Content-Based Image Retrieval -A New Emerging Trend For Image Description

Manish Pundlik
Reader ,School of Computers & Electronics, IPS Academy, Indore
Kavita choudhary
,School of Computers & Electronics, IPS Academy, Indore

Abstract

The difficulties faced by text-based retrieval became more and more severe. The efficient management of the rapidly expanding visual information became an urgent problem. This need formed the driving force behind the emergence of content-based image retrieval techniques.

Many different approaches for content-based image retrieval have been proposed in the literature. Successful approaches consider not only simple features like color, but also take the structural relationship between objects into account. The retrieval process in order to generate perceptually and semantically more meaningful retrieval results. We present a new approach that significantly automates the examination process by relying on image analysis techniques. The general approach is to use previously identified content (e.g., contraband images) and to perform feature extraction, which captures mathematically the essential properties of the images. Based on this analysis, we build a feature set database that allows us to automatically scan a target machine for images that are similar to the ones in the database.

1. INTRODUCTION

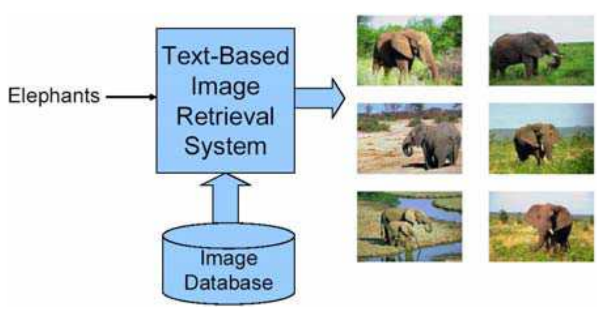
Content-based image retrieval, a technique which uses visual contents to search images from large scale image databases according to users' interests, has been an active and fast advancing research area since the 1990s. During the past decade, remarkable progress has been made in both theoretical research and system development e.g. Digital forensic investigations often require the examination of pictures found on the target media. Two typical tasks in that respect are the identification of contraband images and the identification of case-specific images, the presence of which can establish a fact or a logical link relevant to the investigation. The essential problem is that current forensic tools are ill-equipped to help the investigator given the scale of the task. To illustrate, we will collect thousands of image files on a randomly selected machine in our computing lab. Even if the investigator spends on average a fraction of a second on each image, it will still require several hours of routine, tedious work to browse through all of them. We will make the examiner's task even more difficult by removing any incentive for users to delete images. Thus, it is not unreasonable to expect that the hard drive of the average home user will contain hundreds of thousands of images. If we consider a target such as a web hosting service that can have tens of millions of images, the problem of examining all images becomes virtually intractable and investigators will need some means to narrow down the search space.

Similar problems in traditional forensics (e.g. fingerprint identification) have been tackled by building large reference databases that allow evidence from previous cases to be automatically searched. Clearly, a system capable of automatically identifying contraband images found on target media by cross referencing a database of known images could be of significant help to investigators. The problem, however, is that unlike other forensic artifacts, contraband images typically cannot be stored, even by law enforcement agencies, for future reference. Aside from the legal barriers, building a sizeable reference database to be used on a routine basis by numerous agencies would be a challenging task. From a technical point of view, the storage and bandwidth requirements would be staggering. Scalability would be difficult to achieve as replication and distribution of such highly sensitive material would have to be limited. Finally, a potential security breach at such a storage facility or misuse by authorized personnel can only be compared to a nuclear accident as far as the public outcry is concerned.

2. Content-Based Image Retrieval

2.1 Overview

Depending on the query formats, image retrieval algorithms roughly belong to two categories: text-based approaches and content-based methods (see Figure 1). The text-based approaches associate keywords with each stored image. These keywords are typically generated manually. Image retrieval then becomes a standard database management problem However; manual annotation for a large collection of images is not always available. Further, it may be difficult to describe image content with a small set of keywords. This motivates research on content-based image retrieval (CBIR), where retrieval of images is guided by providing a query image or a sketch generated by a user (e.g., a sketch of a horse).



