Content based Image Indexing & Retrieval

Pratap Singh Patwal[1], Dr. A.K. Srivastava[2]

Research Scholar, SGV University, Jaipur[1], Director, RGGI, Meerut[2]

Abstract- Multimedia (Images) are being generated at an enormous rate by sources such as defense and civilian satellites, biomedical imaging, military reconnaissance and surveillance flights and home entertainment systems, scientific experiments, fingerprinting and mug-shot-capturing devices. For example, National Aeronautics and Space Administration (NASA) Earth Observing System will produce about 1 TB of image data per day when completely operational. A content- based image retrieval (CBIR) system is required to efficiently and competently use information from these image data sets. Such a system that helps users who is unfamiliar with the database can retrieve relevant images based on their contents.

Keywords— Content-Based Image Retrieval (CBIR), Web Service, Image Similarity, Information Visualization. TB (TeraByte).

1. INTRODUCTION

Images can be indexed and retrieved by textual descriptions and by image content. In textual queries, words are used to retrieve images or image sets, and with visual queries (Content-based Retrieval), images or image sets are retrieved by visual characteristics such as color, texture, and pattern. Image retrieval strategies based on either text-based or CBIR alone have their restrictions. Image is frequently clear to a broad scope of interpretations, and textual descriptions can simply begin to capture the richness and complexity of the semantic content of a visual image. Human indexing of images is highly labour intensive and restrictive when large databases are required. However, retrieval based on optical characteristics is computationally intense, and has not yet arrived at the point where it can be efficiently used to formulate intellectually subtle queries, especially for non-specialist users.

2. BACKGROUND

The largest amount of multimedia information is contained in the internet often making the task of finding information difficult [1]. Therefore, there is a need of software which can simplify the task of finding information stored in the internet. WebCrawler [2] and Google [3] are popular search engine which aims to index the large portion of internet. These systems attempt to hold information about the internet in a single location to help the users quickly find the information. This allows the system to rapidly refresh the information contained within the search engines. Another important property is that the Google and the WebCrawler search engine use a database management system to simplify the process of retrieving and storing information. Also, both search engine developer recognizes that the internet has varied range of user ability from beginner to expert users.

Presently there is an ever growing interest in researching method for integrating image processing techniques into searching the large multimedia database. This technique aims to improve the quality of search results. The Webseek prototype search engine utilities textual analysis, color matching of images and spatial arrangement of images. Using this system the search engine has the ability to automate the process of analyzing textual, image and video information. Another system using image processing to improve search result quality of Chabot. This system uses keyword search and color analysis of the image. The color analysis is a complex task if finds the predominant color types in an image. Using this technique, the user has the ability to join search queries together by specifying the keywords and color predominance properties. An important common area of interest between the popular search engines and these search engines is the use of information retrieval, database management and user interface design.

2.1 Color Processing

It is well known that color is one of the most prominent queues for human visual perception and, due to this fact; it has seen a lot of research over the years, making color feature extraction relatively easy to implement. Therefore, color, features have been widely researched in the area of content-based image retrieval. The most common representation of color information is in the form of a color histogram which statistically is the probability of any given pixel having a specific intensity in each of the color channels. It was Swain and Ballard, who proposed using histogram intersections as a similarity measure [Swain and Ballard, 1995] [27]. However, given that histograms are sensitive to noise due to their sparse nature, Stricker and Orengo extended this work to use a cumulated histogram.

Stricker and Orengo introduced the use of color moments with a weighted Euclidean distance metric to determine image color similarity. This extension helped to improve the accuracy of the image matching since the histograms were far less susceptible to noise.

In order to improve retrieval efficiency, it is common practice to quantize the image color space into a fixed number of color bins, as was proposed by Smith and Chang [2]. Before quantization, Smith and Chang firstly transformed the color space from RGB to HSV and then quantized the color space into a set of M bins. A binary color set vector would then be determined for each pixel of an image.

