ORIENTAL JOURNAL OF
COMPUTER SCIENCE & TECHNOLOGY
An International Open Free Access, Peer Reviewed Research Journal
Published By: **Techno Research Publishers, Bhopal, India.**
www.computerscijournal.org

ISSN: 0974-6471
April 2015,
Vol. 8, No. (1):
Pgs. 39-42

An Overview on Content - Based Image Retrieval

H. B. BASANTH KUMAR

Pooja Bhagavat Memorial Mahajana PG Centre, DoS in Computer Applications,
KRS Road, Metagalli, Mysuru-570016, Karnataka, India.

(Received: March 10, 2015; Accepted: March 31, 2015)

ABSTRACT

As the propagation of video and image data in digital form has increased, Content Based Image Retrieval (CBIR) has become a prominent research topic. Therefore an important problem that needs to be addressed is fast retrieval of images from large databases. To find images that are perceptually similar to a query image, image retrieval systems attempt to search through a database. CBIR can greatly enhance the accuracy of the information being returned and is an important alternative and complement to traditional text-based image searching. This paper presents an overview on content-based image retrieval.

Key words: Content – based Image Retrieval, Color space, shape.

INTRODUCTION

The term “content-based image retrieval” seems to have originated in 1992 when it was used by T. Kato to describe experiments into automatic retrieval of images from a database, based on the colors and shapes present. Since then, the term has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features.

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 techniques to the image retrieval problem, that is, the problem of searching for digital images in large databases. “Content-based” means that the search analyzes the contents of the image rather than the metadata such as keywords, tags, or descriptions associated with the image. The term “content” in this context might refer to colors, shapes, textures, or any other information that can be derived from the image itself. CBIR is desirable because most web-based image search engines rely purely on metadata and this produces a lot of garbage in the results. Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results[1].

Image Classification

Intensity Images

It represents an image as a matrix where every element has a value corresponding to how bright / dark the pixel at the corresponding position should be colored. There are two ways to represent the brightness of the pixel. They are

Double type

Assign a real number between 0 and 1 for each pixel.

Uint 8

Assign an integer between 0 and 255

Indexed Images

An indexed image stores an image as two matrices. The first matrix has the same size as the image and one number for each pixel. The second matrix is called the color map and its size may be different from the image.

Scaled Indexed Images

A scaled indexed image uses matrix values.

Binary Images

This image format stores an image as a matrix but can only a color pixel either black or white. It assigns a value 0 for black and a value 1 for white[2].

Content – based Image Retrieval

In CBIR each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps.

Feature Extraction

The first step in this process is to extract the image features to a distinguishable extent.

Matching

The second step involves matching these features to yield a result that is visually similar.

Content – based Image Retrieval Process

Basic idea behind CBIR is that, when building an image database, feature vectors from images (the features can be color, shape, texture, region or spatial features, features in some compressed domain, etc.) are to be extracted and then store the vectors in another database for future use. When given a query image its feature vectors are computed. If the distance between feature vectors of the query image and image in the database is small enough, the corresponding image in the database is to be considered as a match to the query. The search is usually based on similarity rather than on exact match and the retrieval results are then ranked accordingly to a similarity index[3].

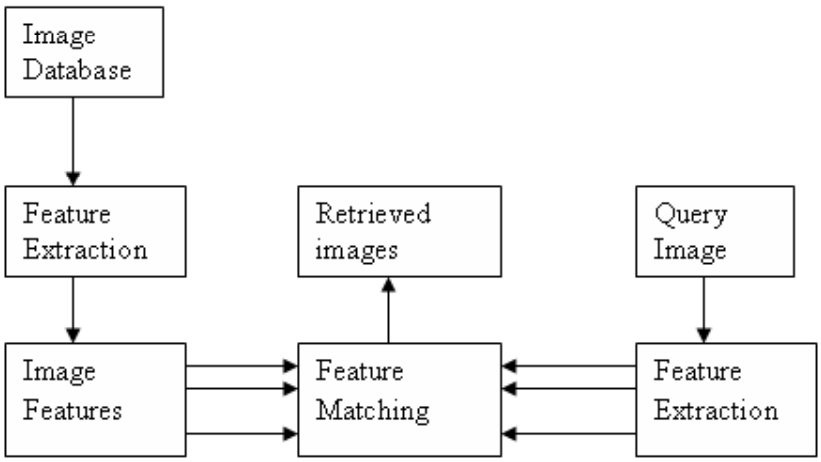

Fig. 1. Process involved in Content – based Image Retrieval

Applications

- Entertainment
- Fashion and publishing
- Architectural and engineering design
- Medical diagnosis
- Geographical information and remote sensing
- Crime detection
- Searching a particular image on the web[2].

