Content Based Image Retrieval: Tools and Techniques

[1]Ishpreet Singh Virk, [2]Raman Maini

[1]University College of Engineering, Punjabi University, Patiala

[2]Associate Professor, University College of Engineering, Punjabi University, Patiala

ishvirk55@gmail.com, research_raman@yahoo.com

Abstract: *This paper provides the comprehensive overview of techniques and tools in the research area of Image retrieval, especially Content-Bases Image Retrieval (CBIR), a vivid research area for the last two decades. However, there are still many challenges to be overcome. Image feature representation and extraction, multi dimensional indexing and retrieval system design are fundamental basis for Content-Bases Image Retrieval (CBIR).We reviews various visual features and their corresponding representation and matching techniques. To provide fast and effective search in large scale image collection, high dimension indexing techniques needs to be implement. State-of-art commercial and research systems and their characteristics are described. Finally, we look at browsing systems as an alternative approach to dealing with large image databases. The hue sphere browsing system organizes images on a spherical visualization space so that visually similar images are located close to each other, and provides powerful interaction tools that aid in exploring the visualized image collections.*

Keywords: *content based, retrieval, feature extraction, indexing, hue sphere.*

1. Introduction

In recent years there is very much increase in the size of digital image collections. Every day, various equipments click giga-bytes of images. A large amount of information (Images) is present there. However, we cannot access or use the information unless it is organized so as to allow efficient browsing, searching, and retrieval. Image retrieval has been a very active research area since the 1970s with the input from two major research areas database management and computer vision. The text-based image retrieval can be traced back to the late 1970s. Avery popular technique of image retrieval, then was to first annotate the images by text and then use text-based database management systems (DBMS) to perform image retrieval Representatives of this approach are [1, 2].However, there exist two major difficulties, one is the vast amount of labor required in manual image annotation and other difficulty, is the subjectivity of human perception in retrieval processes. In the early 1990s, to overcome these difficulties, content-based image retrieval was proposed. That is, instead of being manually annotated by text-based key words, images would be indexed by their own visual content, such as color and texture.

2. Feature Extraction

Feature (content) extraction provides the basis for content-based image retrieval. The goal of feature extraction is to obtain both text-based features (key words, annotations) [3,4] and visual features (color, texture, shape, faces), perceptually relevant representation of the image content. Within the visual feature scope, the features can be further classified as general features and domain specific features. The former include color, texture, and shape features while the latter is application-dependent and may include, for example, human faces and finger prints. Because of perception subjectivity, there does not exist a single best presentation for a given feature. As we will see soon, for any given feature there exist multiple representations which characterize the feature from different perspectives.

2.1 Color and Color Invariants for CBIR

Content-Bases Image Retrieval (CBIR) based on color features [5]–[7] is at the heart of many image database systems. Color descriptors are among the most important features used in image analysis and retrieval. Due to its compact representation and low complexity, direct histogram comparison is a commonly used technique for measuring the color similarity. However, it has many serious drawbacks, including a high degree of dependency on surface characteristics of the captured objects but also on capturing conditions such as scene illumination, color codebook design, and sensitivity to quantization boundaries, and inefficiency in representing images with few dominant colors. Addressing these confounding factors, various color invariants have been introduced. When these factors are not accounted for, Color-based CBIR can perform poorly [8].Two I1 and I2 images taken with a different digital device is composed of sensor responses that can be described by [12]

$$I1 \neq I2 \qquad (1)$$

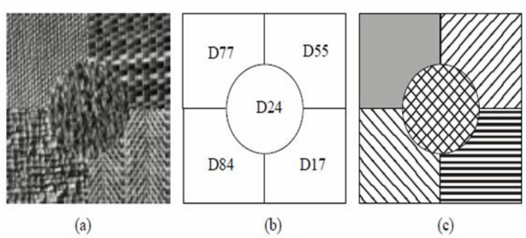
Fortunately the differences in I1 and I2 are not arbitrary. Under reasonable conditions [9], the following relation holds