An image which denoted which color bins were contained within the quantized image. Due to the color set vectors being binary in nature, they could construct a binary tree and performed the retrieval process with a simple binary tree traversal algorithm. The quantization method has also been extended in recent work carried out by Kulkarni, Srinivasan and S. Ray [Kulkarni et al., 1998] whereby a second step called the 'shrink phase' was introduced. During this 'shrink phase', the image was divided up into fixed sized regions and reduced in size. This was achieved by compressing all of the pixels within a given region to a single pixel with an intensity value equal to that of the mean value calculated from all pixels within the specific region. Their results showed that the inclusion of the second step improved the accuracy of retrieval as it reduced the number of false positives retrieved when compared to standard uniform quantization with a global histogram. This work has also been investigated by Faloutsos and others in [Faloutsos et al., 1993] [20].

In work done by Lu, Ooi and Tan, a variation of the shrink phase was used where the image was divided into a quad-tree structure rather than fixed sized blocks [Lu et al., 1994] [8]. Each branch of the quad-tree would then have its local histogram calculated to describe the color content of that region, however, in the research carried out by Smith and Chang, it was discovered that the regular image block sub-division suffered from too many inaccuracies and was computationally expensive. Smith and Chang suggested that the image be segmented by the use of color set back projection to identify high color information areas. These areas would then have their color set and position stored in the retrieval process and although this approach does give greater accuracy, it is well known that at present there are not many reliable segmentation techniques available and thus hard to verify the accuracy of the segmentation using back set projection.

The method proposed by Rickman in [Rickman and Stonham., 1996] was to employ a color tuple histogram approach. In their work they constructed a database containing each of the possible quantized color hue combinations which may be found in a given region. From this database a local histogram was constructed, based upon the quantized hue entries and then used for the similarity matching. Whilst this approach provided adequate results, it did suffer from sensitivity to small transformations within the image regions. To overcome this, Stricker and Dimai proposed extracting the first three color moments from a set of five predefined partially overlapping regions (or fuzzy regions)[Stricker and Orengo, 1995].

A similar technique was used by Pass et al in [Pass et al., 1996], whereby each pixel of a particular color was classified as either incoherent or coherent depending on whether or not it belonged to a large region of similarly color pixels. Using this method they were able to distinguish between clustered and sparse pixels, thus further refining the spatial color information of the regions. Along the same lines, Huang and Rui [Rui et al., 1997] [9] used a Color correlogram to determine the color, layout within the image. This was carried out by firstly constructing a color co-occurrence matrix and then using an auto-correlogram and correlogram as the distance metrics. Their results showed that the retrieval accuracy was improved when compared to standard histogram techniques.

2.2 Shape Processing

Without doubt, shape information is considered one of the most difficult features to extract reliably from digital images. This is due to the fact that there are no mathematical definitions of shape [Scassellati et al., 1994].

In implementing a content-based image retrieval system which employs shape information, it is important that the representation is invariant to transformations such as translation, rotation and scale. One such representation is a region based technique known as moment invariants. It was the early work of Hu [Hu, 1968] [24] which identified seven moment invariants which are invariant to shape transformations and were derived from the second and third moments. Other well known methods of shape represented are area and circularity as well as minimum and maximum axis all of which have been employed in the QBIC retrieval system by IBM [Equitz and Niblick, 1994] [19]. As their respective names suggest, area is the total number of pixels contained within the region. Circularity is the degree to which a particular region approximates a circle and can be expressed as the variance of a boundary pixel's distance to the centroid. Minimum and maximum axes are the axis of minimum and maximum inertia, calculated from the pixel coordinates within a specific region. This requires the central moments of the region to be calculated beforehand, which can also be used to represent shape, however, the central moments are not transforming invariant.

