

Text-based and Content-based, Image Retrievals: A Survey

Shaheen Fatima

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

Abstract: Image retrieval from databases or from the Internet needs an efficient and effective technique due to the explosive growth of digital images. Image retrieval is considered as an area of extensive research, especially in content based image retrieval (CBIR). CBIR retrieves similar images from large image database based on image features, which has been a very active research area recently. The content, that can be derived from image such as color, texture, shape...etc., are called features. This paper will present a survey and discuss the current literature of different types of image retrieval (IR) systems. An overview of the important techniques in image retrieval will be discussed. Finally, some urgent challenges in IR, that have been raised recently, will be presented as well as possible directions for future research.

Keywords: Content-based image retrieval, Text-based image retrieval.

I. INTRODUCTION

In recent years, collections of digital images are created and increased rapidly. In many areas of academia, commerce, government, medicine, and Internet, a huge amount of information is out there. However, we cannot access or make use of this information unless it is organized to allow efficient browsing, searching, and retrieval. One of the main problems is the difficulty of locating a desired image in a large and varied collection. While it is perfectly feasible to identify a desired image from a small collection simply by browsing, more effective techniques are needed with collections containing thousands of items. Image retrieval attracts interest among researchers in the fields of image processing, multimedia, digital libraries, remote sensing, astronomy, database applications, and other related areas [1]. Image retrieval has been a very active research area since 1970s, with the thrust of two major research communities: database management and computer vision[2]. Therefore, image retrieval can be defined as the task of searching for images in an image database.

A. Text-Based Image Retrieval (TBIR)

TBIR can be traced back to the late 1970s. A very popular framework of TBIR was first annotated the images by text and then used text-based database management systems to perform image retrieval[2]. TBIR is used to manually annotate the image in the database with annotations, keywords, or descriptions. This process is used to describe both image contents and other metadata of the image such as: image file name, image and image format, image size, and image dimensions. Then, the user formulates textual or numeric queries to retrieve all images that are satisfying some of the criteria based on these annotations, as shown in Fig. 1. However, there are some drawbacks in TBIR [3]. The first drawback is that the most descriptive annotations must usually be entered manually. Manually annotation for a large image database is

impractical. The second drawback is that the most images are very rich in its content and has more details. The annotator may give different descriptions to images with similar visual contents. Also, textual annotations are language-dependent[4].

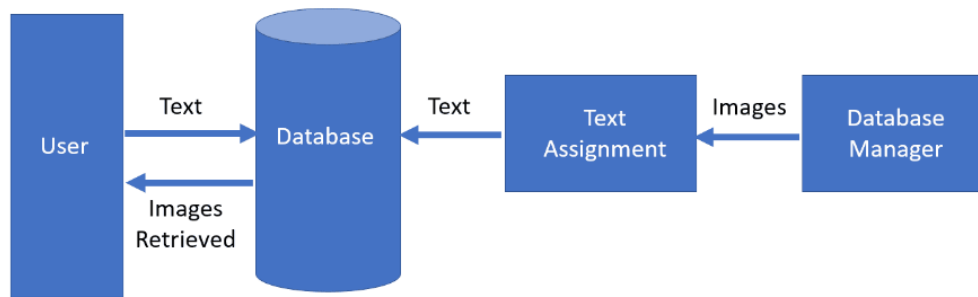


Figure1. A typical Text-Based Image Retrieval system

B. Content-Based Image Retrieval (CBIR)

CBIR is considered as an active and fast advancing research area. It is also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR)[5]. The term CBIR seems to have originated with the work of Kato [6] for the automatic retrieval of the images from a database based on the color and the shape. After that, the CBIR term has widely been used to describe the desired images retrieving process from a large collection of database based on image visual contents, normally called as features (color, shape, texture...etc.)[7]. In the early 1990s, as a result of the advances in the Internet and techniques of digital image production, a huge amount of digital images are produced in sciences, education, medicine, industry, and other fields available to the users that increased dramatically and make the drawbacks faced by TBIR became more and more tough. This needs formed the driving force behind the emergence of CBIR techniques [8]. The advances in CBIR researches mainly contributed by the computer vision community [2]. The used techniques and algorithms originate from many fields such as object recognition and signal processing [7]. However, in the last decade CBIR has received much attention which is motivated by the need to efficiently handle the rapidly growing amount of multimedia data[9]. It covers versatile areas, such as image segmentation, image feature extraction, representation, mapping of features to semantics [10]. Research and development issues in CBIR cover a range of topics, most important are: understanding image users' needs and information-seeking behavior, identification of suitable ways of describing image content, extracting such features from raw images and matching query and stored images in a way that reflects human similarity.

