ISSN: 2456 – 3137
International Journal of Computational Research and Development
Impact Factor 4.775, Special Issue, January - 2017
International Conference on Smart Approaches in Computer Science Research Arena
On 5th January 2017 Organized By
Department of Computer Science, Sri Sarada College for Women (Autonomous), Salem, Tamilnadu

CONTENT BASED IMAGE RETRIEVAL TECHNIQUES: SURVEY

K. Dhanalakshmi* & Dr. S. Lakshmi Prabha**
* Research Scholar, Department of Computer Science, Periyar University, Salem, Tamilnadu
** Assistant Professor, Department of Computer Science, Government Arts College for Women, Salem, Tamilnadu

Cite This Article: K. Dhanalakshmi & Dr. S. Lakshmi Prabha, "Content Based Image Retrieval Techniques: Survey", International Journal of Computational Research and Development, Special Issue, January, Page Number 35-37, 2017.

Abstract:
In recent years, an effective method of image searching and retrieval has been increased. Image retrieval is a primary technique for biomedicine, computer-aided diagnosis, web image searching and plant identification in agricultural images. Content based image retrieval (CBIR) is the process of browsing, searching and retrieving images from a large database of digital images. In this paper analyses the performance of various techniques used in CBIR systems, it deals with the image content itself such as color, texture and shape.

Key Words: CBIR, Color Based Image Retrieval, Texture Based Image Retrieval, Shape Based Image Retrieval & Similarity Matching.

1. Introduction:

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 suitable because searches that rely purely on metadata are dependent on annotation quality and completeness. Having humans manually annotate images by entering keywords or metadata in a large database can be time wasting and may not capture the keywords desired to about the image. The CBIR is a method where relevant images from large scale image databases are searched according to user interest. It has become an active and quick advancing research area since last two decades. During the past decade, remarkable progress has been made in both theoretical research and system development. However, still there is many challenging research problems that continue to attract researchers from multiple disciplines [1]. Early methods used for image retrieval were not based on visual features but based on the textual annotation of images. It means that images were first annotated with text and then searched using a text based approach from traditional database systems. [1]. However, the performance of traditional approach to image retrieval is very sensitive to the keywords employed by the user and the system. Main idea behind CBIR is to analyze image information by low level features of an image, such as color, texture, shape and color to create feature vectors of an image as its index. That features are stored in an image feature database for future use [2]. When a query image is given, the features of the query image are fetched to match the features in the feature database by an algorithm, so that a group of similar images to the query image can be returned as the retrieval images [3]. The most common method for CBIR is using similarity or distance measure. An image distance measures compares the similarity of two images in various dimensions. A distance of 0 signifies an exact match with the query images, a value greater than 0 indicates various degrees of similarities between the images. Search results can be sorted based on their distance to the queried images. Basically, one of the key points for realizing the image retrieval is to fetch appropriate feature vectors to represent image content correctly [3].

2. Objective:

The CBIR is to analyze image information by low level features of an image such as color, texture, shape and color layout etc., and to create feature vectors of an image as its index. When a query image is given, the features of the query image are extracted to match the features in the feature database by a pre-established algorithm, so that a group of similar images to the query image can be returned as the retrieval images.

3. Features for Extraction Methods:

In this step visual information is extracts from the image and saves them as features vectors in a features database. Here we will discuss about only some of the image feature such as color, texture, shape.

