Content-Based Image Retrieval Application: A Review

Mr.Shelar Balaji T[1] Prof. Pergad N .D[2]
[1,2]Department of Electronics & Telecommunication Engineering
[1,2]S.T.B.COE Tuljapur, Maharashtra, India

Abstract— The paper presents a review of different techniques in content-based image retrieval. The paper starts with discussing the fundamental aspects of CBIR. Features for Image Retrieval like color, texture and shape are discussed next. We briefly discuss the similarity measures based on which matches are made and images are retrieved. Another important issue in content-based image retrieval is effective indexing and fast searching of images based on visual features. Dimension reduction and indexing schemes are also discussed. 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. Finally, Relevance feedback is discussed which helps in improving the performance of a CBIR system. Potential uses for CBIR include Architectural and engineering design, collections, Crime, Geographical information and remote sensing systems, Intellectual, Medical, Military, Photograph archives, Nudity-detection filters, Face Finding.

Key words: Content-Based Image Retrieval (CBIR), query by image content (QBIC)

I. INTRODUCTION

Content-based picture recovery (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) may be those requisitions of workstation dream strategies of the picture recovery problem, that is, those issues of seeking to advanced pictures previously, huge databases for an later exploratory diagram of the CBIR field). Content-based picture recovery may be restricted with conventional concept-based methodologies "Content-based" implies that the quest analyzes those substance of the picture instead of those metadata for example, such that keywords, tags, alternately portrayals connected with the picture. Those terms "content" in this setting could allude should colors, shapes, textures, alternately whatever available majority of the data that cam wood be inferred from those pictures itself. CBIR may be alluring as a result searches that depend purely with respect to metadata need aid subject to annotation caliber Also culmination. Hosting people manually clarify pictures by entering keywords or metadata over an expansive database could be time expending Also might not catch those keywords fancied to describe those picture. Those assessment of the adequacy about Pivotal word picture hunt is subjective what's more need not been well-defined. In the same regard, CBIR frameworks need comparable tests over characterizing accomplishment. As information technology proliferates throughout our society, digital images and video or visual objects are becoming as important as traditional textual based information. This phenomenon has several reasons: Demilitarization of imaging and satellite technology, the emergence of the World Wide Web as a digital Communications infrastructure, the impending convergence of computers and television, and the increase in use and availability of digital cameras and video recorders. With the massive growth in the amount of visual information available, there exists a real need for systems to catalogue and provide retrieval from digital image and video libraries. 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. However, there remain many challenging research Problems that continue to attract researchers from multiple disciplines. Information retrieval is the processes of converting are quest for information into a meaningful set of references. Early work on image retrieval can be traced back to the late 1970s. In 1979, a conference on Database Techniques for Pictorial Applications was held in Florence. Since then, the application potential of image database management techniques has attracted the attention of researchers. Early techniques were not generally based on visual features but on the textual annotation of images.

II. FUNDAMENTAL ASPECTS OF CBIR

Previous CBIR systems can be classified into two categories according to the type of queries: text query or pictorial query. In text query based systems, images are characterized 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 natural to the user as well as capturing enough 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 colours and textures, most of which can be extracted automatically, are used. The typical CBIR system performs two major tasks. The first one is feature extraction (FE), where a set of features, called image signature or feature vector, is generated to accurately represent the content of each image in the database. A feature vector is much smaller in size than the original image, typically of the order of hundreds of elements (rather than millions). The second task is similarity measurement (SM), 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.

Fig a. Shows Below the block diagram of Typical CBIR system is shown.