One representation which uses boundaries to describe shapes is the Fourier Descriptor. Earlier work using Fourier Descriptors was carried out by Persian and Fu, whereby the boundaries 2D image coordinates are represented as a 1D complex array and then transformed to obtain the Fourier Descriptors. They discovered that a total of 15 descriptors was adequate to represent each shape accurately. One of the problems with their method, however, was that it was sensitive to digitization noise and required complex mathematical distance metrics to counter the variations in shape translation, scale and rotation. To overcome this problem, Rui [Rui et al., 1996] extended this work and introduced a modified Fourier Descriptor to extract shape information which was invariant to the problems of Fu's work so that a simple Euclidean distance measure could be used to determine similarity.

Another technique which can be used to find geometric lines and shapes within binary images is the Hough Transform. The Hough Transform is essentially a mapping algorithm which processes data from Cartesian image coordinate space into a polar parameter space [Myler and Weeks, 1993]. Using the expression Xcos θ + Ysin θ = P, the points in the X-Y plane can be mapped into the polar plane θ-P. This is done by using an Accumulator matrix

Which is the discretisation of the θ -P space carried out by solving the parametric equation above. The resulting accumulator matrix can then be used by ordering the values from largest to smallest with the largest values corresponding to lines within the image. Line detection is then a matter of finding the local maxima, which generally results in a cluster of points, with the centre point being the straight line representation.

However, one of the major drawbacks with the Hough Transform and why it has not been widely used is that most of the time edge detection techniques do not produce accurate straight lines, resulting in inaccurate shape recognition. In order to overcome this, a fuzzy Hough Transform was proposed by R. Soodamani and Z.Q. Liu [Soodamani and Liu].

Their technique used a standard Hough Transform algorithm, except that when assigning values to the accumulator array they used a fuzzy measure to determine the value for P. From their results this technique shows some improvement over the conventional Hough Transform and looks promising for future expansion.

2.3 Indexing Techniques:

Due to the nature of image features being multi-dimensional it is important that the retrieval process employ an efficient and effective indexing technique. This particular problem of image retrieval has brought together three major research fields, namely Computational Geometry, Database Management and Pattern Recognition. Some previous methods which have been proposed and researched are Bucketing algorithms, k-d tree, priority k-d tree [White and Jain., 1996] [4, [10], R-tree, quad tree, hB tree, K-D-B tree, as well as its variants [Guttman, 1984, Sellis et al., 1987, Greene, 1989, Beckmann et al., 1990, Rui et al., 1997] [15,21].

Two recent additions to these, popular with pattern recognition, are clustering, neural networks and Bayesian networks, which have shown promising results [Duda and Hart, 1973, Zhang and Zhong, 1995, Indrawan, 1998].

The quad tree and k-d tree algorithms have been around for a while and were first introduced back in the 1970's, however, due to their disappointing performance they have not been extended upon. Then, due to the push for spatial indexing required by CAD systems, the R-tree index structure was developed by Guttman in 1984 [Guttman, 1984]. This work was later extended by S Ellis with the R+ tree [Sellis et al., 1987] and also by Greene[Greene, 1989]. There have been many other variants of the R tree scheme, however, it is widely considered that the R* tree is one of the best variations. White and Jain [White and Jain., 1996] performed research which looked into two multi-dimensional indexing techniques known as the VAM k-d tree and the VAMsplit R tree. From their results they concluded that the VAMsplit R tree gave the best performance. Another technique was proposed by Ng and Sedigihiam [Ng and Sidighian, 1996] which involved a three step process;

1. Dimension reduction of the feature vectors.,
2. Evaluation of the existing indexing approaches and
3. Customization of the selected indexing method.

Their results showed that the BA-KD tree gave the best performance.

One of the major problems associated with content-based image retrieval is that the image feature vectors often require non-Euclidean similarity measures and, in some cases are non-numeric. Three approaches which have recently been proposed to overcome these hurdles are clustering, neural networks and Bayesian networks, each of which has shown promising results. Charikar [Charikar et al., 1997] proposed a clustering technique for dynamic information retrieval which had three advantages, namely: dynamic structure, the capability to handle high dimensional data and the probable to deal with non-Euclidean match measures. Rui and Cakrabarti [Rui et al., 1997] later extended this work in the areas of non-Euclidean similarity measures and faster more accurate search strategies.