As illustrated in Fig. 2, a typical CBIR is divided into off-line feature extraction and on-line image retrieval. In off-line stage, the system automatically extracts features of each image in the database and stores them in a features database (features of the images are extracted and represented with feature vectors). In on-line stage, user input an image query to the system. The features of the query image are extracted and represented. The similarity was measured between the feature vector of the query image and the feature vectors of the images in the database. Then,

the retrieval process is performed by applying an indexing scheme to provide an efficient way of searching the image database. Finally, the system returns the images that are most similar to the query image[11].

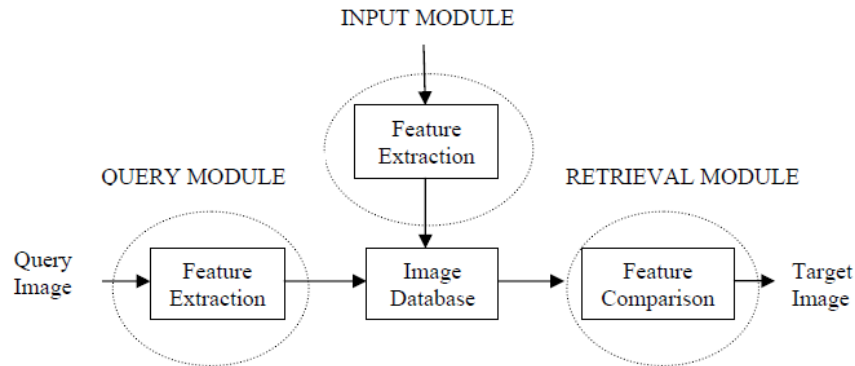


Fig 2. A typical Content-Based Image Retrieval system.

Another technique for image retrieval used to integrate text and image content to enhance the retrieval accuracy. Both the text and content-based techniques have their own characteristics, advantages, and disadvantages. By combining them, parts of their disadvantages can be overcome [12].

II. COMMON FEATURES FOR IMAGE RETRIEVAL

A feature is defined as capturing a certain visual property of an image[13]. In general, image features can be either global or local[14]. The global features describe the visual content of the entire image, whereas local features describe the regions or objects (i.e. a small group of pixels) of the image content. The advantage of global extraction is its high speed for both extracting features and computing similarity. However, global features are often too rigid to represent an image. Specifically, they can be oversensitive to location and hence fail to identify important visual characteristics[13]. Local-feature approaches provide a slightly better retrieval effectiveness than global features[15]. They represent images with multiple points in a feature space in contrast to single point global feature representations. While local approaches provide more robust information, they are more expensive computationally due to the high dimensionality of their feature spaces and usually need nearest neighbors approximation to perform points matching[16,17]. Several important features that can be used in IR will be elucidated in the next subsections.

A. Color Features

The color has widely been used in IR systems, because of its easy and fast computation[18,19]. Color is also an intuitive feature and plays an important role in image matching. Most IR systems use color space, histogram, moments, color coherence vector, and dominant color descriptor represent color. The color histogram is one of the most commonly used idea to use histogram for retrieval comes from Swain and Ballard [19], who realized the power to identify an object using color is much larger than that of a gray scale[7,20]. Although the global color feature is simple to calculate and can provide reasonable discriminating power in image retrieval. It tends to give too many false positives when the image collection is large. Many research results suggested that using color layout is a better solution for image retrieval. To extend the global color feature to a local one, a natural approach is to divide the whole image into sub-blocks and extract color features from each of the sub-blocks. The advantage of this approach is its accuracy while the disadvantage is the general difficult problem of reliable image segmentation[21,22].

B. Texture Features

Texture is a property that represents the surface and structure of an image. Texture can be defined as a regular repetition of an element or pattern on a surface. Image textures are complex visual patterns composed of entities or regions with sub-patterns with the characteristics of brightness, color, shape, size...etc.[10].The commonly known texture descriptors are Wavelet Transform[23],Gabor-filter[24],and Tamura features[25].