$$I2 = I1\ D \qquad (2)$$

Color invariants are color features that do not change with a change in scene illumination. In other words, color invariants are features for which the elements of D in Equation (2) cancel. Several color invariants have been proposed in the literature and have been shown to work well for image retrieval [10] although it should also be noted that the application of color invariants can also lower retrieval performance for image datasets where illumination differences are not an issue [11]. In image retrieval, the color histogram is the most commonly used color feature representation. Statistically, it denotes the joint probability of the intensities of the three color channels. Furthermore, considering that most color histograms are very sparse and thus sensitive to noise, we can use the cumulated color histogram [13]. Besides the color histogram, other color feature representations like color moments and color sets are used in CBIR to overcome the quantization effects in the color histogram [13]. Since most of the information is concentrated on the low-order moments, only the first moment (mean), and the second and third central moments (variance and skewness) were extracted as the color feature representation. Weighted Euclidean distance was used to calculate the color similarity. To provide fast search in large-scale image collections, an approximation to the color histogram [14, 15] can be used. In this method, firstly we transformed the (R, G, B) color space into a uniform space, such as HSV, and then quantized the transformed color space into *M* bins. A color set is defined as a selection of colors from the quantized color space. Because color set feature vectors were binary, a binary search tree was constructed to allow a fast search. The relationship between the proposed color sets and the conventional color histogram was further discussed [15, 16].

2.2 Texture

An image texture is described by the number and types of its (tonal) primitives and the spatial organization or layout of its (tonal) primitives. Image texture, defined as a function of the spatial variation in pixel intensities (gray values), is useful in a variety of applications and has been a subject of intense study by many researchers. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [18].One immediate application of image texture is the recognition of image regions using texture properties [17].



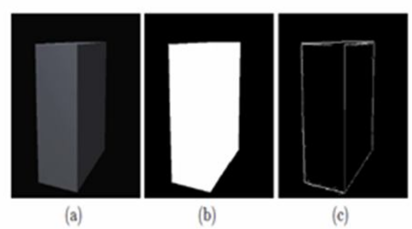


Fig 1 (a) An image consisting of 5 different textured regions: Fig 2. Processing input image

(a) Gay scale (b) Binary image (c) Contour Image

The shape from texture problem is one instance of a general class of vision problems known as “shape from X.” The goal is to extract three-dimensional surface shape from variations in textural properties in the image. A texture image retrieval method based on gray-primitive co-occurrence matrix is used [19]. The method combines the statistical method and structure method, which could describe the image gray level distribution information and local detail and could describe image with less number of features. By analyzing the results of experiments, this method has rotation invariance and certain noise resisting ability.

2.3 Role of Shape in CBIR

In image retrieval system the shape representation should to be invariant to translation, rotation, and scaling. Shape of an image can be obtained by separating the object information from its background details. This is done by converting the given image into a gray scale image and then to binary image.

It is observed that shape features provide a effective method for indexing the images, because many of the images provide very small texture information for the anatomy of interest. In addition to 2D shape representations, there were many methods developed for 3D shape representations. In general, the shape representations can be divided into two categories, boundary-based and region-based. The former uses only the outer boundary of the shape while the latter uses the entire shape region. The most successful representatives for these two categories are Fourier descriptor and moment invariants. The various shape representation and similarity techniques are:

- Geometry based shape: These methods use shape properties like area, perimeter, convexity, orientation etc [21].
- Invariant moments: Several forms of invariant moments are seen. The main idea of moment invariants is to use region-based moments which are invariant to transformations, as the shape feature [22].
- Fourier transform based methods: The main idea of a Fourier descriptor is to use the Fourier transformed boundary as the shape feature. A modified Fourier descriptor which is both robust to noise and invariant to geometric transformations [20].

2.4 Segmentation Methods of Images

Segmentation is very important to image retrieval system. Both the shape feature and the layout feature depend on good segmentation. We will discuss some of the segmentation techniques used in both image retrieval computer vision. In [23] researched a morphological operation (opening and closing) approach in image segmentation. In image retrieval, several systems attempt to perform an automatic segmentation of the images in the collection for feature extraction [24]. After segmentation, the resulting segments can be described by shape features that commonly exist, including those with invariance's with respect to shifts, rotations and scaling [25]. All the above-mentioned algorithms are automatic, but sometimes the automatic segmentation is not always reliable. In [26], a computer-assisted boundary extraction approach, which combined manual inputs from the user with the image edges generated by the computer.