Self Organizing Map (SOM) neural networks was used by Zhang and Zhong [Zhang and Zhong, 1995] [21] to construct a tree indexing structure for image retrieval. SOM neural networks have the advantage in that they support arbitrary similarity measures and from their results the SOM neural network approach shows to be a promising technique. Neural networks are not the only approach which has been borrowed from the Artificial Intelligence field, as Bayesian probability networks have recently seen a lot of research activities for image retrieval.

An example of the use of Bayesian networks can be found in the work done by Vasconcelos and Lippman [Vasconcelos and Lippman, 2000] [27], which presents a Bayesian learning algorithm that uses belief propagation to incorporate user provided feedback during the retrieval session. Their experimental results have shown that the inclusion of the user feedback data and learning during the retrieval session has improved the frequency of convergence to relevant images by the system. Another recent paper by Maria Indrawan [Indrawan, 1998] [23] gave a probabilistic method using Bayesian networks, which showed that the inferential properties of Bayesian networks allowed for automated learning mechanisms and multiple queries using a single network. As well as this, Bayesian networks also allow for multi-level representations of complex objects.

3. IMAGE INDEXING

The objective of image indexing is to retrieve similar images from an image database for a given query image [1]. Each image has its exclusive feature. Hence, image indexing can be carried out by comparing their features, which are extracted from the pictures. The standard of similarity between images may be based on the features such as vividness, intensity, condition, positioning and texture, and above mentioned other image properties. Current Image indexing techniques are of two types [1, 15].

1. Textual
2. Content- based

3.1 Textual

It is real simple techniques; keeping in mind the user approach keywords are employed to a particular picture. These include the following [2,15]:

1. Standard subject headings, Classification, etc.
2. Keyword additions
3. Caption indexing

The problem with this indexing is that it is

1. More prone to inter indexer consistency problems than indexing of textual matter
2. Off-ness, thing-ness, about-ness ambiguities
3. It's a very labor intensive process

3.2 Content-based

It is too known as automated indexing, in this technique images are indexed based on their content like color, pattern, direction, texture, spatial relation, etc. This sort of indexing is taken care by the software itself, algorithms are produced which can differentiate the color, pattern, textures etc. The image retrieved through this technique is known as Content Based Image Retrieval (CBIR).

4. What is CBIR?

The most primitive use of the term CBIR (Content-based Image Retrieval) in the experiments into the automatic retrieval of an image from an image repository with color and shape feature [2,13,15]. The idiom has since been broadly applied to identify the process of retrieving best images from an image dataset on the basis of color, texture & object) that can be automatically retrieved from an image themselves. The features used for retrieval is primitive, but the extraction process must be automatic. Retrieval of images by manually specific keywords are certainly not CBIR as the condition is broadly implicit yet if the keywords explain image content.

CBIR differ from traditional information retrieval in that image database are fundamentally unstructured, since digitized images consist solely of arrays of pixel intensities. One of the key issues with image processing is the demand to extract useful data from the raw data before any kind of analysis about the image's contents is possible [14].

Image databases, therefore, different from text based databases, where the unprocessed stuff is drawn in pixel form CBIR many of its methods from image processing, and is seen by some as a subset of that state. It differs from these surveys, mainly through its stress on the retrieval of picture with similar from an image dataset of substantial size. Image processing covers a much broader area, including image enhancement, compression, transmission, and transformation. While there are grey region the distinction between conventional image analysis and CBIR is usually quite clear-cut. An example, may get this somewhat more readable. Now in these days police forces now use face detection systems. Such systems used in ways. Firstly, the image in front side of the camera may be calculated with a single individual image dataset record to affirm his or her individuality. In scenario two images are best match, a process few observation may call CBIR. Secondly, the whole image database may be searched to find the closest matching images.