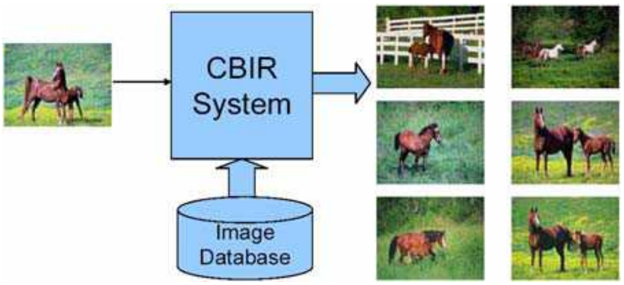


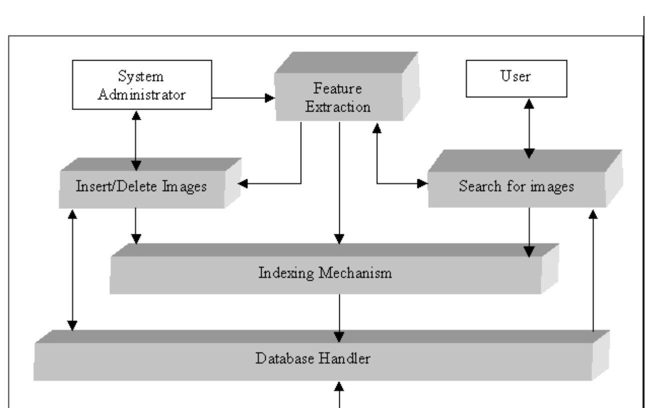
Figure 1: Scheme diagrams of a text-based image retrieval system (top) and a content-based image retrieval system (bottom).

In the past decade, many CBIR systems have been developed. Examples include the IBM QBIC System [FALO94], the MIT Photobook System [PENT96], the Berkeley Chabot [OGLE95] and Blobworld Systems [CARS02], the Virage System [GUPT97], Columbia's VisualSEEK and WebSEEK Systems [SMIT96], the PicHunter System [COX00], UCSB's NeTra System [MA97], UIUC's MARS System [MEHR97], the PicToSeek System [GEVE00], and Stanford's WBIIS [WANG98] and SIMPLIcity Systems [WANG01], to name just a few. From a computational perspective, a typical CBIR system views the query image and the images in the database as a collection of features, and ranks the relevance between the query and any matching image in proportion to a similarity measure calculated from the features. These features are typically extracted from shape, texture, intensity, or color properties of the query image and the images in the database. These features are image signatures and characterize the content of images, with the similarity measure quantifying the resemblance in content features between a pair of images.

Similarity comparison is an important issue in CBIR. In general, the comparison is performed either globally, using techniques such as histogram matching and color layout indexing, or locally, based on decomposed regions (objects). As a relatively mature method, histogram matching has been applied in many general-purpose image retrieval systems such as IBM QBIC, MIT Photobook, Virage System, and Columbia VisualSEEK and WebSEEK. A major drawback of the global histogram search lies in its sensitivity to intensity variations, color distortions, and cropping.

In a human visual system, although color and texture are fundamental aspects of visual perceptions, human discernment of certain visual contents could potentially be associated with interesting classes of objects, or semantic meanings of objects in the image. A region-based retrieval system segments images into regions (objects), and retrieves images based [2] on the similarity between regions. If image segmentation is ideal, it is relatively easy for the system to identify objects in the image and to match similar objects from different images.

2.2 Sample CBIR architecture



3. Methodology: Content-Based Image Retrieval

In content-based image retrieval the use of simple features likes color, shape or texture is not sufficient. Instead, the ultimate goal is to capture the content of an image via extracting the objects of the image. Usually images contain an inherent structure which may be hierarchical. An example can be seen in following figure. In the following, we describe two models for image representation and similarity measurement [2] which take structural as well as content features like color into account.

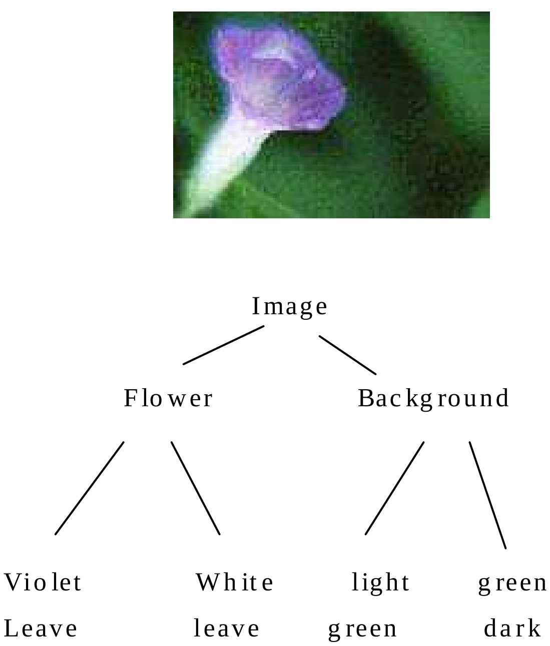


Fig. 3. An image and its inherent structure.

3.1 Image Representation as Containment Trees

One way to model images for content-based retrieval is the use of trees representing the structural and content information of the images. In this section, we describe, how the structure of images can be extracted automatically based on the color of its segments. Additionally we show how the similarity between two such trees can be measured.

