

A Survey on Text and Content Based Image Retrieval System for Image Mining

T. Karthikeyan¹, P. Manikandaprabhu², S. Nithya³

¹ Associate Professor, ^{2,3} Research Scholar,

Department of Computer Science, PSG College of Arts and Science, Coimbatore, Tamilnadu, India.

Abstract— In this paper, we make a survey about text and content based image retrieval system. Image retrieval is performed by matching the features of a query image with those in the image database. It can be classified as text-based and content-based. The text-based Image retrieval applies traditional text retrieval techniques to image annotations. The content-based Image retrieval apply image processing techniques to first extract image features and then retrieve relevant images based on the match of these features. Feature extraction is the process of extracting image features to a distinguishable extent to extract the image content by using colors, textures or shapes based. Similarity measures are used to determine how similar or dissimilar in the given query image and the image database collections.

Keywords— Image retrieval, text-based Image retrieval, content-based Image retrieval, Feature extraction, colors, shapes, textures and similarity measures.

I. INTRODUCTION

Digital images are composed of pixels. Each pixel represents the colour at a single point of the image. Rectangular array of pixels are called as a bitmap or a digital image. Advance development of image procurement and storage technology have lead to marvellous development in very huge and detailed image databases [1]. A massive volume of image data such as digital photographs, medical images and satellite images are generated every day [2].

Image mining can automatically extract meaningful information from a huge of image data are increasingly in demand. It is an interdisciplinary venture that essentially draws upon expertise in artificial intelligence, computer vision, content based image retrieval, database, data mining, digital image processing and machine learning.

Image mining frameworks [3] are grouped into two broad categories: function-driven and information-driven. The problem of image mining combines the areas of content-based image retrieval, data mining, image understanding and databases. Image mining techniques include image retrieval, image classification, image clustering, image segmentation, object recognition and association rule mining.

Image Retrieval is performed by matching the features of a query image with those in the image database. The collection of images in the web are growing larger and becoming more diverse. Retrieving images from such large collections is a

challenging problem. The research communities study about image retrieval from various angles are text based and content based. The text based Image retrieval is used for traditional text retrieval techniques to image annotations. The content-based Image retrieval apply image processing techniques to first extract image features and then retrieve relevant images based on the match of these features.

The rest of this paper is organized as follows. Section 2 discusses about the related work of image retrieval. Section 3 and 4 gives a text and content based image retrieval. Section 5 discussed the conclusion.

II. RELATED WORK

Digital images are currently widely used in medicine, fashion, architecture, face recognition, finger print recognition and bio-metrics etc. Recently, Digital image collections are rapidly increased very huge level. That image contains a huge amount of information. Conversely, we cannot make sure that information is useful unless it is implemented so we need sufficient browsing, searching, and retrieving the images.

Retrieving image has become a very dynamic research area. Two major research communities such as database management and computer vision have study image retrieval from various ways such as text based and content based. Late 1970s, the text-based image retrieval had been traced back. A very popular framework of image retrieval was to annotate the images by keyword and they used text based database management system for operating image retrieval. Two broad surveys on this topic are [4, 5]. Emergence of large-scale image collections in the early 1990s, the major difficulties are manual image annotation is also accurate. To avoid this situation, content-based image retrieval was improved. It means, instead of using text based key words, images should be defined by their visual contents as colour and texture. Many techniques in this research area have been developed for many image retrieving systems as research and commercial, have been built. It has established a general framework of image retrieval. In this paper we will focus our effort mainly to the content-based image retrieval.

Many content-based image retrieval systems have been recently proposed: Chabot[6], MARS [7], Netra [8], Photobook [9], QBIC [10], Surfimage[12], SWIM [13], Virage [14], Visualeek [15] and WebSeek[16]. These systems follow the paradigm of representing images using a set of

attributes, such as color, texture and shape, which are archived along with the images.

III. TEXT-BASED IMAGE RETRIEVAL

Text-based image retrieval [4, 5] can be based on annotations that were manually added for disclosing the images (keywords, descriptions), or on collateral text that is ‘accidentally’ available with an image (captions, subtitles, nearby text). It applies traditional text retrieval techniques to image annotations or descriptions. Most of the image retrieval systems are text-based, but images frequently have little or no accompanying textual information.

Keywords are words or phrases that are described content. They can be used as metadata to describe images, text documents, database records, and Web pages. Assign keywords for an image allows one to retrieve, index, organize and understand large collections of image data. Keywords are used on the Web in two different ways: i) Keywords as a search terms for search engines ii) Keywords that identify the content of the website. An annotation is metadata attached to text, image, or other data. It refers to a specific part of the original data or image. Keyword annotation is the traditional text based image retrieval paradigm. In this approach, the images are first annotated manually by keywords. They can then be retrieved by their corresponding annotations. As the size of image repositories increases, the keyword annotation approach becomes infeasible. Text-based Image Retrieval has some limitations.

