CONTENT BASED IMAGE RETRIEVAL USING LOCAL FEATURES OF IMAGE

B.Hemanthi[1], A.Obulesh[2], Dr.V.Vijay Kumar[3],Dr.G.Vishnu Murthy[4]
[1] *M.Tech Student,CSE, Anurag Group Of Institutions, Telangana,India*
[2] *Assistant Professor,CSE, Anurag Group Of Institutions, Telangana,India*
[3] *Professor&DEAN, CSE, Anurag Group Of Institutions, Telangana,India*
[4] *Professor&HOD, CSE, Anurag Group Of Institutions, Telangana,India*

ABSTRACT

Recently, digital content has become a significant and inevitable asset for any enterprise and the need for visual content management is on the rise as well. There has been an increase in attention towards the automated management and retrieval of digital images owing to the drastic development in the number and size of image databases. A significant and increasingly popular approach that aids in the retrieval of image data from a huge collection is called Content-based image retrieval (CBIR).

Content-based image retrieval has attracted voluminous research in the last decade paving way for development of numerous techniques and systems besides creating interest on fields that support these systems. CBIR indexes the images based on the features obtained from visual content so as to facilitate speedy retrieval. This report discuss the state of the art of the content based image retrieval highlighting the main components and reviewing various approaches employed at each stage, while enhancing the main challenges and key contributions.

Keyword: - *content, image retrieval, color, shape, texture, indexing, similarity*

---

1. INTRODUCTION

Content-based image retrieval (**CBIR**), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases.

"Content-based" means that the search will analyze the actual contents of the image. The term 'content' in this context might refer colors, shapes, textures, or any other information that can be derived from the image itself. Without the ability to examine image content, searches must rely on metadata such as captions or keywords. Such metadata must be generated by a human and stored alongside each image in the database.

An image retrieval system returns a set of images from a collection of images in the database to meet users' demand with similarity evaluations such as image content similarity, edge pattern similarity, color similarity,etc. An image retrieval system offers an efficient way to access, browse, and retrieve a set of similar images in the real-time applications. Several approaches have been developed to capture the information of image contents by directly computing the image features from an image [1].

The image features are directly constructed from the typical Block Truncation Coding (BTC) or half toning-based BTC compressed data stream without performing the decoding procedure. These image retrieval schemes involve two phases, indexing and searching, to retrieve a set of similar images from the database. The indexing phase extracts the image features from all of the images in the database which is later stored in database as feature vector. In the searching phase, the retrieval system derives the image features from an image submitted by a user (as query image) [2].

2.REVIEW OF LITERATURE

Content-Based Image Retrieval Using Features Extraction:

The use of images in human communication is hardly new. Content-based image retrieval (CBIR), a technique for retrieving images on the basis of automatically-derived features such as color, texture and shape. Feature (content) extraction is the basis of content-based image retrieval. In a broad sense, features may include both text-based features (key words, annotations) and visual features (color, texture, shape, faces) [3].

The history of the content-based image retrieval can be divided into three phases:

· The retrieval based on artificial notes.

· The retrieval based on vision character of image contents.

· The retrieval based on image semantic features

All current CBIR systems, allows users to formulate queries by submitting an example of the type of image being sought, though some offer alternatives such as selection from a palette or sketch input. The system then identifies those stored images whose feature values match those of the query most closely, and displays thumbnails of these images on the screen Error! Reference source not found. Some of the more commonly used types of feature used for image retrieval are described below.

3.COLOR RETRIEVAL

The color feature is one of the most widely used visual features in image retrieval. It is relatively robust to background complication and independent of image size and orientation. 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. Several methods for retrieving images on the basis of color similarity have been described in the literature, but most are variations on the same basic idea. 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 matching process then retrieves those images whose color histograms match those of the query most closely. The matching technique most commonly used, histogram intersection, was first developed by Swain and Ballard [1991]. Variants of this technique are now used in a high proportion of current CBIR systems [4].

4. TEXTURE RETRIEVAL

Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity [5]. 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). A variety of techniques has been used for measuring texture similarity. Alternative methods of texture analysis for retrieval include the use of Gabor filters and fractals. Texture queries can be formulated in a similar manner to color queries, by selecting examples of desired textures from a palette, or by supplying an example query image. The system then retrieves images with texture measures most similar in value to the query. A recent extension of the technique is the texture thesaurus, which retrieves textured regions in images on the basis of similarity to automatically-derived code words representing important classes of texture within the collection [4].

4.1 Shape retrieval

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. Shape matching of three-dimensional objects is a more challenging task – particularly where only a single 2-D view of the object in question is available [4]. 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 [5].

5.REPORT ON PRESENT INVESTIGATION

5.1 Problem Definition

Problems with traditional methods of image indexing [Enser,1995] have led to the rise of interest in techniques for retrieving images on the basis of automatically-derived features such as color, texture and shape.