Transforming an Image into a Containment Tree To utilize the inherent structure of images for content-based retrieval, we model them as so called containment trees. Containment trees model the hierarchical containment of image regions within others. To extract the containment tree of an image we first segment the image based on the colors of the regions using a region growing algorithm. The resulting segments are attributed with their color and size relative to the complete image. In a second step, the containment hierarchy is extracted from the set of segments by determining which regions are completely contained in other regions.

Measuring the distance between two Containment Trees To measure the similarity of containment trees, special similarity measures for attributed trees are necessary. A successful similarity measure for attributed trees [4,9] is the edit distance. Well known from string matching the edit distance is the minimal number of edit operations necessary to transform one tree into the other. The basic form allows two edit operations, i.e. the insertion and the deletion of a node. In the case of attributed nodes the change of a node label is introduced as a third basic operation. A great advantage of using the edit distance as a similarity measure is that along with the distance value, a mapping between the nodes in the two trees is provided in terms of the edit sequence. The mapping can be visualized and can serve as an explanation of the similarity distance to the user.

However, as the computation of the edit-distance is NP-complete, constrained edit distances like the degree-2 edit distance have been introduced. They were successfully applied to trees for web site analysis, structural similarity of XML documents, shape recognition or chemical substructure search.

Efficient Similarity Search for Containment Trees While yielding good results, the degree-2 edit distance is still computationally complex and, therefore, of limited benefit for searching or clustering in large databases. A filter and refinement architecture for the degree-2 edit[5] distance is presented to overcome this problem. A set of new filter methods for structural and for content-based information as well as ways to flexibly combine different filter criteria are presented.

3.2 Image Representation as Segmentation Graphs

Graphs are another way to model images for content-based similarity search. They were successfully used for shape retrieval, object recognition or face recognition. In this section, we describe a content-based image retrieval system based on graphs which are extracted from images in a similar way as the trees in the preceding section.

Transforming an Image into a Segmentation Graph To extract graphs from the images, they are segmented with a region growing technique and neighboring segments are connected by edges to represent the neighboring relationship. Each segment is assigned four attribute values, which are the size, the height and width of the bounding box and the color of the segment. The values of the first three attributes are expressed as a percentage relative to the image size, height and width in order to make the measure invariant to scaling.

Measuring the distance between two Segmentation Graphs Most known similarity measures for attributed graphs are either limited to a special type of graph or are computationally extremely complex, i.e. NP-complete. Therefore they are unsuitable for searching or clustering large collections.

.

the edge matching distance is a meaningful similarity measure for attributed graphs and that it enables efficient clustering of structured data.

Efficient Similarity Search for Segmentation Graphs. There is also a filter-refinement architecture and an accompanying set of filter methods presented to reduce [6] the number of necessary distance calculations during similarity search. We employ the same approach to ensure efficient query processing in our experiments.

References:

[1] Arbter, W. E. Snyder, H. Burkhardt, and G. Hirzinger, "Application of affine-invariant Fourier descriptors to recognition of 3D objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 640-647, 1990.

[2] E. M. Arkin, L.P. Chew, D..P. Huttenlocher, K. Kedem, and J.S.B. Mitchell, "An efficiently computable metric for comparing polygonal shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 3, pp. 209-226, 1991.

[3] J. Assfalg, A. D. Bimbo,J.A. Catalan, and J.S. Jin, "Dimension reduction of texture features for image retrieval using hybrid associative neural networks," *IEEE International Conference on Multimedia and Expo*, Vol.2, pp.1211 -1214, 2000

[4]Levenshtein, V.: Binary codes capable of correcting deletions, insertions and reversals. Soviet Physics-Doklady **10** (1966) 707–710

[5]Kailing, K., Kriegel, H.P., Sch¨onauer, S., Seidl, T.: Efficient similarity search for hierachical data in large databases. In: Proc. 9th Int. Conf. on Extending Database Technology (EDBT 2004). (2004) 676–693

[6] Michael Swain and Dana Ballard. Color indexing. *International Journal of Computer Vision*,7(1):11{32, 1991.

[7]R.C. Veltkamp and M. Tanase. Content-based image retrieval systems:A survey. Technical report, Department of Computer Science, Utrecht University, October 2000

[8]S. K. Chang, and A. Hsu, "Image information systems: where do we go from here?" *IEEE Trans. on Knowledge and Data Engineering*, Vol.5, No.5, pp. 431-442, Oct.1992.

[9]Wagner, R.A., Fisher, M.J.: The string-to-string correction problem. Journal of the ACM **21** (1974) 168–173