A. Color Feature Based Image Retrieval Methods: Color is one of the most widely used low-level visual features and is invariant to image size and orientation Color histogram is invariant to orientation and scale and this makes it powerful in image classification. Useful in classifying objects based on color. Typical color images consist of three color planes: red, green and blue. They can be treated as three separate gray-scale images. Color and Edge Directivity Descriptor Extraction (CEDD) technique, which extracts the color and edge of an image in a database. The following feature extraction techniques are used to extract the low level color feature in the database. Color layout Descriptor (CLD) extraction technique which extracts the color and texture value of an image in a database. Edge Histogram Descriptor (EHD) extraction technique which extracts an image shape in a database. Auto color correlogram extraction technique, which extracts the spatial color value in a database. This technique is compared with the average precision value of the existing techniques, which gives the average precision value is 58%. CEDD is effective and efficient technique for image indexing and image retrieval [4]. There are two features of colors that is global and black color histogram based CBIR. A color histogram is a frequency statistic for different colors in a certain color space. The advantage is that it describes the global color distribution for images. It is especially fit for those images difficult to segment and leave spatial locations. However, its disadvantage is that it cannot describe the local distribution of the image in color space and the spatial position of all color. It means that the color histogram cannot give specific objects or things in the image. The color space needs to be divided into several small ranges in order to compute the color histogram. All interval is considered as a bin.

ISSN: 2456 – 3137
International Journal of Computational Research and Development
Impact Factor 4.775, Special Issue, January - 2017
International Conference on Smart Approaches in Computer Science Research Arena
On 5th January 2017 Organized By
Department of Computer Science, Sri Sarada College for Women (Autonomous), Salem, Tamilnadu

Thus, the color is quantized. The color histogram can be calculated in counting pixels where the colors fall into each interval. Color features consist of global color histogram and block color histogram [5].

The color auto-corrlogram, which captures the spatial correlation between the similar colors at a distance 1 is computed for H and S components of an image. Since, the smallest distance gave the most detailed local properties of the image, the proposed system fix the distance to 1[7].

B. Texture Based Image Retrieval Methods: Texture is also one of the most used low level visual features that refer to innate surface properties of an object and their relationship to the surrounding environment. Texture is a very important characteristics for the analysis of many types of images that appears everywhere in nature like natural images, remote sensing images and medical images. Texture can be defined as superficial phenomenon of human visual systems of natural objects. Texture can be attributed to almost everything in nature and also its texture structure of any image is include repeated pattern of all most all of the parts. Texture is commonly known as 'texels'. Texture can be identified by everyone but it is not easy to define. Texture does not take place over a point but it rather take place over a region. Texture can be analyzed by quantitative and qualitative analysis [8].

Tamura Texture Feature: According to quantitative analysis one of the first descriptions given by the Tamura proposed six textural properties and gave descriptions common over all texture patterns in Broadtz's photographic images. These are six different texture features given by tamura Coarseness, Contrast, Directionality, Line-Likeness, Regularity and Roughness [8].

Haralick Texture Feature: Gray Level Co-occurrence Matrix (GLCM) is a statistical method for evaluat texture features that consider the spatial relationship of pixels, also known as Gray Level Spatial Dependence. In this, a GLCM matrix is created by calculating how frequently a pixel with the intensity value i take place in a specific spatial relationship to a pixel with the value j. GLCM consists of frequencies at which two pixels are separated by a certain vector occur in the image. GLCM properties by which the distribution in the matrix will depends on the distance and angular or directions like horizontal, vertical, diagonal, anti-diagonal relationship between the pixels. Haralick [16] proposed 28 kinds of textural features each extracted from the Gray Level Co-occurrence Matrix. Among them five features Contrast, Correlation, Entropy, Energy and Homogeneity [8].

C. Shape Based Image Retrieval Methods: Shape features depend on a silhouette (outline) of an image. Object shape features can also be used to give powerful information, because humans can find out the objects only from their shapes. Basically, the shape have semantic information of object, and it is different from other elementary visual features are color or texture features. Shape representation methods include Fourier descriptors, polygonal approximation, invariant moments, B-splines, deformable templates and curvature scale space (CSS).

Hu-Moment Shape Features: Seven properties related to connected region that are invariant to rotation, scaling, and translation (RTS) and are also known as Algebraic Moment Invariants. Moment invariants that are computed from each of the window are used to form feature vectors. They define simply calculated set properties of region that can be used for class identification and also identification of shape, and this classic technique for generating invariants in terms of algebraic was originally proposed by Hu [8]. The author discussed about the two shape features such as region and contour Shape [8].

