A New Approach for CBIR Feedback based Image Classifier

Neetesh Gupta
Asst.Prof. Dept. of IT
TIT, Bhopal (M.P.)-India

Dr. R.K.Singh
Vice Chancellor
ITM University, Gurgaon, India

P.K.Dey
Registrar
Jodhpur National University
Jodhpur, India

ABSTRACT

Recent years have seen a rapid increase in the size of digital image collections. This ever increasing amount of multimedia data creates a need for new sophisticated methods to retrieve the information one is looking for. The classical approach alone cannot keep up with the rapid growth of available data anymore. Thus content-based image retrieval attracted many researchers of various fields. There exist many systems for image retrieval meanwhile. Retrieval of Images from Image archive using Suitable features extracted from the content of Image is currently an active research area. The CBIR problem is identified because there is a need to retrieve the huge databases having images efficiently and effectively. For the purpose of content-based image retrieval (CBIR) an up-to-date comparison of state-of-the-art low-level color and texture feature extraction approach is discussed. In this paper we propose A New Approach for CBIR with interactive user feedback based image classification by Using Suitable Classifier .This Approach is applied to improve retrieval performance. Our aim is to select the most informative images with respect to the query image by ranking the retrieved images. This approach uses suitable feedback to repeatedly train the Histogram Intersection Kernel based Classifier. Proposed Approach retrieves mostly informative and correlated images.

Keywords

CBIR, Feature Extraction, Correlation Coefficient, Classifier.

1. INTRODUCTION

With advances in the computer technologies and the advent of the World-Wide Web, there has been an explosion in the amount and complexity of digital data being generated, stored, transmitted, analyzed, and accessed [1]. Much of this information is multimedia in nature, including digital images, video, audio, graphics, and text data. In order to make use of this vast amount of data, efficient and effective techniques to retrieve multimedia information based on its content need to be developed [2]. Among the various media types, images are of prime importance. Not only it is the most widely used media type besides text, but it is also one of the most widely used bases for representing and retrieving videos and other multimedia information. There are two types of image retrieval technique. The first one is the text based image retrieval where each image in the database is labeled by keywords describing the image. For searching for an image, user formulates the query using keywords which best describes his query image. Since image databases are very large, so annotating such a huge collection is a very tedious task. Also subjective description of a query image may differ from person to person due to differences in human perception. Keyword annotation is the traditional image retrieval paradigm [3], [4]. In this approach, the images are first annotated manually by keywords. They can then be retrieved by their corresponding annotations. However, there are three

main difficulties with this approach, i.e. the large amount of manual effort required in developing the annotations, the differences in interpretation of image contents, and inconsistency of the keyword assignments.

To overcome the difficulties of the annotation based approach [1], an alternative mechanism, Content-Based Image Retrieval (CBIR), has been proposed in the early 1990's. Besides using human-assigned keywords, CBIR systems use the visual content of the images, such as color, texture, and shape features, as the image index. This greatly alleviates the difficulties of the pure annotation based approach, since the feature extraction process can be made automatic and the image's own content is always consistent [1]. The CBIR is the technique to map each of the images in the database to a feature space and then retrieve based on the feature of the query image. The main challenge of CBIR is to bridge the semantic gap between low level content descriptors with high level concepts (like faces, flowers, architectures etc). Content-based image retrieval, or CBIR, systems have gained popularity because of their objective means of assessing image content. Textual annotation, in contrast, is plagued by inconsistencies of the human annotator. Perhaps more important, in many domains, text cannot accurately capture the visual attributes of an image in the user's mind [5]. In CBIR, the query is an image that the user presents as an example of what he or she is looking for. The goal of CBIR is to retrieve images that are visually similar to the query image. To this end, a CBIR system applies image processing and computer vision algorithms to extract a vector of features from each of the database images. Typical features might include pixel color histograms, gray scale histograms, texture features, and edge-content measures. If the feature vectors are composed of n elements, each vector can be interpreted as a point in n-dimensional space. CBIR systems rest on the assumption that points in this space that are proximal to the point represented by the query image's vector should correspond to the feature vectors of images that are visually similar to the query. Similarity to the query is computed using either a default or a user-defined similarity metric [6]. For example, in the popular k nearest neighbor retrieval method, the vectors with the minimal Euclidean distance to the query vector are retrieved. (Note that some type of normalization of feature scales is required prior to this step.) In the context of image retrieval, an unweighted nearest neighbor [7], [8] retriever has at least one drawback: it assumes that all features are equally relevant. Some features may be meaningless to the user or may be highly correlated with other features. Some features may be very useful for certain queries but may lose significance for other queries. The notion of "similar" in the mind of the user may fluctuate depending on the query, the history of retrievals observed, and the user. If there is a significant discrepancy between the similarity as calculated by the system and the notion of similarity in the user's mind, the results are destined to be unsatisfactory .This problem has