A. Limitations of Text-based Image Retrieval

- The task of visualizing image content is highly subjective.
- Accompany the relevant search results; it could be a large number of irrelevant search results which may be the precision of the text based search can be low.
- In many times, a few words cannot accurately describe the image content, and many words have multiple meanings.
- The textual descriptions provided by an annotator should be different from the other user. A picture defines various things for different people. It can also mean different things to the same person at different time.
- It must be a variety of inconsistencies for user text queries and image descriptions.

IV. CONTENT-BASED IMAGE RETRIEVAL

Problems with text-based retrieval, we will use the content-based image retrieval (CBIR) [20, 21, 22] is the application of computer vision for retrieving the images which means searching the digital images in massive databases is very difficult. The term “Content based” means that it will search the actual content of an image. Information retrieval means the process of converting a request for information into a meaningful set of reference. CBIR is a technology that in principle helps organize digital image archives according to their visual content. This system distinguishes the different regions present in an image based on their similarity in color, texture, shape, etc. and decides the similarity between two

images by reckoning the closeness of these different regions.

Content Based Image Retrieval [20, 21, 22] systems can be classified into two ways according to the type of queries: text based query and pictorial based query. In text query based query, images are defined by text information such as keywords and captions. Text features are powerful as a query, if appropriate text descriptions are given for images in an image database. However, giving appropriate descriptions must be done manually in general and it is time consuming. There are many ways one can pose a visual query. A good query method will be effective to the user as well as capturing information from the user to extract meaningful results. In pictorial query based systems, an example of the desired image is used as a query. To retrieve similar images with the example, image features such as colors and textures, most of which can be extracted automatically when it is used.

The CBIR system [21] provides two major responsibilities. One is feature extraction [11,21] is developed accurately to define the content of every image in the database. It is much smaller in size than the original image, typically of the order of hundreds of elements. The second one is similarity measurement, where a distance between the query image and each image in the database using their signatures is computed so that the top “closest” images can be retrieved.

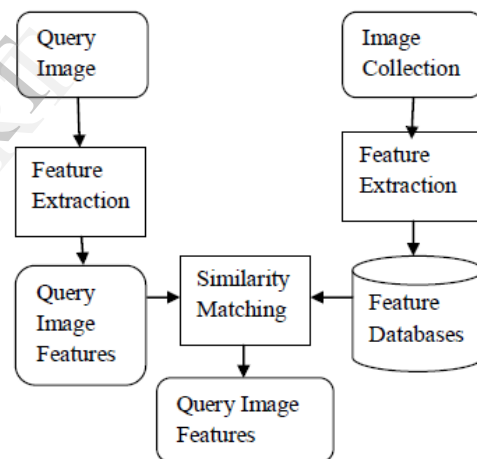


Figure 1: CBIR System Architecture

A. Feature Extraction

Feature extraction [11, 21] is the beginning of content based image retrieval. It is the process of extracting image features to a distinguishable extent. It is a group of features called image signature. It is carried out by using colors, textures or shapes. Once obtained, visual features act as inputs to subsequent image analysis tasks such as similarity estimation, concept detection, or annotation. A feature is referred to capture a certain visual property of an image, which covers wholly for the entire image. There are various kinds of primitive features to represent an image as color, texture and shape. One type of features can only represent part of the image properties, a lot of work done on the combination of these features. However, there is no single “best” feature that gives accurate results in any general setting. Usually, a

combination of features is minimally needed to provide adequate retrieval results.

- **Color**

The first and most straightforward feature for indexing and retrieving images is color. Color [19] is an immediately perceivable visual feature when looking at an image. It is mainly used for image similarity retrieval. Color space is used to represent color images. However, RGB space denotes the gray level intensity is represented as the sum of red, green and blue gray level intensities Color moments have been successfully used in many retrieval systems especially when the image contains just the object.

The first order (mean), the second (variance) and the third order (skewness) color moments have been proved to be efficient and effective in representing color distributions of images. Every image inserted to the collection is analyzed to estimates a color histogram which defines the quantity of pixels of color within the image. Color histogram of an image is a description of the colors present in an image and in what quantities. They are computationally efficient to compute and insensitive to small perturbations in camera position. The Color Structure Descriptor determines an image for color distribution of the image and the local spatial structure.