Region Shape: The shape of an object may made up of a single region or a set of regions as well as some holes in that object. Since the Region Shape descriptor, based on the moment invariants makes use of all pixels constituting the shape within a frame, it can tell of any shape. The shape does not have to be a simple shape with a single connected region, but it can also be a complex shape consisting of holes in the object or several disjoint regions. The benefit of the Region Shape descriptor are that in addition to its ability to tell of diverse shapes efficiently it is also robust to minor deformations along the boundary of the object. The feature extraction and matching job are straightforward. Since they have low order of computational complexities they are suitable for shape tracking in the video sequences [9].

Contour Shape: The Contour Shape descriptor captures characteristics of a shape based on its contour. It depend on the so-called Curvature Scale Space CSS representation, which captures perceptually meaningful features of the shape. The descriptor essentially represents the points of high curvature along the contour (position of the point and value of the curvature). This representation has a number of essential properties, namely, it captures characteristic features of the shape, enabling efficient similarity based retrieval. It is also robust to non-rigid motion [9].

4. Similarity Matching:

In this step involves the matching of the features is visually similar with the use of similarity measure method called as Distance method. Some of the distances method available such as

- ✓ Euclidean distance
- ✓ City Block Distance
- ✓ Canberra Distance

The performance of a CBIR system mainly depends on the image representation and similarity matching function employed .Based Image recovery system which evaluates the similarity of each image in its data accumulate to a query image in terms of various visual features and return the image with desired range of similarity[10]. Euclidian distance to find the difference between the images. Euclidian Distance matrix is mostly used for similarity measurement in context retrieval of image from database because of its higher accuracy and effectiveness. It measures the distance between the two feature vectors of images by calculating the square root of the sum of the squared absolute differences and is calculated and is denoted by ED [8].

An efficient KNN and SVM to extract features according to data set using Auto calculate the feature weight by neural network [10].

5. Conclusion:

The main goal in image analysis is to extract useful information for solving application-based problems. The performance of traditional approach to image retrieval is very sensitive to the keywords. In future we will try to improve the performance of the new methodologies to improve the performance of image retrieval process.

6. References:

1. F. Long, H. Zhang and D. D. Feng, "Fundamentals of content-based image retrieval", in Multi-media Information Retrieval and Management, Sprin-ger Berlin Heidelberg, (2003), pp. 1-26.
2. J. Kang and W. Zhang, "A framework for image retrieval with hybrid features", Control and Decision Conference (CCDC), 2012 24th Chinese, IEEE, (2012), pp. 1326-1330.
3. Savita Gandhani and Nandini Singhal, "Content Based Image Retrieval: Survey and Comparison of CBIR System based on Combined Features" Image Processing and Pattern Recognition, IJSIP, (2015), pp.417-422.
4. K. Jayanthi and M. Karthikeyan, "Color and Edge Directive Descriptor Feature Extraction Technique for Content Based Image Retrieval System", Middle-East Journal of Scientific Research ,(2015)
5. J. Yue, Z. Li, L. Liu and Z. Fu. "Content-based image retrieval using color and texture fused features", Mathematical and Computer Modelling, vol. 54, no. 3, (2011), pp. 1121-1127.
6. X. Y. Wang, Y. J. Yu and H. Y. Yang, "An effective image retrieval scheme using color, texture and shape features", Computer Standards & Interfaces, vol. 33, (2011), pp. 59-68.
7. K.Seetharaman and S.Sathiamoorthy, "A Framework for Color Image Retrieval Using Full Range Gaussian Morkov Random Field Model and Multi-Class SVM Learning Approach", IJIRA, (2014).
8. Neelima Bagri and Punit Kumar Johari, "A Comparative Study on Feature Extraction using Texture and Shape for Content Based Image Retrieval", IJAST, (2015), pp.41-52.
9. Satish Tunga, D. Jayadevappa and C. Gururaj, "A Comparative Study of Content Based Image Retrieval Trends and Approaches", IJIP, (2015).
10. Sandeep Singh and Er. Rachna Rajput, "Content Based Image Retrieval using SVM, NN and KNN Classification", IJARCCE, (2015).