served as the impetus for what is known as relevance feedback [8] [6].

Relevance feedback retrieval systems prompt the user for feedback on retrieval results and then use this feedback on subsequent retrievals with the goal of increasing retrieval Performance [10]. A typical user-system session is as follows: A user presents an image query to the system where upon the system retrieves a fixed number of images using a default Similarity metric [11]. The user then rates each returned result with respect to how useful the result is for his or her retrieval task at hand. Ratings may be simply "relevant" or "not Relevant" or may have finer gradations of relevancy such as "somewhat relevant," "not Sure and "somewhat irrelevant." The relevance feedback algorithm uses this feedback Information to select another set of images to retrieve for the user; whether the new and previous sets are disjoint depends on the particular system. The system's goal is to effectively infer which images in the database are of interest to the user based on this feedback. The user could then rate these images in the second set in a similar way and the process may iterate indefinitely in this closed-loop fashion [8,] [1].

A relevance feedback retrieval system has a number of design requirements that allow the system to function in an efficient online manner.

- After each iteration, when a set of images is retrieved, the system must require a reasonable amount of feedback. If the user needs to labor over providing feedback for numerous images after each iteration, they will tire quickly and not be satisfied with the process.

- The system must produce acceptable results after only a few iterations. If large numbers of iterations are required, the user will also tire.

- Feature extraction should be completed in a short period of time to prevent user frustration

Reasonable query feature extraction time, feedback per iteration, and number of iterations are a necessity. Exact bounds on what constitutes "reasonable" are domain and user specific, but rough bounds are determined by common sense In this dissertation , we present a new relevance feedback retrieval system that uses machine Learning to infer which images in the database would be of most interest to the user at a Particular point in time[10].

2. RELATED WORK

2.1 Color Feature

Color is the most popularly used features in image retrieval and indexing. On the other hand, due to its inherent nature of inaccuracy in description of the same semantic content by different color quantization and /or by the uncertainty of human perception, it is important to capture this inaccuracy when defining the features. We apply fuzzy logic to the traditional color histogram to help capture this uncertainty in color indexing [2], [5].In image retrieval systems color histogram is the most commonly used feature. The main reason is that it is independent of image size and orientation. Also it is one of the most straight-forward features utilized by humans for visual recognition and discrimination. Statistically, it denotes the joint probability of the intensities of the three color channels. Once the image is segmented, from each region the color histogram is extracted. The major statistical data that are extracted are histogram mean, standard deviation, and median for each color channel i.e. Red, Green, and Blue. So totally $3 \times 3 = 9$ features per segment are

obtained. All the segments need not be considered, but only segments that are dominant may be considered. because this would speed up the Calculation and may not significantly affect the end result.

2.2 Texture Feature

Texture is another important property of images. Various texture representations have been investigated in pattern recognition and computer vision. Basically, texture representation methods can be classified into two categories: structural and statistical. Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular [5]. There is no precise definition for texture. However, one can define texture as the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity. Texture determination is ideally suited for medical image retrievals]. In this work, computation of gray level co occurrence matrix is done and from which a number of statistical measures are derived. The autocorrelation function of an image is used to quantify the regularity and the coarseness of a texture. This function is defined for an image *I* as: The notion of texture generally refers to properties of homogeneity direction information.