This is a real example of CBIR. Research and development issues in CBIR cover a range of topics, many shared with mainstream image processing and information retrieval. CBIR includes the following [11, 15]:

i. Understanding image users' needs and information-seeking behavior.
ii. Identification of suitable ways of describing image content.
iii. Extracting such features from raw images.
iv. Providing compact storage for big image databases.
v. Matching query and stored pictures in a manner that reflects human similarity judgments.
vi. Efficiently accessing stored images by content.
vii. Providing usable human interfaces to CBIR systems.

4.1 CBIR Techniques

Most of the CBIR indexing techniques based on color, shape and texture.

4.1.1 Color

Retrieving images based on color similarity are reached by calculating a color histogram for each icon that identifies the ratio of pixels within an image holding color value. Recent research is attempting segments the colors by region and by the visual association among some color regions.

4.1.2 Texture

Texture is a hard concept to present [16]. The detection of explicit textures in an image is proficient primarily by texture as a two-dimensional grey level deviation. The relative intensity of pairs of pixels is considered such that the degree of difference, reliability, coarseness and directionality may be projected by Tamura, Mori & Yamawaki, 1978[28]. Nevertheless, the trouble is in identifying co-pixel variation and associate them with special categories of textures such as glass and metal [16].

4.1.3 Shape

Queries for shapes are mostly accomplished by taking an example image provided by the organization or by accepting the user sketch a pattern. The main mechanisms used for shape retrieval include identification of features such as communication channels, boundaries, aspect ratio, and circularity, and by identifying regions of change or stability via region growing and edge detection. A particular concern is the problem of dealing with images having overlapping or touching shapes.

4.3 Practical Applications of CBIR

A wide scope of potential applications for CBIR technology, has been named. Some of these services are listed below [11].

1. Architectural and engineering, invention
2. Crime prevention or detection
3. Cultural heritage
4. Education and training services
5. Fashion and home design
6. Geographical data and remote sensing systems
7. Home entertainment
8. Intellectual property infringement
9. Journalism and advertising
10. Medical diagnosis
11. The military
12. Web searching tool

Conclusion

Content base image indexing and Content base image retrieval is still a developing science. As an image compression, digital image processing, and image feature extraction techniques become more developed, CBIR maintains a steady pace of development in the research field. Furthermore, the development of powerful Central processing unit and faster and cheaper memories contribute heavily to CBIR development. This development promises an immense range of future applications using CBIR.