A technique for content-based image retrieval (CBIR) for the generation of image content descriptor. In the encoding step, compresses an image block into corresponding quantizers and bitmap image. Two image features are proposed to index an image, namely, color co-occurrence feature (CCF) and bit pattern features (BPF).The CCF and BPF of an image are simply derived from the two quantizers and bitmap, respectively. The ODBTC scheme is not only suited for image compression, because of its simplicity, but also offers a simple and effective descriptor to index images in CBIR system.

1.Color Co-Occurrence Feature (CCF):

The color distribution of the pixels in an image contains huge amount of information about the image contents. The attribute of an image can be acquired from the image color distribution by means of color co-occurrence matrix. This matrix also represents the spatial information of an image. Color Co-occurrence Feature (CCF) can be derived from the color co-occurrence matrix.

2.Bit Pattern Feature (BPF):
BPF characterizes the edges, shape, and image contents. The binary vector quantization produces a representative bit pattern codebook from a set of training bitmap images.

5.2 Methodology

Two image features are proposed to index an image:
1. color co-occurrence feature (CCF)

2. bit pattern features (BPF)

Which are generated directly from the ODBTC encoded data streams without performing the decoding process. The CCF and BPF of an image are simply derived from the two ODBTC quantizers and bitmap. Figure 3.1 shows the Block diagram of the proposed image retrieval method.

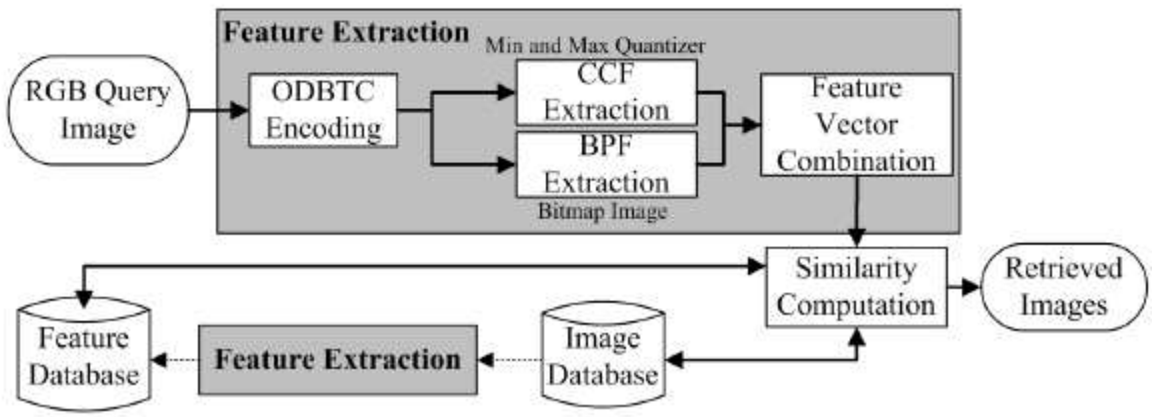

Figure 1 Block diagram of the proposed image retrieval method.

5.3 Ordered- Dither Block Truncation Coding:

The main advantage of the ODBTC image compression is on its low complexity in generating bitmap image by incorporating the Look-Up Table (LUT), and free of mathematical multiplication and division operations on the determination of the two extreme quantizers. The traditional BTC derives the low and high mean values by preserving the first-order moment and second-order moment over each image block, which requires additional computational time. Conversely, ODBTC identifies the minimum and maximum values each image block as opposed to the former low and high mean values calculation, which can further reduce the processing time in the encoding stage. In addition, the ODBTC yields better reconstructed image quality by enjoying the extreme-value dithering effect. At the end of the ODBTC encoding, the bitmap image, bm, the minimum quantizer, Xmin, and maximum quantizer, Xmax ,are obtained and considered as encoded data stream.

5.4 Color Co-occurrence Feature (CCF):

Color Co-occurrence Feature (CCF) can be derived from the color co-occurrence matrix. Figure 2 shows the schematic diagram for deriving the CCF. In the proposed scheme, CCF is computed from the two ODBTC color quantizers. The minimum and maximum color quantizers are firstly indexed using a specific color codebook. The color co-occurrence matrix is subsequently constructed from these indexed values. Subsequently, the CCF is derived from the color co-occurrence matrix at the end of computation. In general, the color indexing process on RGB space can be defined as mapping a RGB pixel of three tuples into a finite subset of codebook index. The CCF calculation is simple, making it more preferable for CBIR task.

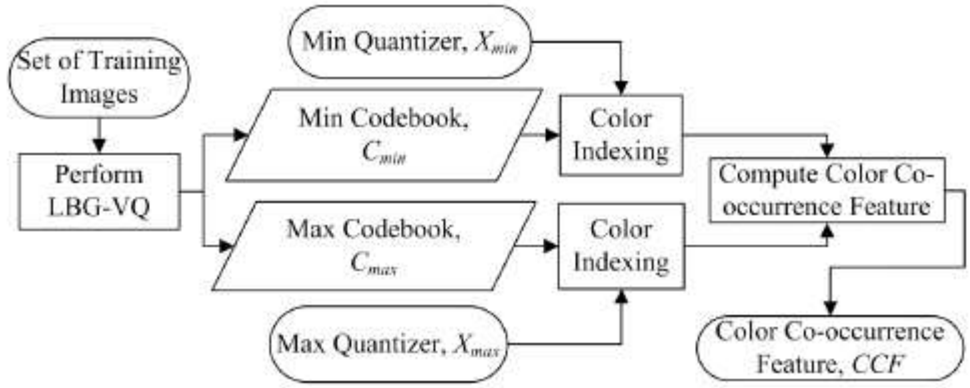

Figure 2 Block diagram for computing the color co-occurrence feature.