2.3 Shape Feature

Shape is used as another feature in image retrieval. However, it is evident that Retrieval by shape is useful only in very restricted environments, which provide a good basis for segmentation (e.g. art items in front of a homogeneous background). Shape descriptors are diverse, e.g. turning angle functions, deformable templates, algebraic moments, and Fourier coefficients [6].

2.4 Combinations of Color, Texture, Shape

Features Similarity is based on visual characteristics such as dominant colors, shapes and textures. Many systems provide the possibility to Combine or select between one or more models. In a combination of color, texture and contour features is used. Extends the color histogram with textural information by weighting each Pixel's contribution with its Laplacian. also provides several different techniques for image retrieval. These methods use the feedback from the user in order to automatically select or weight the different models, such that the requirements of the user are best fulfilled. Several systems perform segmentation and characterize each region by color, texture, and shape. As stated before these systems crucially depend on a good segmentation which is not always given. This can result in retrieval results which seem to have nothing in common with the query. Among various flowers the system returns an image that does not display a star-shaped object but for which the segmentation delivered a star-like region as an artificial boundary within a textured region. We therefore consider segmentation-based methods reliable only for applications that allow for a precise segmentation [5], [12].

2.5 Similarity Computation

Similarity measurement is a key to CBIR algorithms. These algorithms search image database to find images similar to a given query, so, they should be able to evaluate the amount of similarities between images,. Therefore, feature vectors, extracted from the database image and from the query, are often passed through the distance function d. The aim of any distance function (or similarity measure) is to calculate how

close the feature vectors are to each other. There exist several common techniques for measuring the distance (dissimilarity) between two N-dimensional feature vector f and g. Each metric has some important characteristics related to an application. The following are the most important metrics used in the literature [10], [11].

- Euclidean metric
  The Euclidean metric is obtained as:

$$D_{eucld}(f,g) = \sqrt{\sum_{i=1}^{N}(f_i - g_i)^2} \dots \dots (1)$$

- Coefficient of Correlation

$$D_{corr}(f,g) = \frac{\sum (f_i - f)(g_i - g)}{\sqrt{\sum (f_i - f)^2 \sum (g_i - g)^2}} \dots \dots (2)$$

2.6 Image Classification

It is processing techniques which apply quantitative methods to the values in a digital yield or remotely sensed scene to group pixels with similar digital number values into feature classes or categories [1],[6] Image classification and annotation are important problems in computer vision, but rarely considered together. Intuitively, annotations provide evidence for the class label, and the class label provides evidence for annotations. Image classification is conducted in three modes: supervised, unsupervised, and hybrid. General, a supervised classification requires the manual identification of known surface features within the imagery and then using a statistical package to determine the spectral signature of the identified feature. The spectral fingerprints of the identified features are then used to classify the rest of the image. An unsupervised classification scheme uses spatial statistics (e.g. the ISODATA algorithm) to classify the image into a predetermined number of categories (classes). These classes are statistically significant within the imagery, but may not represent actual surface features of interest. Hybrid classification uses both techniques to make the process more efficient and accurate. We can use SVM (Support Vector Machine) for classification. Support Vector Machine, an important machine learning technique has been used efficiently for variety of classification purposes like object based image analysis, hand written digit recognition, and image segmentation among others. SVM can be used efficiently as a binary classifier as well as multi class classifier purpose. A Support Vector Machine (SVM) performs classification by constructing an *N*-dimensional hyper plane that optimally separates the data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network. Support Vector Machine (SVM) models are a close cousin to classical multilayer perceptron neural networks. Using a kernel function, SVM's are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training. Linear SVM is the simplest type of SVM classifier which separates the points into two classes using a linear hyper plane that will maximize the margin between the positive and negative set. The concept of non linearity comes when the data points cannot be classified into two different classes using a simple hyper plane, rather a nonlinear curve is required. In this case, the data points are mapped non-linearly to a higher dimensional space so that they become linearly separable.