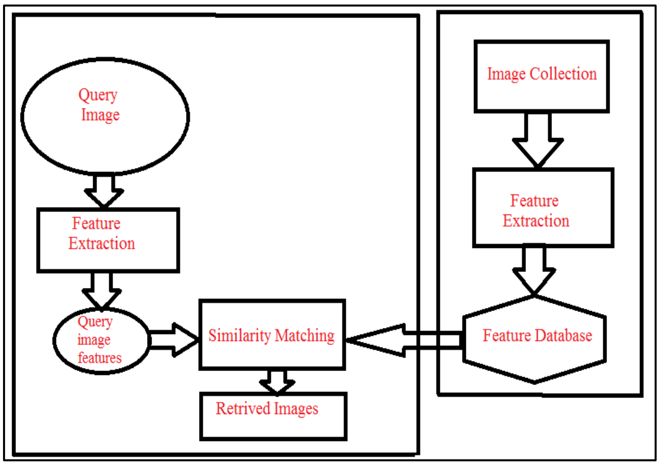

Fig. 1: block diagram of Typical CBIR system is shown.

III. Feature Representations

Feature extraction and representation is the fundamental process behind CBIR systems. As mentioned, features are properties of the image extracted with image processing algorithms, such as colour, texture, shape, and edge information. Our discussion will focus on three general features representations that have been extensively studied in the literature: colour, texture, and shape. 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 since perceptual subjectivity permeates throughout this problem.

A. Color Feature

The first and most straightforward feature for indexing and retrieving images is color, the basic constituent of images (we consider grayscale a color). All other information computed by image processing algorithms start with the color information contained in an image. Color moments have been successfully used in many retrieval systems (like QBIC [19, 20]), 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 [21]. Each image added to the collection is analyzed to compute a color histogram which shows the proportion of pixels of each color within the image. The color histogram for each image is then stored in the database. The color histogram of an image is a description of the colours present in an image and in what quantities. They are computationally efficient to compute and insensitive to small perturbations in camera position. At search time, the user can either specify the desired proportion of each color (75% olive green and 25% red, for example), or submit an example image from which a color histogram is calculated.

Either way, the matching process then retrieves those images whose color histograms match those of the query most closely. When an image database contains a large number of images, histogram comparison will saturate the discrimination. To solve this problem, the joint histogram technique is introduced [22]. However, it is possible that two images have the same color histogram even though they have completely different appearances, because a single color histogram extracted from an image, which is used in most of histogram-based image retrieval systems, lacks spatial information of colors in the image[4][5]. To overcome this problem, Yamamoto et al [3] propose a content based image retrieval system which takes account of the spatial information of colours by using multiple histograms. The proposed system roughly captures spatial information of colors by dividing an image into two rectangular sub-images recursively. The proposed method divides an image into dominant two regions using a straight line vertically or horizontally, even when the image has three or more color regions and the shape of each region is not rectangular. In each sub-image, the division process continues recursively until each region has a homogeneous color distribution or the size of each region becomes smaller than a given threshold value. As a result, a binary tree which roughly represents the color distribution of the image is derived. The tree structure facilitates the evaluation of similarity among images. Another technique in which the position of each colour also plays an important role in the retrieval of images is proposed by Zhang Lei et al [6]. A fast algorithm is proposed which could include several spatial features of color in an image for retrieval. These features are area and position, which mean the zero-order and the first order moments, respectively. By computing the moments of each color region, one can evaluate the similarity of two images according to the weight of each factor. G. Pass et al. [24] propose a different way of incorporating spatial information into the color histogram, Color coherence vectors (CVV), was proposed. The color correlogram [23] was proposed to characterize not only the color distributions of pixels, but also the spatial correlation of pairs of colors.