2.5 Image Compression and Compression Domain CBIR

We have still limited storage space, transmission bandwidth (Internet) and *transmission* time means images needs to store in compressed form. Image compression, which is applied to the most images, leads to both image processing overheads but also to a small but noticeable drop/overhead in image retrieval performance. It is also notice that image features of compressed images are different from those extracted from their original image. Today most of the images are compressed and stored using JPEG [26]. JPEG uses the discrete cosine transform followed by quantization in the frequency domain to achieve high compression ratios.Although most images exist only in compressed form, almost all CBIR techniques operate in the pixel domain. In contrast, compressed domain techniques operate directly on the compressed data without the need for decompression [26]. Compressed domain CBIR can be performed either based on existing compression formats such as JPEG [26] or vector quantization or employing so-called 4-th criterion compression techniques where the compressed information is directly visually meaningful [27].

We have explored many visual features both in image retrieval applications and computer vision. For each visual feature, there may be multiple representations which model the human perception of that feature from different perspectives. What features and representations should be used in image retrieval is application dependent.

3. Multi Dimenstional Indexing

Before we utilize any indexing technique, it is beneficial to first perform dimension reduction. We have two approaches have appeared in the literature, i.e. Karhunen–Loeve transform (KLT) and column-wise clustering. From experimental results showed that most of the visual feature vectors can be considerably reduced in dimension without significant degradation in retrieval quality [28].

3.1 Indexing Technique

After extracting the feature vectors of the image, we need to select appropriate multidimensional indexing algorithms to index the reduced but still high dimensional feature vectors of the image. There is much work done in this area, the popular multidimensional indexing techniques are bucketing algorithm, k-d tree, priority k-d tree , quad-tree, K-D-B tree, B-tree, R-tree and its variants R^{+}-tree and R^{*}-tree. Very good reviews and comparisons of various indexing techniques in image retrieval can be found in [29]. In [30], Ng and Sedighian proposed a three-step strategy towards image retrieval indexing, i.e. dimension reduction, evaluation of existing indexing approaches, and customization of the selected indexing approach. After dimension reduction using the Eigen image approach, the following three characteristics of the dimension-reduced data can be used to select good existing indexing algorithms:

- The new dimension components are ranked by decreasing variance,
- The dynamic ranges of the dimensions are known,
- The dimensionality is still fairly high.

4. Comparison Technique Used

Image retrieval system often used measurement systems such as Euclidean vector space model [31] for measuring distances between a query images (represented by its visual feature) with possible results representing all images as feature vectors in an n-dimensional vector space. Some other distance measures exist e.g. city-block distance, Mahalanobis distance [31]. Still the use high dimensional feature space has shown problem and needs care about choosing distance measure technique

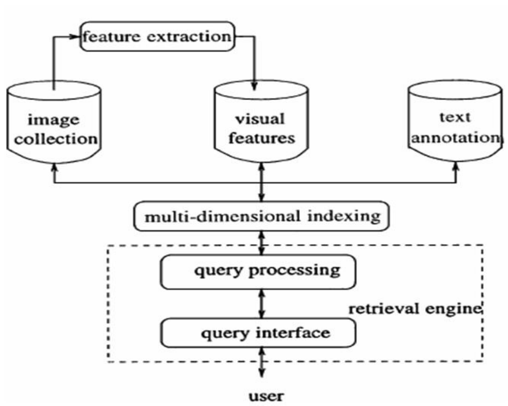


FIG. 3. Image retrieval system architecture.

