www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume - 3 Issue - 8 August, 2014 Page No. 7680-7684

QBIC, MARS and VIPER: A Review on Content Based Image Retrieval Techniques

Amit Singla[1] , Meenakshi Garg [2]

[1]Guru Kashi University, Department of CSE,
Talwandi Sabo, Punjab, India
amitsingla2212@gmail.com

[2]Guru Kashi University, Department of CSE,
Talwandi Sabo, Punjab, India
manu8708@gmail.com

Abstract: ***With the widespread of the network and development of Image technology, the older information retrieval techniques do not meet today's demand. Recently, the content-based image retrieval has become the most developing topic and the techniques of content-based image retrieval have acquired great attention. In this paper, the basic components of content-based image retrieval system are discussed. Image retrieval methods based on color, texture, shape and semantic image are studied discussed and compared. Other related techniques such as relevance feedback and performance evaluation are also discussed. In the end of paper the problems and challenges are fetched out of each available technique. In recent areas of commerce, government, academics, and hospitals, large collections of digital images are being created and stored. Large amount of the digital image collections are the product of digitalizing existing collections of hardcopy photographs, diagrams, drawings, paintings, and prints. Usually, the only way of searching these collections was by keyword indexing, or simply by browsing related words. Digital image database however, open the way to content-based searching. In this paper we review some technical aspects of current content-based image retrieval systems.***

Keywords: Image retrieval, content-based image retrieval, color, texture, shape and semantic-based image retrieval.

1. Introduction

Human being precieve images, sounds and any other informations by sight, hearing, perception and analysis. Humans judge similarity of images and sounds according to their semantic contents, for instance the searching for a car's picture is based on its appearence characterstics or other visible contents. So the retrieval methods based on text or keywords for the digital multimedia apparently can't meet the demand that human being get multimedia information exactly [1].

With the growth of multimedia information available on the Web , human beings' thirst for accurate, precise and fast retrieval, we will penetrate deeply into this area. In this paper, we discuss techniques which are based on features extracted from the contents of multimedia information.These retrieval techniques are becoming the focus of the academic research as well.

2. RELATED WORK

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.

The image retrieval that is based on artificial notes labels images by using text first. In fact it has already changed image retrieval into traditional keywords retrieval. Now, there are two problems remaining in this method. Firstly, it leads too heavy usablity of the CPU time and Bandwidth usage. Secondly, it remains subjective and uncertain. Also the image retrieval that is based on artificial notes still remains insufficient, the further study that adapts vision image features has been come up and become the main study. The characterstic of this method is image feature extraction, whether the retrieval is good or not depends on the accuracy of the features extracted. So the research based on vision features is becoming the topic of focus in the academic community. The features of vision can be classified by semantic hierarchy into two levels: Middle-level features and Low-level features. Low-level features includes color, texture and inflexion. Middle-level features involves shape description and object feature [1, 3, 5, 10].

2.1 TEXT-BASED APPROACH

Indexing by input of keywords and descriptions. Eg. (Google, Yahoo, etc.)

• Advantages:
– Easy to implement
– Fast retrieval
– Web image search (surrounding text)

• Disadvantages:
– Manual annotation is not always available
– Manual annotation is impossible for a large DB
– Manual annotation is not accurate
– A picture is worth a thousand words
– Surrounding text may not describe the image

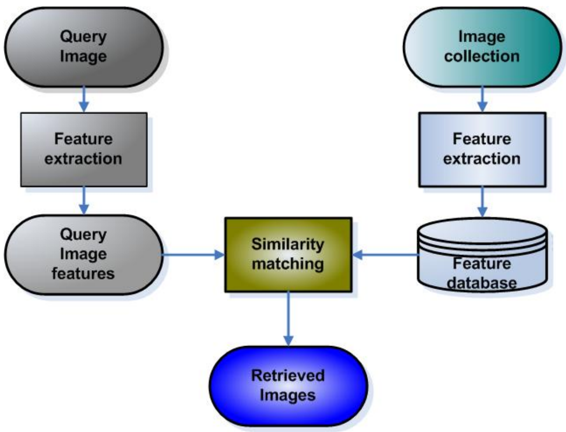

Figure 2.1: Content Based Image Retrieval basic structure.