Different types of CBIR

General

Here, matching a query image to an arbitrary collection of images.

Application Specific

In this step, query image is matched to a collection of images of a specific type.
Ex: finger prints, face detection etc.

Feature extraction

There are 2 types of visual feature in CBIR. They are

Primitive features

Primitive features which includes color texture and shape.

Domain specific features

Which are applications specific such as X-ray Images, finger prints etc[2].

Color

Color is most widely used low-level feature extraction. It is represented in the form of color components. Color feature includes

1. Conventional color histogram (CCH).
2. Fuzzy color histogram (FCH).
3. Color correlogram (CC) and a more recent.
4. Color – shape based feature.

Each color based feature extraction follows the following four methods.

- Selection of the color space.
- Quantization of the color space.
- Extraction of the color feature.
- Derive an appropriate distance function[4].

Texture

A variety of techniques have been used for measuring texture such as co-occurrence matrix, Fractals, Gabor filters, variations of wavelet transform. Further techniques have also been developed for describing the local patterns using texture spectrum, for characterization of texture using a composition of edge information and co-occurrence matrix properties. The identification of specific textures in an image is achieved primarily by modeling texture as a two-dimensional gray level variation[5].

Shape

Shape is one of the important features and contains the most attractive visual information for human perception. We use the term shape to refer to the information that can be deduced directly from images and that cannot be represented by color or texture; as such, shape defines a complementary space to color and texture. Shape representations techniques used in similarity retrieval are generally characterized as being region based and boundary based. Boundary based shape representation only uses the outer boundary of the shape. Methods that extract region based features take into account all the pixels within the shape. They map each shape into a fixed sized grid or circle to achieve scale, rotation and translation invariance. This normalized shape is viewed as a probability density of a two-dimensional variable, from which orthogonal moments that describe some global properties of the shape can be computed. Low level visual features are used for representation and retrieval of images[6].

CONCLUSION

Content Based Image Retrieval (CBIR) is a significant and increasingly popular approach that helps in the retrieval of image data from a huge collection. Image representation based on certain features helps in retrieval process**.** This paper has focused on the process involved in content-based image retrieval (CBIR), different types of CBIR, feature extraction and applications of CBIR.

REFERENCES

1. http://en.wikipedia.org/wiki/Content-based_image_retrieval
2. Mohd. Danish, Ritika Rawat, Ratika Sharma, "A Survey: Content Based Image Retrieval Based on Color, Texture, Shape & Neuro Fuzzy", *Intl. Jrnl. of Engineering Research and Application,* ISSN: 2248 – 9622, **3**(5) pp.839 – 844 (2013).
3. P.S.Suhasini , Dr. K.Sri Rama Krishna, Dr. I. V. Murali Krishna, "CBIR Using color histogram processing", *Jrnl. of Theoretical and Applied Information Technology,* **6**(1), pp. 116 – 122.
4. Michele Saad, "Content–Based Image Retrieval–Literature Survey", EE 381K: Multi – dimensional Digital Signal Processing, March 18 (2008).
5. Dr. H.B.Kekre, Sudeep D. Thepade, Tanuja K. Sarode and Vashali Suryawanshi, "Image Retrieval using Texture Features extracted from GLCM, LBG and KPE", *Intl. Jrnl. of Computer Theory and Engineering,* **2**(5), October, 2010, pp. 1793-8201.
6. Sh. Akbarpour, "A Review On Content Based Image Retrieval In Medical Diagnosis", *Intl. Jrnl. on Technical and Physical Problems of Engineering (IJTPE),* **5**(15) pp. 148-153 (2013).