understanding how similarity searching works opens new vistas for online searchers and others interested in information retrieval.

FIGURE 1



Result of similarity search based equally on composition and structure

VIRAGE

(http://www.virage.com)

Virage's VIR (Visual Information Retrieval) image engine offers four visual search attributes (color, composition, texture, and structure). Each can be assigned a weight from 0 to 10. Searching by color is the most straightforward. The software examines the hue, tint, and saturation of the base image selected, and looks for other images in the collection that have the most similar combination of these color attributes. Composition looks for similarity in the relative location of the colored areas. Texture examines granularity, pattern variations, and repetitiveness of patterns. Structure similarity searches recognize similar boundary characteristics of objects with clearly defined edges.

You may set one or more attribute weights to optimize your search. It takes some experimentation to find the best balance, but the search process is very fast and you may choose to look at 3, 6, 9, 12, 15, or 18 thumbnails in the result matrix.

Color similarity works perfectly well. It takes only a second or two to find images that have similar color combinations as, for example, New England foliage in the fall. If all the other attributes are set to zero, images of any objects may show up as long as they have the same hues and tints, such as a scarf or a cloth. Setting composition and structure at mid-weight along with color at maximum weight will increase the likelihood of retrieving images of trees and forests with the orange, brown, and auburn mixture characteristic of New Hampshire in September.

Emphasizing structure retrieves images where the shape of the object is similar to that of the base image. In my tests, it did not find as strikingly similar images as the other image search engines, but it may have to do with the size of the image collection at its demo site.

Emphasizing colors based on the American flag would retrieve those flags that have blue, red, and white hues, irrespective of shapes or forms.

It is important to understand that the similarity of an image shape has nothing to do with the object. A map of Italy will retrieve an image of a Gucci boot, just as a map of Scandinavia will find a picture of a lion resting after a good meal.

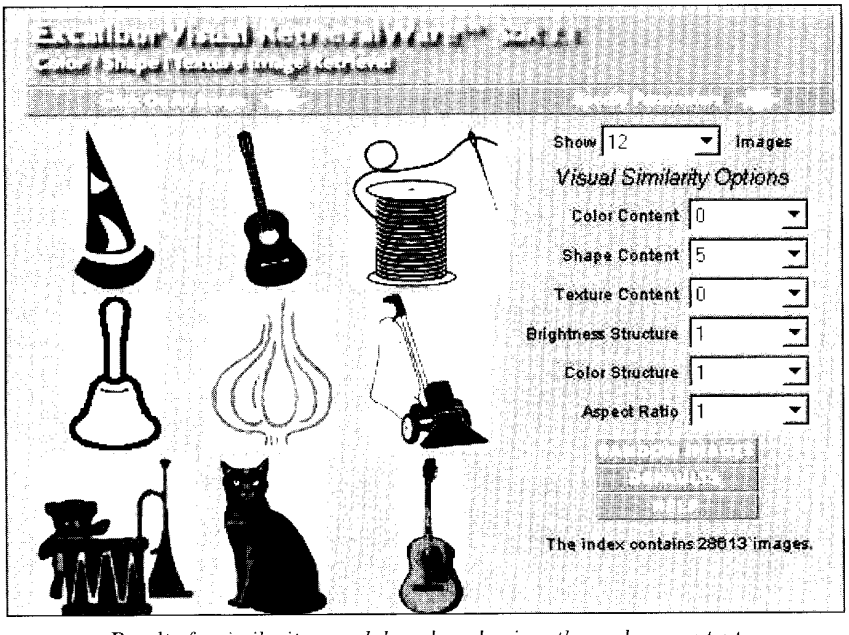
The most telling test searches were related to national flags (Figure 1). These are very appropriate for testing different settings since there are many flags that use stars, stripes, and two or three colors. Emphasizing colors based on the American flag would retrieve those flags that have blue, red, and white hues, irrespective of shapes or forms. Emphasizing structure based on the same flag would retrieve flags sporting stars and stripes in any color, such as the flags of many African and Middle Eastern countries with stars and stripe motives in green and yellow colors. The flags also serve as an excellent example to see how their sequencing changes by varying the weight for any of the four attributes. The order of the thumbnails in the result matrix will go by decreasing similarity. Clicking on the caption brings up a few details of the image, including the similarity ratio as calculated by Virage.