2.2 CONTENT-BASED APPROACH

Index images using reference images .

• Advantages
– Visual features, such as color, texture, and shape information of images are extracted automatically.
– Similarities of images are based on the distances between features [6]

3. VARIOUS CBIR SYSTEMS

3.1 QBIC or Query by Image Content

It was developed by IBM, Almaden Research Centre [5,7], which allow users to graphically pose and refine queries based on multiple visual properties such as colour, texture and shape. It supports queries based on input images, user-constructed sketches, and selected colour and texture patterns present in an image.

3.1.1 Features

Color features computed are: the 3D average color vector of an object or the whole image in RGB, YIQ, Lab, and Munsell color space and a 256-dimensional RGB color histogram. If x is an n dimensional color histogram and C= [c_1 c_2 ......c_n] is a 3xn matrix whose columns represent the RGB values of the n quantized colors, the average color vector X_{avg} is Cx. The texture features used in QBIC are modified versions of the coarseness, contrast, and directionality features. The shape features of QBIC consist of shape area, circularity, and eccentricity, major axis orientation and a set of algebraic moment invariants. The major axis orientation and the eccentricity are computed from the second order covariance matrix of the boundary pixels: the major axis orientation as the direction of the largest eigen vector and eccentricity as the ratio of the smallest eigen value to the largest one. For the database images, these shape features are extracted for all the object contours, semi automatically computed in the database population step. In this process, the user enters an approximate object outline, which is automatically aligned with the nearby image edges, using the active contours technique. In this object identification step, the user can also associate text to the outlined objects. The 18 algebraic moment invariants are the eigen values of the matrices $M_{[2,2]}$, $M_{[2,3]}$ x $M_{[3,2]}$, $M_{[3,3]}$, $M_{[3,4]}$ x $M_{[4,3]}$, $M_{[4,4]}$, $M_{[4,5]}$ x $M_{[5,4]}$, where the elements of $M_{[i,j]}$ are scaled factors of the central moments. QBIC also implemented a method of retrieving images based on a rough user sketch. For this purpose, a reduced binary map of edge points represents images in the database. This is obtained as follows: first, the color image is converted to a single band luminance; using a Canny edge detector, the binary edge image is computed and is next reduced to size 64 × 64 . Finally this reduced image is thinned.

3.1.2 Querying

QBIC allows queries based on example images, user-constructed sketches or/and selected color and texture patterns. In the last case, the user chooses colors or textures from a sampler. The percentage of a desired color in an image is adjusted by moving sliders. Matching For the average color, the distance between a query object and database object is a weighted Euclidean distance, where the weights are the inverse standard deviation for each component over the samples in the database. In matching two color histograms, two distance measures are used: one low dimensional, easy to compute (the average color distance) and one much more computationally expensive (the quadratic histogram distance). The first one (which is computed for all the images in the database) acts as a filter, limiting the expensive matching computation to the small set of images retrieved by the first matching. The 36 average color distance is

$$d_{avg}^2(x,y) = (x_{avg} - y_{avg})^t (x_{avg} - y_{avg})$$

The histogram quadratic distances is given by

$$d_{hist}^2(x,y) = (x - y)^t A (x - y)$$

Where the symmetric color similarity matrix A is given by

$$a_{ij} = 1 - d_{ij} / d_{\max}$$

with d_{ij} being the L_2 distance between the colors i and j in the RGB space and $d_{\max} = \max_{i,j} d_{ij}$. The texture distance is a weighted Euclidean distance, with the weighting factors being the inverse variances for each of the three texture components over the entire database. Two shapes are matched also by a similar weighted Euclidean distance between shape feature vectors. In a query by sketch, after reducing the binary sketch image drawn by the user to size 64 x 64 , a correlation based matching is performed, a kind of template matching. This is done by partitioning the user sketch into 8 x 8 blocks of 8 x 8 pixels and fined the maximum correlation of each block of the sketch within a search area of 16 x 16 pixels in the image database (this is done by shifting the 8 x 8 block in the search area). This local correlation score is computed on the pixel level using logical operations. The matching score of a database image is the sum of the correlation scores of all local blocks.