C. Shape Features

Shape can generally be defined as the description of an object regardless of its position, orientation, and size. Therefore, shape features should be invariant to translation, rotation, and scale for an effective IR. In the direction of using shape as an image feature, it is necessary to determine object or region boundaries in the image and this is a challenge[26]. Compared with color and texture features, shape features are usually described after images have been segmented into regions or objects. Since robust and accurate image segmentation is difficult to achieve, the use of shape features for image retrieval has been limited to special applications where objects or regions are readily available. In general, the shape representations can be divided into two categories, boundary-based that uses only the outer boundary of the shape and region-based that uses the entire shape region[27].

III. COMMON TRADITIONAL AND CURRENT RESEARCH IN IMAGE RETRIEVAL

There have been numerous previous efforts in image retrieval. This section gives a brief overview of some previous work targeting the TBIR, CBIR.

A. TBIR Systems

The text-based approach is a traditional simple keyword based search. The images are indexed according to the content, like the caption of the image; filename, title of the web page, and alternate tag...etc. and stored in the database. Processing a user query could involve a stop word

removal and stemming. Some of the keyword based image retrieval approaches are bag of words. Image retrieval is then shifted to standard database management capability combined with information retrieval techniques. Some commercial image search engines, such as Google Image Search and Yahoo Image Search, are keyword-based image retrieval systems [28].

Liet al.[29] proposed a TBIR approach that can effectively exploit loosely labeled Web images to learn robust SVM classifiers.

B. CBIR Systems

This section discusses CBIR from three different aspects: common features, local features, and hybrid techniques. Several methods are proposed for images retrieval based on similarity between common features, such as color, shape...etc. This is done by retrieving the stored images from a collection by comparing the stored features in the database with automatically extracted features of the processed image. The use of local features in image retrieval have been proposed to resist many problems, such as the illumination changes, rotation, viewpoint variations...etc. The interest points are defined as the salient image patches that contain rich local information about an image [30].

IV. IR CHALLENGES AND RESEARCH DIRECTIONS

Bridging the semantic gap for image retrieval still considered a big challenge. Even though there are a lot of efforts and works on image retrieval research, but it is not enough to provide satisfactory performance. However, there are still some spaces, which need to be improved besides the challenges that is associated with mapping low level to high-level concepts. Also overcome of the semantic gap in the broad domain database is complex because the images in broad domains can be described using various concepts. There are need to see better support for the image retrieval based semantic concept with a focus on the retrieval by abstract attributes, involving a significant amount of high-level reasoning about the meaning and purpose of the objects. In addition, the extracted semantic features should be applied for any kind of image collection. Moreover, there is need to effective ways retrieve of similar images that are conform to human perception and without human interference.

V. CONCLUSION

This paper attempted to provide an overview of the most common techniques of different types of image retrieval systems. Most systems used low-level features, few systems used semantic feature. Global features fail to identify important visual characteristics of images but it's very efficient in computation and storage due to its compact representation. From another perspective, local features that can be extracted from images handle partial image matching or searching for images that contain the same object or same scene with different viewpoints, different scale, changes in illumination, oclusions,,...etc. Therefore, local features can identify important visual characteristics of images but it is more expensive computationally. As a survey, it is very difficult to include each and every aspect of all works. However, this paper focused to give overview of the most common traditional and modern types of IR approaches.

