

Content Based Image Retrieval System- A State of Art Approach

Shaheen Fatima

Associate Professor of Electronics, Government College, Kalaburagi, Karnataka India.

Abstract: The emergence of multimedia, the availability of large image archives, and the rapid growth of the digital image database have attracted significant research efforts in providing tools for effective retrieval and management of visual data. A major approach directed towards achieving this goal is to use visual contents of the image data to segment, index and retrieve relevant images from the image database. The commonest approaches use the so-called Content-Based Image Retrieval (CBIR) systems. The goal of these systems is to operate on collections of images and, in response to visual queries, extract similar images. The application potential of CBIR for fast and effective retrieval of images is enormous, expanding the use of computer technology to a management tool.

Keywords: Image retrieval; feature extraction; similarity measures; Content-based image retrieval.

Introduction

In the early 1990s, as a result of advances in the Internet and new digital image sensor technologies, the volume of digital images produced by scientific, educational, entertainment, medical, industrial, and other applications available to users increased dramatically. Storage of such image data is relatively straightforward, but an accessing and searching image database is intrinsically harder than their textual counterparts. A major approach directed towards achieving this goal is to use low-level visual features of the image data to segment, index and retrieve relevant images from the image database. Recent CBIR systems based on features like color, shape, texture, spatial layout, object motion, etc., are cited in [1, 2]. Of all the visual features, color is the most dominant and distinguishing one in almost all applications.

Content-Based Image Retrieval

Content Based Image Retrieval describes the process of retrieving desired images from a large image database on the basis of features (such as color, texture and shape) that can be extracted from the images themselves. In typical content based image retrieval systems (Fig. 1), the visual contents of the images in the database are extracted and described by multi-dimensional feature vectors. To retrieve images, users provide the retrieval system with example images or sketched figures. The system then changes these examples into its internal representation of feature vectors. The similarities /distances between the feature vectors of the query example or sketch and those of the images in the database are then calculated and retrieval is performed with the aid of an indexing scheme. The indexing scheme provides an efficient way to search for the image database.

CBIR vs Text-Based Image Retrieval

Depending on the query formats, image retrieval algorithms roughly belong to two categories: text-based approaches and content-based methods (Fig. 1). The text-based approaches associate keywords with each stored image. These keywords are typically generated manually. There are two main disadvantages in this approach.

One is that it requires a huge amount of human labor in the manual annotation especially when the image collection is large. The other one is that it is hard to precisely annotate the rich content of an image by humans due to perception subjectivity [3, 4]. 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). In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps:

- **Feature Extraction:** The first step in the process is to extract image features to a distinguishable extent.
- **Matching:** The second step involves matching these features to yield a result that is visually similar.



Fig1. Schematic diagrams of a text-based image retrieval system (left) and a content-based image retrieval system (right)

Block diagram of CBIR system

The general image retrieval system usually consists of three main modules, i.e., input module, query module, and retrieval module [5]. The block diagram of the retrieval process is shown in Fig. 2. In the query module, a query image is inputted, and its feature vector is extracted. While in the input module, the feature vector is extracted from the input image and the feature vector of each image in the image database is stored in the feature database. During the retrieval process, the feature vector of the query image is compared with each vector of the images in the database, and the corresponding target (similar) images are outputted.