3.1.3 Indexing

QBIC was one of the first systems that applied multidimensional indexing to enhance the speed performance of the system. The average color and the texture features (both 3D vectors) are indexed using trees. The 18 dimensional moment-based shape feature vector is first reduced using the KL transform and then indexed by using R*-trees.

3.1.4 Result presentation

The best matches are presented in decreasing similarity order with the matching score.

3.2 Multimedia Analysis and Retrieval System (MARS) [9, 10, 11]

It was developed by the Beckman Institute for Advanced Science and Technology, University of Illinois. It supports colour, spatial layout, texture and shape matching.

3.2.1 Features

The system supports queries on combinations of low-level features (color, texture, shape) and textual descriptions. Color is represented using a 2D histogram over the HS coordinates of the HSV space. Texture is represented by two histograms, one measuring the coarseness and the other one the directionality of the image, and one scalar defining the contrast. In order to extract the color/texture layout, the image is divided into 5× 5 subimages. For each subimage a color histogram is computed. For the texture of a subimage, a vector based on wavelet coefficients is used. The object in an image is segmented out in two phases. First, a k-means clustering method in the color-texture space is applied, then the regions detected are grouped by an attraction based method. This consists of choosing a number of attractor regions and associating each region with the attractor that has the largest attraction to it. The attraction between two regions, i and j, is defined as $F_{ij} = M_i M_j / d_{ij}^2$, where M_i M_j are the sizes of the two regions and d_{ij} is the Euclidean distance between the two regions in the spatial-color-texture space. In the MARS system, five attractors are used: one for each corner of the image (background attractors) and one in the center of the image (the objects attractor). This is consistent with the fact that their database consists of images of single objects. The shape of the boundary of the extracted object is represented by means of Fourier Descriptors (FD).

3.2.2 Querying

Complex queries can be formulated using boolean operators. The desired features can be specified either by example (pointing an image database that has such a property) or direct (for example, by choosing colors from a palette or textures from an available set of patterns).

3.2.3 Matching

The similarity distance between two color histograms is computed by histogram intersection. The similarity between two textures of the whole image is determined by a weighted sum of the Euclidean distance between contrasts and the histogram intersection distances of the other two components, after a normalization of the three similarities. For computing the texture similarity between two corresponding subimages, the Euclidean distance between the vector representations is used. A weighted sum of the 5x5 color/texture similarities is used to compute the color/texture layout distance between two images. The similarity measure between two FD shape representations is a weighted sum of the standard deviations of $ratio(k) = M_2(k) / M_1(k)$ and $Shift(k) = \theta_2(k) - \theta_1(k) - \psi$, $k = -N_c, \ldots\ldots, N_c$ where $M_i(k)$and $\theta_i(k)$ are the magnitude and the phase angle of the FD coefficients ψ is the difference of the major axis orientations of the two shapes and Nc is the number of FD coefficients. Each query has a query tree associated. In a query tree, the leaves represent the feature vectors (the terms of the boolean expression defining the query) while the internal nodes correspond to boolean operators or more complex terms indicating a query by object. Individual queries on each of the query terms are made. The tree is evaluated bottom-up: each internal node receives from each child a list of ranked images and combines these lists, after a normalization process, according to the weights on the parent-child links.

3.3 Visual Information Processing for Enhanced Retrieval (VIPER)[14]

It was developed at the Computer Vision Group, University of Geneva. It supports colour and texture matching.