EXCALIBUR
(http://www.excalib.com)

Visual RetrievalWare is only one of the search programs offered by Excalibur. This is the search engine licensed by—among others—Yahoo! and Infoseek. However, neither implementation offers the user-controlled setting of image attribute searching that is the strength of this program. At its demo site, you may test the sophistication of the search parameters.

Excalibur offers six attributes for image searching: color content, shape content, texture content, brightness structure, color structure, and aspect ratio.

FIGURE 2



Result of a similarity search based predominantly on shape content

Color content is the measure of the base image's colors and their ratio, but not their location, which is controlled by the color structure parameter. The shape content attribute is a measure of the relative orientation, curvature, and contrast of lines in the image. Texture content is the measure of the smoothness or roughness of the image, sort of the tactile surface characteristics of a painting. Brightness structure measures the brightness of the pixel combinations that make up the picture. This can be a powerful tool to find images similar to the beautifully blurry photographs of David Hamilton.

Each attribute can be weighted on a scale of 0 to 5. For example, in searching primarily by shape content, the weight of that attribute is set to 5 with the rest of them set to 0 or 1. Varying the weight will yield different results from the set of over 28,000 images.

Figure 2 illustrates the images chosen based on the image in the leftmost top corner of the thumbnail matrix. Although at first glance, the content of the images are enormously different from the base image of the wizard's hat (a guitar, a cat, a vacuum cleaner), their shape and the proportion of the components are indeed remarkably similar in the majority of images. It is important to understand that the similarity of an image shape has nothing to do with the object. A map of Italy will retrieve an image of a Gucci boot, just as a map of Scandinavia will find a picture of a lion resting after a good meal.

QBIC
(http://wwwqbic.almaden.ibm.com)

QBIC (Query By Image Content) is the software developed by IBM's famous Almaden laboratory. It is perhaps the most widely licensed software by museums and other organizations, such as the San Francisco Museum of Modern Art or the French Ministry of Culture, that want to offer searchable art collections by similarity measures.

Although only three attributes (color layout, color percentage, and texture) can be used as search criteria for most of the prototype collections,

→

and not all of them can be customized or controlled by the user, this is impressively effective software. Color percentage defines the ratio of hues and tints in the base image and looks for images with similar color ratios. Color layout goes one step further by looking for images not only with similar colors, but also with colors in similar location in images. Texture that I found the most elusive in all of the software is a mixture of coarseness, contrast, and presence or absence of directionality of lines in images.

. . .a look at IBM's trademark image collection on its demo site reveals how important similarity searching in business (and science) applications can be.

IBM's demo site contains the two collections that demonstrate best the power of the software in finding similar images. One is the stamp collection, which allows the user to create a keyword search to limit the domain of similarity search. For example, you may specify "poet" to retrieve all the stamps that show the portrait of a poet. This is a classic keyword search that has nothing to do with image attributes. However, clicking on any of the buttons listed on the side of the stamp to be used as the base model will retrieve stamps that have similar color proportions, layout, or texture.

GETTING DOWN TO BUSINESS

So far, similarity software may sound like high-tech software gadgets for artsy folks, but a look at IBM's trademark image collection on its demo site reveals how important similarity searching in business (and science) applications can be. Developing, introducing, and protecting a distinctive sign or logo is a costly process, but it may be worth it as people often identify a brand or a product by the logo. The trademark of trademarks is the novel and unique use of typographical and design elements, and less often, color. These are not easy to describe, but everyone can easily recognize the characteristic shape and concatenation of the letters in the Coca Cola trademark, or the shape of the Continental and Nike logos.

Among the first steps in developing a trademark is to find if there are similar logos or designs. This can be a costly process. A single search by a trademark research company is about $300. It is cheaper if you do it yourself at the U.S. Patent and Trademark Office using their automated search service at about $40 per hour, but it is available only in the Virginia office. The CD-ROM trademark collection is rather expensive and excels in smart text searching. Consulting the trademark archive on microfiche is free, but it is not easy.

IBM revolutionized patent searching when it made available its version of the U.S. patent records on the Web free of charge. It is possible that IBM will do it again with trademarks. This is important because while there are good patent databases online with sophisticated search facilities, trademark searching depends on structure and pattern recognition that none of the currently existing trademark databases offer. True, trademarks are often assigned quite descriptive design codes, such as the ones below for the trademark shown in Figure 3.