REFERENCES

- [1] H.W. Hui, D. Mohamad, and N.A. Ismail, "Approaches, challenges and future direction of image retrieval," *Journal of Computing*, vol. 2, 2010, pp. 193–199.
- [2] Y. Rui, T.S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *Journal of visual communication and image representation*, vol. 10, 1999, pp. 39–62.
- [3] S.-K. Chang and A. Hsu, "Image information systems: where do we go from here?," *IEEE transactions on Knowledge and Data Engineering*, vol. 4, 1992, pp. 431–442.
- [4] H. Tamura and N. Yokoya, "Image database systems: A survey," *Pattern recognition*, vol. 17, 1984, pp. 29–43.
- [5] G. Mailaivasan Parthiban Karthikram, "Tag based image retrieval (TBIR) using automatic image annotation," *International Journal of Research in Engineering and Technology*, vol. 03, 2014.
- [6] T. Kato, "Database architecture for content-based image retrieval," *SPIE/IS&T 1992 Symposium on Electronic Imaging: Science and Technology*, International Society for Optics and Photonics, 1992, pp. 112–123.
- [7] M. Singha and K. Hemachandran, "Content based image retrieval using color and texture," *Signal & Image Processing: An International Journal (SIPIJ)*, vol. 3, 2012, pp. 39–57.
- [8] S. Mitra and T. Acharya, *Data mining: multimedia, soft computing, and bioinformatics*, John Wiley & Sons, 2005.
- [9] N. Chauhan and M. Goyani, "Enhanced Multistage Content Based Image Retrieval," *IJCSMC*, vol. 2, Issue. 5, 2013, pp. 175–179.
- [10] N. Chauhan and M. Goyani, "Enhanced Multistage Content Based Image Retrieval," *IJCSMC*, vol. 2, Issue. 5, 2013, pp. 175–179.
- [11] H. Zhang H. Dagan F. Long and D. Fen, *Fundamentals of content based image retrieval*, in D. Feng, W. Siu, H. Zhang (Eds.): "Multimedia Information Retrieval and Management. Technological Fundamentals and Applications., Springer-Verlag, .
- [12] B. Dinakaran, J. Annapurna, and C.A. Kumar, "Interactive image retrieval using text and image content," *Cybern Inf Tech*, vol. 10, 2010, pp. 20–30.
- [13] R. Datta, D. Joshi, J. Li, and J.Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys (CSUR)*, vol. 40, 2008, p. 5.
- [14] A. Halawani, A. Teynor, L. Setia, G. Brunner, and H. Burkhardt, "Fundamentals and Applications of Image Retrieval: An Overview.," *Datenbank-Spektrum*, vol. 18, 2006, pp. 14–23.
- [15] M. Aly, P. Welinder, M. Munich, and P. Perona, "Automatic discovery of image families: Global vs. local features," *Image Processing (ICIP), 2009 16th IEEE International Conference on, IEEE, 2009*, pp. 777–780.
- [16] A. Arampatzis, K. Zagoris, and S.A. Chatzichristofis, "Dynamic two-stage image retrieval from large multimedia databases," *Information Processing & Management*, vol. 49, 2013, pp. 274–285.
- [17] A. Popescu, P.-A. Moëllic, I. Kanellos, and R. Landais, "Lightweight web image reranking," *Proceedings of the 17th ACM international conference on Multimedia*, ACM, 2009, pp. 657–660.

- [18] Y.-K. Chan and C.-Y. Chen, "Image retrieval system based on color-complexity and color-spatial features," *Journal of Systems and Software*, vol. 71, 2004, pp. 65–70.
- [19] J.M. Fuertes, M. Lucena, N. Perez de la Blanca, and J. Chamorro-Martínez, "A scheme of colour image retrieval from databases," *Pattern Recognition Letters*, vol. 22, 2001, pp. 323–337.
- [20] M.J. Swain and D.H. Ballard, "Color indexing," *International journal of computer vision*, vol. 7, 1991, pp. 11–32.
- [21] T.S. Chua, K.-L. Tan, and B.C. Ooi, "Fast signature-based color-spatial image retrieval," *Multimedia Computing and Systems' 97. Proceedings.*, IEEE International Conference on, IEEE, 1997, pp. 362–369.
- [22] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, "Efficient and effective querying by image content," *Journal of intelligent information systems*, vol. 3, 1994, pp. 231–262
- [23] J.R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases," *Image Processing, 1994. Proceedings. ICIP-94.*, IEEE International Conference, IEEE, 1994, pp. 407–411.
- [24] P. Wu, B. Manjunath, S. Newsam, and H. Shin, "A texture descriptor for browsing and similarity retrieval," *Signal Processing: Image Communication*, vol. 16, 2000, pp. 33–43.
- [25] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 8, 1978, pp. 460–473.
- [26] D. Zhang and G. Lu, "Shape-based image retrieval using generic Fourier descriptor," *Signal Processing: Image Communication*, vol. 17, 2002, pp. 825–848.
- [27] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern recognition*, vol. 37, 2004, pp. 1–19.