3.3.1 Features

A basic concept is that of a primitive, which denotes a feature’s type, computation and matching distance. Five abstract data types are defined: global values and histograms, local values and histograms, and graphs. The VIR Image Engine provides a set of general primitives, such as global color, local color, texture and shapes. Apart from these, various domain specific primitives can be created when developing an application. When defining such a primitive, the developer supplies a function for computing the primitive’s feature data from the raw image.

3.3.2 Querying and Result presentation

The VIR Image Engine provides a set of GUI tools necessary for the development of a user interface. These include facilities for image insertion, image query, weight adjustment for re-query, inclusion of keywords, and support for several popular image file formats. Another available component, the query canvas, allows queries-by-sketch; it consists of a bitmap editor where the user can sketch a picture with drawing tools and color it using the colors from a palette. Also, the user can bring onto the canvas an image from an existing collection and modify it using the same drawing tools. Queries can be performed on various user-defined combinations of primitives.

3.3.3 Matching

When defining a new primitive, a function for computing the similarity between two sets of feature data previously extracted must also be supplied by the developer. When comparing two images, for each primitive in the current query combination, a similarity score is computed using the distance function defined within the primitive. These individual scores are combined in an overall score using a set of weights in a way characteristic to the application. This score is then stored in a score structure, which contains also the individual similarity scores for each primitive. This allows a quick recompilation of the overall score for a new set of weights.

4. RESULT ANALYSIS

Research in content-based image retrieval (CBIR) today is an extremely active discipline. There are already review articles containing references to a large number of systems and description of the technology implemented. This paper proposes a content-based image retrieval system for images from the databases. For evaluation of retrieval performance, MPEG group have defined an evaluation metric called Averaged Normalized Modified Retrieval Rate (ANMRR) in order to measure the performance of retrieval. It was developed on the basis of the specification of a data set, a query set and the corresponding ground-truth data, which is a set of visually similar images for a given query image.

4.1 Color

Computing distance measures based on color similarity is achieved by computing a color histogram for each image that identifies the proportion of pixels within an image holding specific values (that humans express as colors). Current research is attempting to segment color proportion by region and by spatial relationship among several color regions [20]. Examining images based on the colors they contain is one of the most widely used techniques because it does not depend on image size or orientation. Color searches will usually involve comparing color histograms, though this is not the only technique in practice.

4.2 Texture

Texture measures look for visual patterns in images and how they are spatially defined. Textures are represented by texels which are then placed into a number of sets, depending on how many textures are detected in the image. These sets not only define the texture, but also where in the image the texture is located. Texture is a difficult concept to represent. The identification of specific textures in an image is achieved primarily by modeling texture as a two-dimensional gray level variation. The relative brightness of pairs of pixels is computed such that degree of contrast, regularity, coarseness and directionality may be estimated (Tamura, Mori & Yamawaki, 1978). However, the problem is in identifying patterns of co-pixel variation and associating them with particular classes of textures such as silky, or rough.

4.3 Shape

Shape does not refer to the shape of an image but to the shape of a particular region that is being sought out. Shapes will often be determined first applying segmentation or edge detection to an image. Other methods like [Tushabe and Wilkinson 2008] use shape filters to identify given shapes of an image. In some case accurate shape detection will require human intervention because methods like segmentation are very difficult to completely automate.

5. CONCLUSION

The features of content based image retrieval techniques are classified into the low level classes: Color, texture, and shape, and the higher level classes: Layout and face detection. In this paper, we compared various image retrieval techniques based on Color, Texture, Shape and Semantic. Indeed, for small collections of images, an indexing data structure is not needed, and a linear search can be sufficiently fast. Most computers can perform simple matching of hundreds of images in near real time. It is widely recognized that most current content-based image retrieval systems work with low level features (color, texture, shape), and that next generation systems should operate at a higher semantic level. This can be achieved in future by letting the system recognize images and photographs.

References

[1] R. Brunelli and O. Mich, "Histograms Analysis for Image Retrieval," Pattern Recognition, Vol.34, No.8, pp1625–1637,2001.