The major aspects for feature extraction are the choice of a color space. It is a multidimensional space in which the various dimensions determine the different components of colors [4]. It is three dimensional. Example of a color space is denoted as RGB, which assigns for each pixel a three element vector providing an individual color intensity of the three primary colors such as red, green and blue. The space spanned by the R, G, and B values completely represents visible colors in the 3D RGB color space. Retrieving images based on colour similarity is achieved by computing a colour histogram for individual image that visualize the quantity of pixels within an image holding particular values.

- **Shape**

Shape [19] mainly defined as the characteristic surface configuration of an object; an outline or contour. It shows that objects are mostly familiar by their shape. Shape feature alone provides capability to recognize objects and retrieve similar images on the basis of their contents. A number of features qualities of object shape are estimated for every object recognized in the stored image. Queries for shape retrieval to be defined by giving an example for each image to act as the query. Retrieving those stored images whose features are closely matches with that particular query image. Shape feature are commonly defined in two ways – global features defines aspect-ratio, circularity and moment invariants and local features defines group of consecutive boundary.

Li and Ma [17] discussed that the geometric moment's method or region-based and the fourier descriptor or boundary-based were related by a simple linear transformation. Babu *et al.* [18] compared the performance of boundary-based representations such as chain code, Fourier descriptor and UNL Fourier descriptor, region-based representations such as moment invariants, Zernike moments

and pseudo-Zernike moments and combined representations such as moment invariants and Fourier descriptor, moment invariants and UNL Fourier descriptor. Their experiments showed that the combined representations outperformed the simple representations.

- **Texture**

Image texture [19] is a widely used and primitive visual feature of an image. Texture feature plays important role to separate regions. It refers to the visual patterns that have property of homogeneity or arrangement that do not result from the presence of only a single colour or intensity. This is widely used because it is based on human texture representation. Various texture representations have been investigated in both pattern recognition and computer vision. It focuses property of nearly all surfaces such as clouds, trees, bricks, hair, and fabric. Textures are determined by texels which must be fixed into a many sets, based on how many textures which are detected in the image. It not only defines the texture, but also defines the image in which where the texture is located.

The six texture properties were coarse, contrast, directional, line likeness, regularity and roughness. The most common measures for capturing the images are wavelets and Gabor filters. Which try to retrieve the image or image parts characteristics with reference to the changes in certain directions and the scale of the images. This is most useful for region or images with homogeneous texture.



Figure 2: Examples for texture

Texture is a difficult concept to visualize. The specific textures in an image are defined primarily by modelling the texture in a two dimensional gray level variation. The accurate brightness of set of pixels is estimated as degree of contrast, regular, coarseness and directional may be estimated. It matches texture regions in images to words representing texture attributes.

B. Similarity Measures

It involves matching these features to get a result which is visually similar. Instead of exact matching, content-based image retrieval calculates visual similarities between a query image and images in a huge database. The result is not a single image but a list of images ranked by their similarities based on the query image. Many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features in recent years. Distance method is used for similarity measures. Different similarity/distance

measures will affect retrieval performances of an image retrieval system significantly.

Similarity measures for color features are- Histogram Quadratic Distance Measure, Integrated Histogram Bin Matching, Histogram intersection, Histogram Euclidean distance, Minkowskimetric, Manhattan distance, Canberra distance, Angular distance, czekanonski coefficient, Inner product, Dice coefficient, Cosine coefficient, Jaccard coefficient.

Similarity measurement for texture features are- Kull back-leiber distance, Tree structured wavelet transform, Generalized Gaussian density, Histogram method, wavelet transform, Pyramid structured wavelet transform, Multiresolution simultaneous autoregressive model, weighted Euclidean distance, Monte-Carlo method and Earth movers distance shows higher accuracy and flexibility in focusing texture information.

Similarity measurement for shape features are- Perceptual distance, Polygon approximation method, Fourier descriptor method, Time Wrapping, Angular distance, Inner product, Dice coefficient, Ray distance and Ordinal co-relation. DTW develop a comfortable distance calculation scheme which is enough with the human visual system in perceiving shape similarity.

V. CONCLUSION

In this paper, we focussed on Image retrieval such as text based and content based image retrieval. Text- based image retrieval has some limitations such as task of determining image content is highly perspective. So overcomes this problem, we will discuss the CBIR system. CBIR is a fast developing technology with considerable potential. Research in CBIR has been focused on image processing, low level feature extraction and so on. It has been believed that CBIR provides maximum support in bridging 'semantic gap' between low level feature and richness of human semantics. Feature extraction is the process of extracting image features to a distinguishable extent. CBIR system distinguishes the different regions present in an image based on their similarity in colour, texture and shape. CBIR technology has been used in many application areas such as fingerprint identification, biodiversity, digital libraries, crime prevention, medicine, historical research. Similarity measures are used to determine how similar or dissimilar in the given query image and image database collections. In this paper we focused on the study of content based image retrieval and future enhancement is to implement the content based image retrieval in medical field.