B. Texture

In computer vision, texture is defined as all what is left after color and local shape have been considered or it is defined by such terms as structure and randomness.The ability to retrieve images on the basis of texture similarity may not seem very useful. But the ability to match on texture similarity can often be useful in distinguishing between areas of images with similar color (such as sky and sea, or leaves and grass). Basically, texture representation methods can be classified into two categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including Fourier power spectra, co-occurrence matrices, shift-invariant principal component analysis (SPCA), Tamura feature, Wold decomposition, Markov random field, fractal model, and multi-resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity. In the context of image retrieval, research is mostly directed toward statistical or generative methods for the characterization of patches. Basic texture properties include the Markovian analysis, dating back to Haralick in 1973, and generalized versions thereof [7], [8]. In retrieval, the property is computed in a sliding mask for localization [9], [10]. Wavelets [12] have received wide attention. They have often been considered for their locality and their compression efficiency. Many wavelet transforms are generated by groups of dilations or dilations and rotations that have been said to have some semantic correspondent. The lowest levels of the wavelet transforms [12], [11] have been applied to texture representation [13], [14] sometimes in conjunction with Markovian analysis [15]. Other transforms have also been explored, most notably fractals [16]. A solid comparative study on texture classification from mostly transform-based properties can be found in [17]. The Tamura features, including coarseness, contrast, directionality, linelikeness, regularity, and roughness, are designed in accordance with psychological studies on the human perception of texture. The first three components of Tamura features have been used in some early wellknown image retrieval systems, such as QBIC [19, 20] and Photobook .Wold decomposition [26, 27] provides another approach to describing textures in terms of perceptual properties. The Gabor filter has been widely used to extract image features, especially texture features . G. Pass et al. [18] propose a novel method to describe spatial features in a more accurate way. Additionally, this model is invariant to scaling, rotation and shifting. In the proposed method segmentations are objects of the images and all images are segmented into several pieces and ROI (Region of Interest) technique is applied to extract the ROI region to enhance the user interaction.

C. *Shape*

The ability to retrieve by shape is perhaps the most obvious requirement at the primitive level. Unlike texture, shape is a fairly well-defined concept – and there is considerable evidence that natural objects are primarily recognized by their shape A number of features characteristic of object shape (but independent of size or orientation) are computed for every object identified within each stored image. Queries are then answered by computing the same set of features for the query image, and retrieving those stored images whose features most closely match those of the query. Two main types of shape feature are commonly used – global features such as aspect ratio, circularity and moment invariants and local features such as sets of consecutive boundary segments. Alternative methods proposed for shape matching have included elastic deformation of templates, comparison of directional histograms of edges extracted from the image. Queries to shape retrieval systems are formulated either by identifying an example image to act as the query, or as a user-drawn sketch. The state-of-art methods for shape description can be categorized into either boundary-based (rectilinear shapes, polygonal approximation , finite element models, and Fourier-based shape descriptors) or region-based methods (statistical moments). A good shape representation feature for an object should be invariant to translation, rotation and scaling.

IV. SIMILARITY MEASURES AND INDEXING SCHEMES

A. *Similarity/Distance Measures*

Instead of exact matching, content-based image retrieval calculates visual similarities between a query image and images in a database. Accordingly, the retrieval result is not a single image but a list of images ranked by their similarities with the query image. Many similarity measures have been developed for image retrieval based on empirical estimates of the distribution of features in recent years. Different similarity/distance measures will affect retrieval performances of an image retrieval system significantly. The different distance measures used for matching are as follows: 1. Minkowski-Form Distance 2. Quadratic Form (QF) Distance 3. Mahalanobis Distance 4. Kullback-Leibler (KL) Divergence and JeffreyDivergence (JD) 5. Euclidean distance

B. *Indexing Scheme*

Another important issue in content-based image retrieval is effective indexing and fast searching of images based on visual features. Because the feature vectors of images tend to have high dimensionality and therefore are not well suited to traditional indexing structures, dimension reduction is usually used before setting up an efficient indexing scheme.One of the technique commonly used for dimension reduction is PCA. The QBIC system uses PCA to reduce a 20-dimensional shape feature vector to two or three dimensions [19] [20]. Apart from PCA and KL transformation, neural network has also been demonstrated to be a useful tool for dimension reduction of features. After dimension reduction, the multi-dimensional data are indexed. A number of approaches have been proposed for this purpose, including R-tree (particularly, R-tree), linear quad-trees, K-d-B tree and grid files. Most of these multi-dimensional indexing methods have reasonable performance for a small number of dimensions (up to 20), but explore exponentially with the increasing of the dimensionality and eventually reduce to sequential searching. Furthermore, these indexing schemes assume that the underlying feature comparison is based on the Euclidean distance, which is not necessarily true for many image retrieval applications. One attempt to solve the indexing problems is to use hierarchical indexing scheme based on the Self-Organization Map (SOM) proposed in.