4.1 Evaluation methods of CBIR

The most common evaluation measures used in CBIR are precision and recall usually presented as a precision vs. recall graph (PR graph). Researchers are familiar with PR graphs and can extract information from them without interpretation problems.

$$\text{precision} = \frac{\text{no. relevant items retrieved}}{\text{no. items retrieved}} \tag{3}$$

$$\text{recall} = \frac{\text{no. relevant items retrieved}}{\text{no. relevant items}} \tag{4}$$

5. EXITING CBIR SYSTEMS

Since the early 1990s, content-based image retrieval has become a very vivid research area. Many image retrieval systems, both commercial and research, have been built. Most image retrieval systems support one or more of the following options [32]:

- Random browsing
- Search by example
- Search by sketch

- Search by text (including key word or speech)
- Navigation with customized image categories.

5.1 QBIC

QBIC [31] stands for Query By Image Content is the first commercial content-based image retrieval system (CBIR). It provide framework and techniques basis for many image retrieval systems. QBIC supports queries based on example images, user-constructed sketches and drawings, and selected color and texture patterns, etc. The color feature used in QBIC are the average (R,G,B), (Y,i,q), (L,a,b), and MTM mathematical transform to Munsell) coordinates, and a k-element color histogram. In its new system, text-based key word search can be combined with content-based similarity search. Its texture feature is an improved version of the Tamura texture representation i.e. combinations of coarseness, contrast, and directionality. Its shape feature consists of shape area, circularity, eccentricity, major axis orientation, and a set of algebraic moment invariants. The on-line QBIC demo is at http://wwwqbic.almaden.ibm.com/

5.2 Virage

Virage is a content-based image search engine developed at Virage Inc. Similar to QBIC; Virage[4] supports visual queries based on color, composition (color layout), texture, and structure (object boundary information). But Virage goes one step further than QBIC. It also supports arbitrary combinations of the above four atomic queries. The users can adjust the weights associated with the atomic features according to their own emphasis. The corresponding demos of Virage are at http://www.virage.com/cgi-bin/query-e.

5.3 Retrieval Ware

Retrieval Ware is a content-based image retrieval engine developed by Excalibur Technologies Corp. From one of its early publications, we can see that its emphasis was in neural nets to image retrieval. Its more recent search engine uses color, shape, texture, brightness, color layout, and aspect ratio of the image, as the query features .It also supports the combinations of these features and allows the users to adjust the weights associated with each feature. Its demo page is at http://vrw.excalib.com/cgi- bin/sdk/cst/cst2.bat.

5.4 Netra

Netra is a prototype image retrieval system developed in the UCSB Alexandria Digital Library (ADL) project. Netra uses color, texture, shape, and spatial location information in the segmented image regions to search and retrieve similar regions from the database. The on-line demo is at http://vivaldi.ece.ucsb.edu/Netra/.

5.5 Photobook

Photobook [33] is a set of interactive tools for browsing and searching images developed at the MIT Media Lab. Photobook consists of three subbooks from which shape, texture, and face features are extracted, respectively. Users can then query, based on the corresponding features in each of the three subbooks. The motivation of this was based on the observation that there was no single feature which can best model images from each and every domain.

5.6 MARS

MARS (multimedia analysis and retrieval system) was developed at University of Illinois at Urbana-Champaign [20].MARS differs from other systems in both the research scope and the techniques used. It is an interdisciplinary research effort involving multiple research communities: computer vision, database management system (DBMS), and information retrieval (IR). The main focus of MARS is not on finding a single "best" feature representation, but rather on how to organize various visual features into a meaningful retrieval architecture which can dynamically adapt to different applications and different users. MARS formally proposes relevance feedback architecture in image retrieval and integrates such a technique at various levels during retrieval, including query vector refinement, automatic matching tool selection, and automatic feature adaption. The on-line demo is at http://jadzia.ifp.uiuc.edu:8000.

6. IMAGE DATABASE BROWSING

Browsing systems provide an overview of the image database to allow for intuitive navigation through the image collection. This is particularly the case when images are arranged according to mutual similarity [36], that is, visually similar images are located close to each other in the browsing interface. One of the main challenges of such image database navigation systems is the limited screen size. The hue sphere browsing system [34]–[35] provides a very efficient and effective approach to image database browsing. It employs a spherical visualization space since users will be familiar with the concept of a globe and how to navigate on one.