3. PROPOSED APPROACH

The proposed method shows some relevant images to the user based on the query image the user selects. We have proposed a new approach for Content based image retrieval system feedback based image classifier in efficient manner. In Proposed system we have prepared a database consisting a number of images. The user has to input test image (query image) and selected object images. Low level Features like color and texture are then extracted for test (query) image and object images. Similarities distances are measured and calculating linear Coefficient of Correlation between test (query) and object images. On the basis of coefficient of correlation Training set is formed into two categories (relevant and irrelevant). Then using those labeling the histogram intersection kernel based classifier is trained. These feedback rounds go on iteratively and each time a refinement of the images shown to the user is done. Finally those images which are judged by the classifier as positive are taken. The query image is plotted in the positive set. Now those images which are nearest to the query image based on a histogram intersection based similarity ranking are taken and shown to the user. According to this proposed Approach is following in next steps.

Step 1: Preprocessing

- Input various object Images Oi.

- Create Bm1*1 Block Matrices.

- Calculate Mean μ of Block Matrices and Concatenate.

Step 2: Feature Extraction

- Convert Bm1*1 Block Matrices "f & g" RGB from space to HSV from space, where f and g represent the average values of vectors.

- Extract feature vector V_j from HSV space. and combined all color and texture features.

Step 3: Similarity Computation

Calculate Euclidean Distance then get Euclidean (D)

$$D_{eucld}(f,g) = \sqrt{\sum_{i=1}^{N}(f_i - g_i)^2} \dots \dots (1)$$

(where f and g represent the average values of feature vectors respectively.)

While ($T_i == O_i$) // where T_i is the test image & O_i is the object images.

Repeat above procedure for n object images. Now we have 'N' object image and its Euclidean distance Matrices.

Calculate correlation coefficient

$$D_{corr}(f,g) = \frac{\sum (f_i - f)(g_i - g)}{\sqrt{\sum (f_i - f)^2 \sum (g_i - g)^2}} \dots \dots (2)$$

Step 1: Relevance Feedback

- Mark all Images as Relevant Images as well as Irrelevant Images and Forms set Rr and Ri. (Where Rr ε Tr and Ri ε Tr) and Tr is Training Set.

Step 5:

- The set Rr is modified by including the query image selected by user in it.

- Calculate feature vector of Rr and Ri. Feature vectors of Fr forms the positive set and feature vectors of Fi forms the negative set of data points for training the classifier.

- These sets are then given as input to classifier (IKSVM). It draws the hyper plane on the basis of them.

- The feature set of Tr is calculated and then fed to classifier for classification so that the hyperplane formed can separate data points in training set as positive or negative. Those data points which falls on the positive side of the hyper plane forms a set F_{Tr} and similarly the negative side forms another set F_{Ti}. Since an image is a data point in a feature space we use the term data point and feature vector interchangeably.

- This Process is iterated many times.

Step 6:

- Collect the set F_{Tr} from the last iteration of training.

- Plot all the data points in F_{Tr} in the same space.

- Now, say user wants g numbers of images from the database which are most relevant to the query image, then on the basis of histogram intersection based similarity measure calculate the distance of query data point with data points in the set F_{Tr}.

- Sort in descending order the distances as two of the most relevant images will have histogram distance close to 1 and the dissimilar ones will have distance close to 0 and then retrieve first P no. images from that sorted list.

4. CONCLUSION AND FUTURE WORK

Over segmenting the images using an automatic algorithm and letting the user group the smaller regions into objects, can be adopted as an alternative partitioning method. In this way, we may obtain a more accurate partitioning compared to our manual coarse segmentation over a rectangular grid. The current feature set can be extended to cover a broader class of descriptors, including higher level of texture and shape descriptors. Lastly the attribute mapping of our sliding areas can be enhanced by allowing the user formulate queries in the form "look for this on this kind of background" to benefit from the background's distinctiveness. Interactive Content Based Image Retrieval System which aims at selecting the most informative images with respect to the query image by ranking the retrieved images. System which is more robust and computationally efficient. Complexity will be reduced in terms of resources.