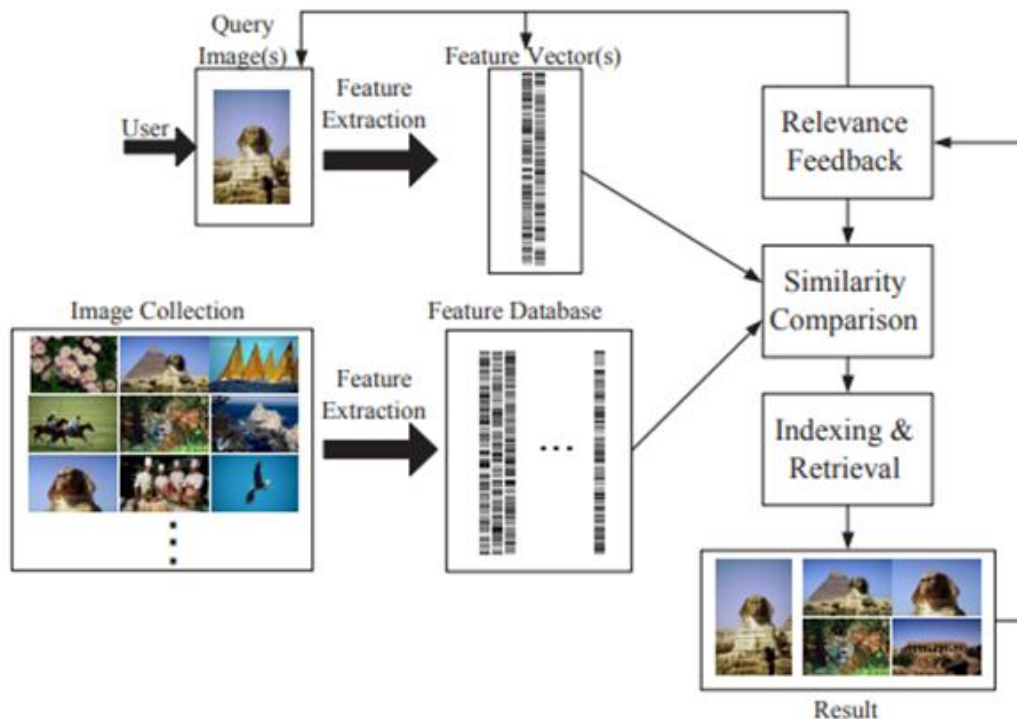


Fig.2. Block diagram of content-based image retrieval system

Query by Image Contents

For content-based image retrieval, user interaction with the retrieval system is crucial since flexible formation and modification of queries can only be obtained by involving the user in the retrieval procedure. User interfaces in image retrieval systems typically consist of a query formulation part and a result presentation part. Technique of query by image contents is used when images are retrieved on the basis of low-level features such as to color, texture and shape features. There are two most common methods in query by image contents, namely query by visual example and query by image specification.

□ Query by visual: With query by example, the user searches with a query image (supplied by the user or chosen from a random set), and the software finds images similar to it based on various low-level criteria. A number of retrieved images will be displayed in order of similarity. For example QBIC system uses this method of retrieval [6].

□ Query by image specification: A user can directly set the values of specific options. For example, a system allows users to specify the proportions of colors desired (e.g. “80% red, 20% blue”). Similarly, the system VisualSEEK allows users to specify multiple regions in the query, e.g. “find images which contain a region of blue sky and ocean on the top and bottom respectively” [7].

Feature Extraction Techniques

Feature extraction is the basis of content-based image retrieval. CBIR operates on a principle of retrieving stored images from a database by comparing features automatically extracted from the images themselves. The commonest features used are mathematical measures of color, texture or shape.

Color

Color is the most extensively used visual content for image retrieval. It is relatively robust to background complication and independent of and orientation image size. The most common primary colors in computing are red, green and blue (e.g. colors used in a monitor). Usually colors are defined in three dimensional color spaces. These could either be RGB (Red, Green and Blue), HSV (Hue, Saturation and Value) or HSB (Hue, Saturation and Brightness). Most image formats such as *jpg*, *bmp*, *gif* use the RGB color space to store information. The primary reason for this is because it retains compatibility with computer displays. However, the RGB space has the major drawback in that it is not perceptually uniform. In image retrieval systems, color histogram is the most commonly used feature representation. The color histogram describes the proportion of pixels of each color in an image with simple and computationally effective manner. Color histogram is obtained by quantizing image colors into discrete levels and then counting the number of times each discrete color occurs in the image. At search time, the user can specify the desired proportion of each color (65% red and 35% green, for example), or submit an example image from which a color histogram is calculated. During retrieval, the histogram of a query image is compared with the histogram of all the images in the database. Swain and Ballard [8] proposed the histogram intersection as the histogram intersection as the similarity measure for color histogram. Besides color histograms, several other color feature extraction techniques have been applied in image retrieval. These include color sets and color moments. To overcome the quantization effects in the color histogram, Stricker and Orengo [9] proposed the color moments approach. The basic idea behind this approach behind this approach is that any color distribution can be characterized by its moments.

Texture

Texture refers to the visual patterns with properties of homogeneity that do not result from the presence of a single color or intensity [10] [11]. It is that innate property of all surfaces that describes visual patterns such as; clouds, leaves, bricks, fabric, etc. It contains information about the structural arrangement of surfaces and their relationship to the surrounding environment. Texture properties include coarseness, contrast, directionality, regularity and roughness. The texture analysis methods can be divided into statistical, structural and spectral approaches.