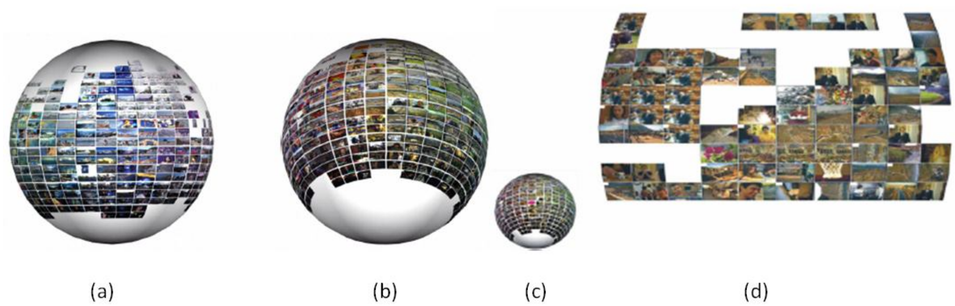


Fig. 4. (a) Hue sphere view of the MPEG-7 colour dataset, (b) View after rotation and tilt operations, (c), (d) View after zoom operation.

7. Conclusion

Due to very compact representation and low complexity, color quantization with respect to the fixed codebook, followed by the histogram computation is the most often used feature extraction technique in image database retrieval**.** The visual feature database stores the visual features extracted from the images using techniques described in Section 2. This is the information needed to support content-based image retrieval. Image compression was shown to have a negative impact on retrieval performance, and we have consequently looked into compressed domain retrieval techniques where CBIR is performed directly on compressed data without the need for decoding. Future efforts have to be presented to select a reduced set of data that could represent the entire data (image). Finally, we have explored image browsing systems as an alternative approach of handling large image collections.

8. References

[1] N. S. Chang and K. S. Fu, A Relational Database System for Images, TR-EE 79-28, Purdue University, May 1979.

[2] S.-K. Chang, Pictorial data-base systems, IEEE Computer, 1981.

[3] A. D. Narasimhalu, Special section on content-based retrieval, Multimedia Systems, 1995.

[4] Special issue on visual information management (R. Jain, Guest Ed.), Comm. ACM, Dec. 1997.

[5] M.J. Swain and D.H. Ballard, “Color indexing,” Int. Journal of Computer Vision, vol. 7, no. 11, pp. 11–32, 1991.

[6] G. Pass and R. Zabih, “Histogram refinement for content-based image retrieval,” in 3rd IEEE Workshop on Applications of Computer Vision,1996, pp. 96–102.

[7] [9] L. Cinque, S. Levialdi, and A. Pellicano, “Color-based image retrieval using spatial-chromatic histograms,” in IEEE Int. Conf. Multimedia Computing and Systems, 1999, pp. 969–973.

[8] B.V. Funt and G.D. Finlayson, “Color constant color indexing,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 17, no. 5, pp. 522–529, 1991.

[9] J.A. Worthey and M.H. Brill, “Heuristic analysis of Von Kries colour constancy,” Journal Optical Society of America A, vol. 3, no. 10, pp. 1708–1712, 1986.

[10] R. Lenz, L.V. Tran, and P. Meer, “Moment based normalization of colour images,” in IEEE Signal Processing Society Workshop on Multimedia Signal Processing, 1999.

[11] G. Schaefer, “How useful are colour invariants for image retrieval?,” in Int. Conference on Computer Vision and Graphics, Computational Imaging and Vision. Springer, 2006.

[12] G. Schaefer, “Content based retrieval from image databases: Color, compression and browsing” 978-1-4244-5651-2/10 2010 crown

[13] M. Stricker and M. Orengo, Similarity of color images, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995*.

[14] J. R. Smith and S. F. Chang, Transform features for texture classification and discrimination in large image databases, in *Proc. IEEE Int. Conf. on Image Proc., 1994*.

[15] J. R. Smith and Chang, Single color extraction and image query*. IEEE Conf. on ImageProc., 1995*.