261312	Quadrilaterals with bars, bands, and lines
261321	Quadrilaterals that are completely or partially shaded
261325	Quadrilaterals with one or more curved sides
270301	Geometric figures forming letters or numerals

However, the assignment of these design codes is not consistent, and not all records have them.

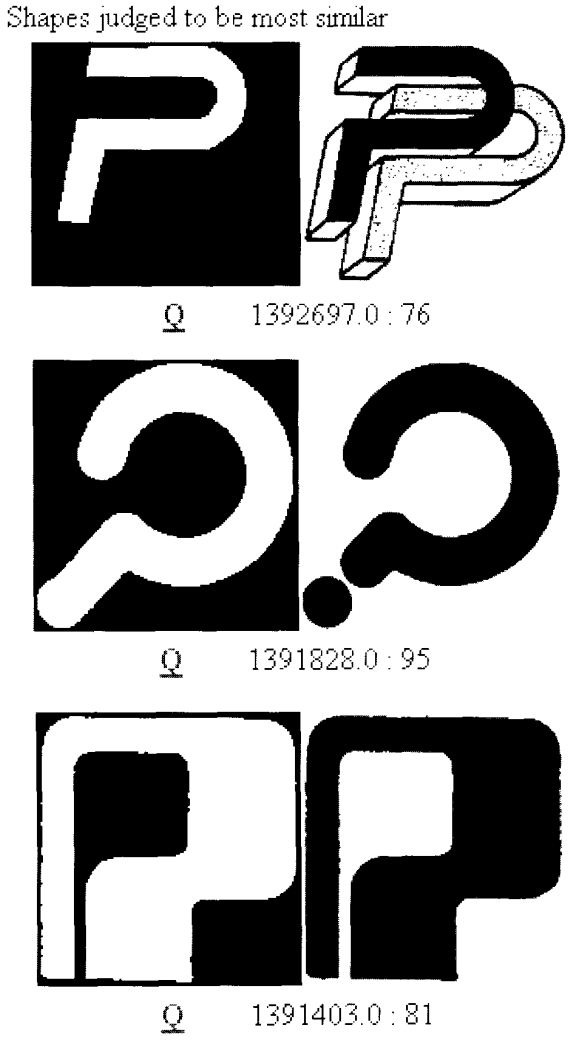
Even if design codes were assigned in a consistent manner, it is far more simple and efficient to search by design pattern. Although IBM's sample collection is very small (1,000 trademarks), the test searches proved the excellent quality of the QBIC software as customized for trademark searching.

The trademark in the right square of Figure 3 was selected as the base figure. By clicking on the base image, its shape component is extracted and displayed to the left of the original trademark, and the software goes out to find images with similar shape components. Figure 4 shows some of the trademarks and their automatically generated shape components that were found to be matching. Considering the small size of the prototype, this is an awesome result.

FIGURE 3

The base image and its shape component

FIGURE 4



Matching trademarks and their shape components

FIGURE 5

Insufficient shape component extraction

This process works best with images that have distinctive silhouettes as is the case in the example in Figures 3 and 4. Trademarks that are mostly distinguished by the content and format of text components don't lend themselves well for similarity searching, yet. Figure 5 shows two trademarks and their shape components, illustrating the limitations of purely shape-based searching.

Just as indexing of textual information on hundreds of millions of Web pages is mostly done automatically, indexing of images will follow this model, especially for images stored on Web pages. Image collections are likely to use both human indexing and classification and automatic extraction of visual attributes to facilitate finding highly relevant images swiftly. As this happens, image collections will expand exponentially, creating a need for alternate and more accurate methods of searching.

Searching for images by similarity has promising applications in many areas. In law enforcement, the sketch of a suspect can be compared against tens of thousands of digitally stored mug shots in a few minutes. A physician will be able to let the computer compare an X-ray of an emergency patient's broken collar with those in the digital archive that includes X-rays, diagnosis, and operating procedures for similar conditions. By the time digitization of visual documents will be prevalent, the search tools will be there for efficient retrieval by visual attributes.

REFERENCE

[1] Berinstein, Paula. "Turning Visual: Image Search Engines on the Web." *ONLINE* (May/June 1998): pp. 37-42.

Communications to the author should be sent to ***Péter Jacsó****, Associate Professor, Department of Information and Computer Sciences, University of Hawaii, 1030 Aoloa Place 207-A, Kailua, HI 96734; Fax 808/956-5835; jacso@hawaii.edu.*