REFERENCES

- [1] O.R. Zajane, J. Han Z.N. Li and J. Hou, "Mining multimedia data", Proc. of SIGMOD, 1998.
- [2] J. Zhang, H. Wynne, and M. L. Lee, "Image mining: Issues, frameworks, and techniques", Proc. 2nd Int. Workshop Multimedia Data Mining, pp.13 -20, 2001.
- [3] T.Karthikeyan, P.Manikandaprabhu, "Function and Information Driven Frameworks for Image Mining - A Review", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Vol.2, Issue 11, pp.4202-4206, Nov. 2013.
- [4] S.K. Chang and A. Hsu, "Image information systems: Where do we go from here?", IEEE Trans. on Knowledge and Data Engineering 4(5), 1992.
- [5] H. Tamura and N. Yokoya, "Image database systems: A survey", Pattern Recognition 17(1), 1984.
- [6] V. Ogle and M. Stonebraker, "Chabot: Retrieval from a relational database of images", IEEE Computer, 28(9):40-48, 1995.
- [7] S. Mehrotra, Y. Rui, M. Ortega and T. S. Huang, "Supporting content based queries over images in MARS", Proc. IEEE Int. Conf. Multimedia Computing Systems, ON, Canada, pp. 632-633, June. 1997.
- [8] W. Y. Ma and B. S. Manjunath, "Netra A toolbox for navigating image databases", Proc. IEEE Int. Conf. Image Processing, Santa Barbara, CA, vol. 1, pp. 568-571, Oct. 1997.
- [9] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Content-based manipulation of image databases," Proc. SPIE Storage Retrieval Image Video Databases II, pp. 34-47, Feb. 1994.
- [10] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack, D. Petkovic, and W. Equitz, "Efficient and effective query by image content," J.Intell. Inform. Syst., vol. 3, pp. 231-262, 1994.
- [11] Yong Rui and Thomas S. Huang, "Image Retrieval in Current Techniques, Promising Directions, and Open Issues", Journal of Visual Communication and Image Representation 10, 39-62, 1999.
- [12] hahab Nastar, Matthias Mitschke, Christophe Meilhac, and Nozha Boujemaa. "Surfimage: A flexible content-based image retrieval system", Proc. ACM International Multimedia Conference, Bristol, England, pp. 339-344, Sep. 1998.
- [13] H. J. Zhang, C. Y. Low, S. W. Smoliar, and J. H. Wu, "Video parsing retrieval and browsing: An integrated and content-based solution", Proc. ACM Multimedia '95, San Francisco, CA, pp.15-24, Nov. 1995.
- [14] A. Hampapur, A. Gupta, B. Horowitz, C. F. Shu, C. Fuller, J. Bach, M.Gorkani, and R. Jain, "Virage video engine", Proc. SPIE Storage Retrieval Image Video Databases V, San Jose, CA, pp. 188-197, Feb. 1997.
- [15] J. R. Smith and S. F. Chang, "A fully automated content based image query system", Proc. ACM Multimedia, Boston, MA, pp. 87-98, Nov. 1996.
- [16] J. R. Smith and S.F. Chang, Visually searching the web for content, IEEE Multimedia Magazine 4(3), 12-20, 1997.
- [17] B. Li and S. D. Ma, "On the relation between region and contour representation", Proc. IEEE Int. Conf. on Image Proc., 1995.
- [18] B. M. Mehre, M. Kankanhalli, and W. F. Lee, "Shape measures for CBIR: A comparison", Information Processing & Management 33(3), 1997.
- [19] Ritendra Datta, Dhiraj Joshi, Jia Li, And James Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age", ACM Computing Surveys, Vol. 40, No. 2, Article 5, April 2008.
- [20] A. W. Smeulders, S. Santini, A. Gupta, and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans. Pattern Analysis and Machine Intelligence, 22(12):1349-1380, 2000.
- [21] Datta R, Li J, Wang J Z. " Content-based Image Retrieval – Approaches and Trends of the New Age", ACM Intl. Workshop on Multimedia Information Retrieval, Singapore, ACM Multimedia, 2005.
- [22] Michael S. Lew, Nicu Sebe, Chabane Djeraba and Ramesh Jain, "Content based Multimedia Information Retrieval in State of the Art and Challenge", ACM Transactions on Multimedia Computing, Communications and Applications, Feb. 2006.