[16] J. R. Smith and S.-F. Chang, Tools and techniques for color image retrieval, in *IS & T/SPIE Proceedings, Vol. 2670, Storage & Retrieval for Image and Video Databases IV, 1995*.

[17] The Handbook of Pattern Recognition and Computer Vision (2nd Edition), by C. H. Chen, L. F. Pau, P. S. P. Wang (eds.), pp. 207-248, World Scientific Publishing Co., 1998.

[18] R. M. Haralick, K. Shanmugam, and I. Dinstein, Texture features for image classification, IEEE Trans. On Sys. Man. and Cyb. SMC-3(6), 1973.

[19] Texture Image Retrieval Based on Gray-Primitive Co-occurrenceMatrix by WeiWANG, SUZUKI, Fuji REN.

[20] Y. Rui, A. C. She, and T. S. Huang, Modified fourier descriptors for shape representation—a practical approach, in Proc. of First International Workshop on Image Databases and Multi Media Search, 1996.

[21] B. Li and S. D. Ma, On the relation between region and contour representation, in Proc. IEEE Int. Conf. on Image Proc., 1995.

[22] D. Copper and Z. Lei, On representation and invariant recognition of complex objects based on patches and parts, in Lecture Notes in Computer Science Series, 3D Object Representation for Computer

Vision (M. Hebert, J. Ponce, T. Boult, and A. Gross, Eds.), pp. 139–153, Springer-Verlag, New York/Berlin, 1995.

[23] M. Lybanon, S. Lea, and S. Himes, Segmentation of diverse image types using opening and closing, in Proc. IEEE Int. Conf. on Image Proc., 1994.

[24] L. Lucchese, S. K. Mitra, Unsupervised segmentation of color images based on k-means clustering in the chromaticity plane, in: IEEE Workshop on Content{based Access of Image and Video Libraries (CBAIVL'99), Fort Collins, Colorado, USA, 1999, pp. 74-78.

[25] S. Loncaric, A survey of shape analysis techniques, Pattern Recognition 31 No 8 1998) 983-1001.

[26] G.K. Wallace, "The JPEG still picture compression standard," Communications of the ACM, vol. 34, no. 4, pp. 30–44, 1991.

[27] R.W. Picard, "Content access for image/video coding: The fourth criterion," Tech. Rep. 195, MIT Media Lab, 1994.

[28] D. White and R. Jain, Similarity indexing: Algorithms and performance, in Proc. SPIE Storage and Retrieval for Image and Video Databases, 1996.

[29] D. White and R. Jain, Algorithms and strategies for similarity retrieval, in *TR VCL-96-101*, University of California, San Diego, 1996.

[30] R. Ng and A. Sedighian, Evaluating multi-dimensional indexing structures for images transformed by principal component analysis, in *Proc. SPIE Storage and retrieval for Image and Video Databases, 1996.*

[31] W. Niblack, R. Barber, W. Equitz, M. D. Flickner, E. H. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, G. Taubin, QBIC project: querying images by content, using color, texture, and shape, in: W. Niblack (Ed.), Storage and Retrieval for Image and Video Databases, Vol. 1908 of SPIE Proceedings,1993, pp. 173-187.

[32] S.-F. Chang, A. Eleftheriadis, and R. McClintock, Next-generation content representation, creation and searching for new media applications in education, *IEEE Proceedings, 1998*, to appear.

[33] A. Pentland, R. W. Picard, and S. Sclaroff, Photobook: Content-based manipulation of image databases, International Journal of Computer Vision, 1996.

[34] G. Schaefer and S. Ruszala, "Image database navigation: A globe-al approach," in Int. Symposium on Visual Computing, 2005, vol. 3804 of Springer Lecture Notes on Computer Science, pp. 279–286.

[35] G. Schaefer and S. Ruszala, "Hierarchical image database navigation on a hue sphere," in Int. Symposium on Visual Computing, 2006, vol.

[36] K. Rodden, D. Basalaj, W. Sinclair, and K. Wood, "Evaluating a visualisation of image similarity as a tool for image browsing," in IEEE Symposium on Information Visualization, 1999, pp. 36–43.

* * * * *