□ Statistical techniques characterize texture using the statistical properties of the gray levels of the pixels comprising an image. Normally, in images, there is periodic occurrence of certain gray levels. The spatial distribution of gray levels is calculated.

□ *Structural* techniques characterize texture as being composed of texels (texture elements). These texels are arranged regularly on a surface according to some specific arrangement rules.

In *spectral* approach, texture description is done by Fourier transform of an image and then group the transformed data in a way that it gives some set of measurements.

Statistical techniques are most important for texture classification because it is these techniques that result in computing texture properties. Some of the statistical representations of texture are co-occurrence matrix, tamura texture and wavelet transform. Co-occurrence matrix was originally proposed by R.M. Haralick [12]. This technique constructs a co-occurrence matrix on the basis of orientation and the distance between the pixels. Then meaningful statistics are extracted from matrix as the texture representation

Shape

Shape is an important criterion for matching objects based on their profile and physical structure. Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out. Shape features can represent spatial information that is not represented by color or texture. It contains all the geometrical information of an object in the image which does not change generally change even when orientation or location of the object are changed. Shape is a key attribute of segmented image region, and its efficient representation plays an important role in image retrieval. Some simple shape features are the perimeter, area, eccentricity, symmetry, etc. To describe shape features, it is essential to segment the image to obtain regions or objects. In general, shape representations can be divided into two categories, boundary-based and region-based.

Similarity Measures

The objective of a CBIR query is to efficiently search and retrieve images from a database that are similar to the query image specified by a user. Exactness, however, is a precise concept. The most common method for comparing two images in content based image retrieval (typically an example image and an image from the database) is using an image distance measure. An image distance measure compares the similarity of two images in various dimensions such as color, texture, shape, and others. For example a distance of 0 signifies an exact match with the query, with respect to the dimensions that were considered. A value greater than 0 indicates various degrees of similarities between the images.

Semantic Gap

Low-level features such as color, texture and shape can be easily extracted from images. However, they do not coincide with the high-level concepts such as emotions, objects, events, or activities as conveyed by an image. This is referred to the so-called semantic gap. It shows the difference between high-level concepts and the relatively limited descriptive power of low-level visual features. Normally humans recognize objects by using prior knowledge on different objects. This knowledge can be based on personal preferences, previous experiences of similar situations, cultural context, etc. This kind of information is hard to incorporate in CBIR systems. This means that such CBIR systems can only retrieve images using low-level features and it is

not possible to find images by putting query like “give me images containing grass and tiger” [13].

Relevance Feedback

Relevance feedback has been proposed as an important technique to boost the retrieval performance in content based image retrieval. Generally, the CBIR system cannot return all the relevant information accurately in its first response, and so the system has to ask users to provide positive (relevant) and negative (irrelevant) examples as feedback in an iterative manner. In principle, relevance feedback is based on learning an ‘optimal’ correspondence between the high-level semantic concepts and the low-level internal features in each query session. Initially developed in document, relevance feedback was transformed and introduced into content-based multimedia retrieval, mainly content-based image retrieval (CBIR), during early and mid 1990s. Relevance feedback provides a way to bridge the gap between semantic searching and low level features. A detailed review of relevance feedback techniques in content-based image retrieval is published by Zhou and Huang [14].

Evaluation Strategies

Performance evaluation has been a challenging issue in the field of content-based retrieval, primarily because of the difficulty associated with calculating quantitative measures to evaluate the quality of retrieval. There are several feature extraction techniques and retrieval techniques which have been developed, so it is better to compare and select the ones which give superior performance. For any information retrieval system, a strategy for evaluation involves determining the following aspects.

- The dataset should be large enough for the evaluation to be statistically significant. Also, the dataset should be general enough to cover a large range of semantics from a human point-of-view.
- The evaluation criteria should try to model human requirements from a population perspective. The evaluation metrics should depend on the objective of the CBIR system. Hence, the performance criteria for algorithms may be different for different purposes.
- The performance evaluation should be comprehensive. The comprehensiveness entails that all the targets of the CBIR system are considered.
- The metrics definition should be quantitative, normative, objective, and compatible with the human vision Evaluation

Performance Metrics

CBIR is essentially an information retrieval problem. Two of the most popular evaluation measures are the precision and recall [8]. The precision measures the proportion of the total images retrieved which are relevant to the query.

precision = The recall measure is defined as the fraction of the all relevant images.

recall = High precision means that less irrelevant images are returned or more relevant images are retrieved, while high recall indicates that few relevant images are missed. Another way of

presenting the performance of the system is by plotting precision and recall graph, in which precision values are plotted against values of recall. These graphs give a clear idea about the system performance. Some systems use average precision which provides a single value to compare the retrieval performance. In average precision, first precision is calculated for each retrieved relevant image. Then average precision measure is obtained by averaging these precisions over the total number of relevant objects.