References

1. W. Grosky and R. Mehrotra, Guest Editors, Special Issue on "Image Database Management", Computer, Vol. 22, No. 12, Dec. 1989.
2. Smith, J. and Chang, S. [1995] In Proc. IEEE Int. Conf. on Image Proc., pages 528-531, 1995. URL: citeseer.ist.psu.edu/smith95single.html
3. A.D. Narasimhalu, Guest Editor, Special Issue on Content-Based Reaieva1, ACMMultimedia Systems, Vol. 3, No. 1, Feb. 1995.
4. A. Ralescu and R. Jain, Guest Editors, Special Issue on "Advances in Visual Information Management Systems, Intelligent information Systems", Vol. 3, No. 3, July 1994.
5. V. Gudivada, V. Raghavan, and K. Vanapipat, "A Unified Approach to Data Modeling and Retrieval for a Class of Image Database Applications," in Multimedia Database Systems: Issues and Research Directions, S. Jajodia and V. Subrahmanian, Eds., Springer-Verlag, New York, 1995.
6. G. Salton, Automatic Text Processing, Addison-Wesley, Reading, Mass., 1989.
7. G. Jung and V. Gudivada, "Adaptive Query Reformulation in Attribute-Based Image Retrieval," in Third Golden West Int'l Conf. Intelligent System, Kluwer Academic Publishers, Boston, Mass., 1994, pp. 763-774.
8. Jun Yue, Zhenbo Li, Lu Liu, Zetian Fu "Content-based image retrieval using color and texture fused features" ELSEVIER Mathematical and Computer Modelling, Volume 54, Issues 3–4, August 2011, Pages 1121-1127
9. Rui, et al. - 1998 "Supporting ranked boolean similarity queries in mars – Ortega"
10. Del Bimbo, A. (1999). Visual Information Retrieval, Morgan Kaufmann, San Francisco,CA,
11. A. Gupta and R. Jain "Visual Information Retrieval" ACM,40.
12. Guttman, A. "R-tree: a dynamic index structure for spatial searching". In In Proc ACM SIGMOD.
13. G. Pass, R. Zabih, and J. Miller (1996). "Comparing Images Using Color Coherence Vectors". Proc. ACM Conference on Multimedia.
14. J. Dowe (1993). "Content-Based Retrieval in Multimedia Imaging". Proc. SPIE.
15. Beckmann, N., Kriegel, H. P., Schneider, R., and Seeger, B. (1990).
"The R* tree: an efficient and robust access method for points and rectangles". In In Proc ACM SIGMOD.
16. Charikar, M., Chekur, C., Feder, T., and Motwani, R. (1997). Incremental clustering and dynamic information retrieval. In In Proc of the 29th Annual ACM.
17. Duda, R. O. And Hart, P. E. "Pattern classification and scene analysis". John Wiley and Sons, Inc.
18. Equitz, W. And Niblick, W. "Retrieving images from a database using a texture - algorithm for the cubic system". Technical report, IBM.
19. Faloutsos, C., Flickner, M., Niblick, W., Petkovic, D., Equitz, W., and Barber, R. (1993). EÆcient and effective querying by image content. Technical report, IBM.
20. Gonzalez, R. C. And Woods, R. E. (1993). Digital Image Processing. Addison-Wesley Publishing Company, Inc., 3 edition.
21. Zhang, H. and Zhong, D. "A scheme for visual feature based image retrieval ", Proc. SPIE storage and retrieval for image and video databases
22. Hu, M. K. "Visual pattern recognition by moment invariants, computer methods in image analysis". In IRE Transactions on Information Theory, volume 8.
23. Indrawan, M. "A Framework for Information Retrieval based on Bayesian Networks". PhD thesis, Computer Science and Software Engineering. Monash
24. K.Prashanti Jasmine and P.Rajesh Kumar "Integration Of HSV Color Histogram And LMEBP Joint Histogram For Multimedia Image Retrieval" Intelligent computing, networking and informatics, advances in intelligent system and computing 243, DOI 101007/978-81-322-1665-0_76 Springer India 2014.
25. Swain, M. and Ballard, D. [1990]. Indexing via color histograms, AAAI-91 DARPA Image Understanding Workshop, p. 623630.
26. Patil, P.B. "Content Based Image Retrieval with Relevance Feedback Using Riemannian Manifolds" Signal and Image Processing (ICSIP), 2014 Fifth International Conference on Date of Conference:8-10 Jan. 2014 Page(s): 26 - 29 INSPEC Accession Number:14199015 Conference Location: Jeju Island DOI:10.1109/ICSIP.2014.9 Publisher: IEEE.
27. Vasconcelos, Lippman – 2000 "A probabilistic architecture for content-based image retrieval".

28. H.Tamura, S. Mori, and T. Yamawaki. "Texture features corresponding to visual perception".IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-8, no. 6, 1978, 460.

Authors

Mr. Pratap Singh Patwal Received his M. Phil, M. Tech, M.Sc. degree in Computer Science Presently he is working for his Ph.D., as a Research Scholar and his area of interest includes image segmentation, feature extraction, and image searching in large databases, data mining.

Prof. (Dr.) Anoop Kumar Srivastava is working as a Director Radh Govind Group of Institution (R.G.G.I) and Ph.D. in Computer Science from Tata Institute of Fundamental Research, Mumbai, His areas of research interest are Multi Agent Systems, Intelligent Agents, Machine Learning, Programming Languages, Mobile Computing, Adhoc Network, Image Processing.