5.5 Bit Pattern Feature (BPF)

Another feature, namely Bit Pattern Feature (BPF), characterizes the edges, shape, and image contents. Figure 3.3 shows the schematic diagram for deriving the BPF. The binary vector quantization produces a representative bit pattern codebook from a set of training bitmap images obtained from the ODBTC encoding process. Let Q = {Q1, Q2. . .QNb } be the bit pattern codebook consisting Nb binary code word. These bit pattern codebooks are generated using binary vector quantization with soft centroids, and many bitmap images are involved in the training stage.

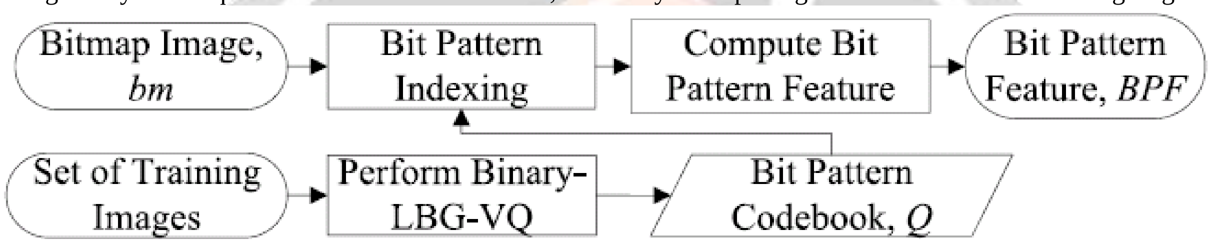

Figure3 Block diagram for computing the bit pattern feature.

The bitmap of each block bm (i, j) is simply indexed based on the similarity measurement between this bitmap and the code word Qq. Subsequently, the BPF is simply derived as the occurrence probability of the bitmap image mapped into the a specific bit pattern codeword Qq. Similar to that of the CCF, the BPF only needs a simple computation, making it suitable for real applications where fast response is required.

5.6 Implementation Plan

Analysis : The proposed image retrieval methods introduce two image features to index an image:

1. color co-occurrence feature (CCF)

2. bit pattern features (BPF)

The similarity between two images (i.e., a query image and the set of images in the database as target image) can be measured using the relative distance measure [6]. The similarity distance plays an important role for retrieving a set of similar images. The query image is firstly encoded with the ODBTC, yielding the corresponding CCF and BPF. The two features are later compared with the features of target images in the database. A set of similar images to the query image is returned and ordered based on their similarity distance score, i.e. the lowest score indicates the most similar image to the query image.

System requirement specification:

a) Software requirement:

 MATLAB 7.0

 MySQL 2000

b) Hardware requirement:

 Pentium 4 2.66 Ghz/AMD Athlon 2800+ processor.

 256MB RAM (Minimum)/512MB RAM (Recommended).

 40 GB+ Hard Disk

6. CONCLUSION

This report reviewed the main components of a content based image retrieval system, including image feature representation, indexing, query processing, and query-image matching and user's interaction, while highlighting the current state of the art and the key-challenges. It has been acknowledged that it remains much room for potential improvement in the development of content based image retrieval system due to semantic gap between image similarity outcome and user's perception. Contributions of soft-computing approaches and natural language processing methods are especially required to narrow this gap. Standardization in the spirit of MPEG-7, which includes both feature descriptor and language annotation for description various entity relationships, is reported as a crucial step in order to increase the interoperability of the various systems.

7. REFERENCES

[1] Content-Based Image Retrieval Using Features Extracted From Half toning-Based Block Truncation Coding Jing-Ming Guo, Senior Member, IEEE, and Heri Prasetyo

[2] Content-Based Image Retrieval Using Error Diffusion Block Truncation Coding Features Jing-Ming Guo, Senior Member, IEEE, Heri Prasetyo,and Jen-Ho Chen

[3] Image Retrieval: Current Techniques, Promising Directions, and Open Issues Yong Rui and Thomas S. Huang Department of ECE & Beckman Institute, University of Illinois.

[4] http://www.engineersgarage.com/contribution/content-based-image-retrieval-matlab-project.

[5] J. R. Smith and S.-F. Chang, Automated binary texture feature sets for image retrieval, in Proc. ICASSP-96,Atlanta, GA, 1996.

[6] G. Qiu, “Color image indexing using BTC,” IEEE Trans. Image Process.,vol. 12, no. 1, pp. 93–101, Jan. 2003.