Research Challenges

The implementation of CBIR systems raises several research challenges. Some of these are:

- Understanding image users' needs and information-seeking behavior
- Investigating new user interfaces for annotating, browsing, and searching based on image content
- Creating formalisms to describe image content descriptions and related services
- New tools for marking/annotating images (and their regions)
- Creating better semantically enriched descriptions.
- Providing compact storage for large image databases
- Matching query and stored images in a way that reflects human similarity judgments
- Efficiently accessing stored images by content
- Providing usable human interfaces to CBIR systems
- Develop tools that automatically extract semantic features from images
- Incorporate classification strategies into the image retrieval process.

Conclusion and Future Scope of Work

The area of content-based image retrieval is a hybrid research area that requires knowledge of both computer vision and of database systems. The application of information theory to image interpretation and CBIR poses many questions for further exploration. The technology is exciting but immature, and few operational image archives have yet shown any serious interest in adoption. The field appears to be generating interesting and valid results, even though it has so far led to few commercial applications. Agencies concerned with technology transfer or dissemination of best practice in fields which could potentially benefit from CBIR (including management of image collections and drawing archives, electronic publishing and multimedia content creation) should consider sponsoring programmes to raise awareness of CBIR technology among leading practitioners in these fields. Interactive performance evaluations including several levels of feedback and user interaction need to be developed. The need for standardized evaluation measures is needed, since several measures are slight variations of the same definition. A frequently updated shared image database and the regular comparison of system performances would be of great benefit to the CBIR community.

References

- [1] V.N.Gudivada and V.V.Raghavan.: Special issue on content-based image retrieval systems - guest eds. IEEE Computer. 28(9) (1995) 18-22.
- [2] M.De Marsicoi, L.Cinque, and S.Levialdi.: Indexing pictorial documents by their content: a survey of current techniques. Image and Vision Computing. 15(2) (1997) 119-141.
- [3] Yong Rui and Thomas S. Huang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues," Journal of Visual Communication and Image Representation, 1999.
- [4] Wei-Ying Ma and HongJiang Zhang, "Content-Based Image Indexing and Retrieval," Handbook of Multimedia Computing, CRC Press, 1999.
- [5] Niblack, W., R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Pektovic, P. Yanker, C. Faloutsos and G. Taubin, The QBIC project: Querying images by content using color, texture, and shape, in Storage and Retrieval for Image and Video Databases, Proc. SPIE 1908, pp.173-187, 1993.
- [6] Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Huang, Q., Dom, B., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., And Yanker, P. 1995. Query by image and video content: The QBIC system. IEEE Comput. 28, 9, 23–32.
- [7] Smith, J. And Chang, S.-F. 1997. Visualseek: A fully automated content-based image query system. In Proceedings of the ACM International Conference on Multimedia
- [8] Mukherjea, S., Hirata, K., And Hara, Y. 1999. Amore: A World Wide Web image retrieval engine. In Proceedings of the International World Wide Web Conference (WWW).
- [9] M. Stricker and M. Orengo. Similarity of color images. In In SPIE Conference on Storage and Retrieval for Image and VideoDatabases III, volume 2420, pages 381392, Feb. 1995.
- [10] J.R. Smith and S.F. Chang, "Automated binary texture feature sets for image retrieval," in Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc., (Atlanta), GA, 1996
- [11] Sharmin Siddique, "A Wavelet Based Technique for Analysis and Classification of Texture Images," Carleton University, Ottawa, Canada, Proj. Rep. 70.593, April 2002
- [12] R. M. Haralick, K. Shanmugam, I. Dinstein, "Texture features for image classification." IEEE Trans. System Man Cybernat, 8 (6), 1973, pp. 610-621
- [13] Pentland, A., Picard, R., And Sclaroff, S. 1994. Photobook: Tools for content-based manipulation of image databases. In Proceedings the Conference on Storage and Retrieval for Image and Video Database II, SPIE, San Jose, CA.
- [14] Zhou, X.S. and Huang, T.S. (2003), "Relevance Feedback in Image Retrieval: A Comprehensive Review," ACM Multimedia Systems Journal, pp. 536–544.