V. USER INTERACTION

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.

A. *Query Specification*

Specifying what kind of images a user wishes to retrieve from the database can be done in many ways. Commonly used query formations are: category browsing, query by concept, query by sketch, and query by example. Category browsing is to browse through the database according to the category of the image. For this purpose, images in the database are classified into different categories according to their semantic or visual content . Query by concept is to retrieve images according to the conceptual description associated with each image in the database. Query by sketch and query by example is to draw a sketch or provide an example image from which images with similar visual features will be extracted from the database.

B. *Relevance Feedback*

Human perception of image similarity is subjective, semantic, and task-dependent. Although content-based methods provide promising directions for image retrieval, generally, the retrieval results based on the similarities of pure visual features are not necessarily perceptually and semantically meaningful.

In addition, each type of visual feature tends to capture only one aspect of image property and it is usually hard for a user to specify clearly how different aspects are combined. To address these problems, interactive relevance feedback, a technique in traditional text-based information retrieval systems, was introduced. With relevance feedback, it is possible to establish the link between high-level concepts and lowlevel features. For a given query, the system first retrieves a list of ranked images according to a predefined similarity metrics. Then, the user marks the retrieved images as relevant (positive examples) to the query or not relevant (negative examples). The system will refine the retrieval results based on the feedback and present a new list of images to the user. Hence, the key issue in relevance feedback is how to incorporate positive and negative examples to refine the query and/or to adjust the similarity measure.

VI. Conclusion

Content Based Image Retrieval is 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. The impetus behind content-based image retrieval is given by the wide availability of digital sensors, the Internet, and the falling price of storage devices. Given the magnitude of these driving forces, it is to us that content-based retrieval will continue to grow in every direction: new audiences, new purposes, new styles of use, new modes of interaction, larger data sets, and new methods to solve the problems. Some fundamental techniques for content-based image retrieval, including visual content description, similarity/distance measures, indexing scheme, user interaction were introduced. General visual features most widely used in content-based image retrieval are color, texture, shape, and spatial information. There are various ways to calculate the similarity distances between visual features. some basic metrics, including the Minkowskiform distance, quadratic form distance, Mahalanobis distance, Kullback-Leibler divergence and Jeffrey divergence were listed. Up to now, the Minkowski and quadratic form distance are the most commonly used distances for image retrieval. Efficient indexing of visual feature vectors is important for image retrieval. To set up an indexing scheme, dimension reduction is usually performed first to reduce the dimensionality of the visual feature vector. Commonly used dimension reduction methods are PCA, ICA, Karhunen-Loeve (KL) transform, and neural network methods. After dimension reduction, an indexing tree is built up.

The most commonly used tree structures are R-tree, quad-tree, K-d-B tree, etc. Image retrieval systems rely heavily on user interaction. On the one hand, images to be retrieved are determined by the user's specification of the query. On the other hand, query results can be refined to include more relevant candidates through the relevance feedback of users. Updating the retrieval results based on the user's feedback can be achieved by updating the images, the feature models, the weights of features in similarity distance, and select different similarity measures. Although content-based retrieval provides an intelligent and automatic solution for efficient searching of images, the majority of current techniques are based on low level features OR current techniques are primarily based on low level features. The similarity measures between visual features do not necessarily match human perception. Users are interested in are semantically and perceptually similar images, the retrieval results of lowlevel feature based retrieval approaches are generally unsatisfactory and often unpredictable. Although relevance feedback provides a way of filling the gap between semantic searching and low-level data processing, this problem remains unsolved and more research is required.