[2] C. S. Fuh, S.W. Cho and K. Essig, "Hierarchical Color Image Region Segmentation for Content-Based Image Retrieval System,"IEEE Transactions on Image Processing, Vol. 9, No. 1, pp. 156–162, Jan. 2000.

[3] M. Adoram and M. S. Lew, "IRUS: Image Retrieval Using Shape," Proceedings of IEEE International Conference on Multimedia Computing and System, Vol. 2, pp. 597–602, 1999.

[4] Yao M, Luo J H. Research on generalized computing system. Journal of System Engineering and Electronics, 1998, 9 (3): 39-43.

[5] K. Fukunaga, Introduction to Statistical Pattern Recognition, San Diego, CA, Academic Press, 1990. [1]

Flicker M, Query by image and video content. The QBIC System IEEE Computer, 1995,28(9): 23-32.

[6] Stricker M A,Orengo M, Similarity of color images , Proc of SPIE, Storage and Retrieval for Image and Video Database, San Jose, CA:s.n, 1995:381-392.

[7] Flickner M, Sawhney H, N aIblack W, et al. Query by image and video content: the QBIC system. IEEE Computer, 1995, 28 (9): 23 -32.

[8] A.Natsev,R.Rastogi and K.Shim."WARLUS: a similarity retrieval algorithm for image database," IEEE Transaction on Knowledge and Data Engineering 16(3), March 2004.

[9] F.Jeng, M.Li, H.-J. Zhang and B. Zhang, "An efficient and effective region-based image retrieval framework," IEEE Transaction on Image Processing 13(5), May 2004.

[10] B. Brandshaw. "Semantic based image retrieval: aprobabilistic approach," proc, ACM Multimedia, October 2000. http://www.cs.virginia.edu/papers/MIS03.pdf

[11] W. Y. Ma. NETRA: A Toolbox for Navigating Large Image Databases. PhD thesis, Dept. of Electrical and Computer Engineering, University of California at Santa Barbara, June 1997.

[12] Wei-Ying Ma and B. S. Manjunath. Netra: A toolbox for navigating large image databases. Multimedia Systems, 7(3): 184–198, 1999.

[13] David McG. Squire, Wolfgang M¨uller, Henning M¨uller, and Thierry Pun. Content-based query ofimage databases: inspirations from text retrieval. Pattern Recognition Letters, 21:1193–1198, 2000.

[14] Gupta. Visual information retrieval: A virage perspective.Technical Report Revision 4, Virage Inc., 9605 Scranton Road, Suite 240, San Diego, CA 92121, 1997.

[15] Michael Ortega,Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra, and Thomas S. Huang. Supporting similarity queries in MARS. In Proceedings of the 5th ACM International Multimedia Conference, Seattle, Washington, 8-14 Nov. '97, pages 403–413, 1997.

[16] R. Smith and S.-F. Chang. Querying by color regions using the VisualSEEk content-based visual query system. In M. T. Maybury, editor, Intelligent Multimedia Information Retrieval. AAAI Press, 1997.

[17] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin. The qbic project: Quering images by content using color, texture, and shape. In Poceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases, 2-3 February '93, San Jose, CA, pages 173–187, 1993.

[18] J. R. Smith and S.-F. Chang. Querying by color regions using the VisualSEEk content-based visual query system. In M. T. Maybury, editor, Intelligent Multimedia Information Retrieval. AAAI Press, 1997.

[19] Gulfishan Firdose Ahmed, Raju Barskar,Jyoti Bharti, Ntin Singh Rajput, ""Content Base Image Retrieval Using Fast Phong Shading", In proceeding of IEEE, International Conference on Computational Intelligence and Communication Networks, CPS and indexed in IEEE Computer Society, pp.419-423, Novmber-2010, Bhopal, India.

Author Profile

Amit Singla received the B.Tech degree in Computer Science Engineering from Punjab Technical University Jalandhar in 2012. Currently he is pursuing M.Tech degree in Computer Science from Guru Kashi University, Talwandi Sabo, Bathinda (Punjab). His research interests include Image processing and Data mining.