References

[1] A.W.M. Smeulders, M.Worring, S. Santini, A. Gupta and R. Jain,"Content-Based Image Retrieval at the End of the Early Years", Pattern Analysis and Machine Intelligence, IEEE Transactions on, Vol. 22, No. 12. (06 December 2000), pp. 1349- 1380
[2] J. M. Zachary, Jr. and S. S. Iyengar, "Content Based Image Retrieval Systems", 1999 IEEE Symposium on Application - Specific Systems and Software Engineering and Technology H. Yamamoto, H. Iwasa, N. Yokoya, and H. Takemura, "ContentBased Similarity Retrieval of Images Based on Spatial Color Distributions", ICIAP '99 Proceedings of the 10th International Conference on Image Analysis and Processing
[3] G. Pass and R. Zabih. Histogram refinenement for content based image retrieval, In Proc. 3rd IEEE Workshop on Applications of Computer Vision, pp.96-102, Dec. 1996.
[4] W.Hsu,T.Chua,H.Pung. An integrated color-spatial approach to content-based image retrieval, In ACM Multimedia Conference,pp.305-313,1995.
[5] Cbir method based on color-spatial feature, Zhang Lei, Lin Fuzong, Zhang Bo, State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[6] S. Krishnamachari and R. Chellappa, ªMultiresolution GaussMarkov Random Field Models for Texture Segmentation,º IEEE Trans. Image Processing, vol. 6, no. 2, 1997.
[7] G.L. Gimel'farb and A.K. Jain, ªOn Retrieving Textured Images from an Image Database,º Pattern Recognition, vol. 29, no. 9, pp. 1,461-1,483, 1996.
[8] C.C. Gottlieb and H.E. Kreyszig, ªTexture Descriptors Based on Co-Occurrences Matrices,º Computer Vision, Graphics, and Image Processing, vol. 51, 1990.
[9] H.C. Lin, L.L. Wang, and S.N. Yang, ªColor Image Retrieval Based on Hidden Markov Models,º IEEE Trans. Image Processing, vol. 6, no. 2, pp. 332-339, 1997.
[10] C.K. Chui, L. Montefusco, and L. Puccio, Wavelets: Theory,Algorithms, and Applications. Academic Press, 1994.
[11] I. Daubechies, Ten Lectures on Wavelets. Philadelphia: SIAM, 1992.
[12] A. Laine and J.,vol. 15, no. 11, pp. 1,186-1,191, Nov. 1993. Fan, Texture Classification by Wavelet Packet

Signature, IEEE Trans. Pattern Analysis and Machine Intelligence
[13] J.R. Smith and S.F. Chang, Automated Binary Feature Sets for Image Retrieval,º Proc. Int'l Conf. Acoustics, Speech, and Signal Processing, 1996.
[14] H. Choi and R. Baraniuk, Multiscale Texture Segmentation Using Wavelet-Domain Hidden Markov Models,º Proc. 32nd Asilomar Conf. Signals, Systems, and Computers, vol. 2, pp. 1,692-1,697, 1998.
[15] L.M. Kaplan et al., ªFast Texture Database Retrieval Using Extended Fractal Features,º Storage and Retrieval for Image and Video Databases, VI, vol. 3,312, pp. 162-173, SPIE Press, 1998
[16] T. Randen and J.H. Husoy, Filtering for Texture Classification: A Comparative Study,º IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 4, pp. 291-310, Apr. 1999.
[17] ROI Image Retrieval Based on the Spatial Structure of Objects, Weiwen ZOU1, Guocan FENG2, 12Mathematics and Computational School, Sun Yat-sen University, Guangzhou, China, 510275 paper: 05170290.
[18] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by image and video content: The QBIC system." IEEE Computer, Vol.28, No.9, pp. 23-32, Sept. 1995.
[19] W. Niblack et al., "Querying images by content, using color, texture, and shape," SPIE Conference on Storage and Retrieval for Image and Video Database, Vol. 1908, pp.173-187, April 1993.
[20] M. Stricker, and M. Orengo, "Similarity of color images," SPIE Storage and Retrieval for Image and Video Databases III, vol. 2185, pp.381-392, Feb. 1995.
[21] G.Pass, and R. Zabith, "Comparing images using joint histograms," Multimedia Systems, Vol.7, pp.234-240, 1999.