5. REFERENCES

[1] S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak "A Universal Model for Content- Based Image Retrieval "Proceedings of World Academy of Science, Engineering and Technology, 36 December 2008, ISSN 2070-3740.

[2] Zurina Muda, Paul H. Lewis, Terry R. Payne, Mark J. Weal. "Enhanced Image Annotations Based on Spatial Information Extraction and Ontologies". In Proceedings of the IEEE International Conference on Signal and Image Processing Application, ICSIPA 2009.

[3] Tienwei Tsai, Te-Wei Chiang, and Yo-Ping Huan, "Image Retrieval Approach Using Distance, Threshold Pruning", IEEE Trans. On Image Processing 2007, Vol.12, PP.241-249.

[4] P Liu, K Jia, Z Wang and Z Lv "A New and Effective Image Retrieval Method Based On Combined Features" Proc. IEEE Int. Conf. On Image and Graphics, August 2007, Vol I pp 786-790.

[5] K.P. Ajitha Gladis, K.Ramar, "A Novel Method for Content Based Image Retrieval Using the Approximation of Statistical Features, Morphological Features and BPN Network", IEEE computer society ICCIMA 2007 ,Vol. 148, PP. 179-184.

[6] J. Bi, Y. Chen and J.Z. Wang. "A sparse support vector machine approach to region-based image categorization". In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR '04), 2004.

[7] Jia Li, James Z. Wang, Gio Wiederhold "IRM: Integrated Region Matching for Image Retrieval" in National Science foundation Digital Library, 2000.

[8] K.-S. Goh, E.Y. Chang, and W.-C. Lai. "Multimodal concept-dependent active learning for image retrieval". In Proceedings of 12th Annual ACM International Conference of Multimedia, 2004.

[9] Hachimura and A. Tojima. Image retrieval based on compositional featureand interactive query specification. In IAPR International Conference on PatterRecognition (ICPR), volume 4, pages 262–266, Barcelona, Spain, September2000.

[10] ChengjunLiu,"ABayesian Discriminating Features Method for Face Detection", IEEE Trans PAMI, Vol. 25, No. 6, June 2003.

[11] D. A. Adjeroh and M. C. Lee, "On ratio-based color indexing," IEEE Trans.Image Processing, , vol. 10, no. 1, pp. 36– 48, 2001.

[12] P. Hong, Q. Tian and T.S. Huang. "Incorporate support vector machines to content-based image retrieval with relevance feedback". In Proceedings of IEEE International Conference on Image Processing, Vancouver, Canada, 2000

[13] Hui Hui Wang, Dzulkifli Mohamad, N.A.Ismail. "Towards Semantic Based Image Retrieval: A Review". In The 2nd International Conference on Digital Image Processing (ICDIP 2010), Singapore, 2010.Proceedings of SPIE, Vol 5511, 2010.

[14] Michele Saad "Low-Level Color and Texture Feature Extraction for Content-Based Image Retrieval" Final Project Report, EE 381K: Multi-Dimensional Digital Signal Processing, May 09, 2008.

[15] S. Newsam and C. Kamath, "Comparing shape and texture features for pattern recognition in simulation data," in SPIE Electronic Imaging, San Jose,CA, January 2005, pp. 106–117.

[16] P. S. Hiremath , Jagadeesh Pujari, "Content Based Image Retrieval using Color, Texture and Shape features", 15th International Conference on Advanced Computing an Communications, IEEE Computer Society 2007, pp. 780-784.

[17] R.Datta, D.Joshi, J.Li, J.Z.Wang. "Image Retrieval: Ideas, Inuences, and Trends of the New Age". ACM Transactions on Computing Surveys, 40(2), 2008.

[18] Ville Viitaniemi and Jorma Laaksonen. "Techniques for still image scene lassification and object detection" In Proc. of ICANN 2006, volume 2, pages 35–44, Athens, Greece, September 2006. Springer.

AUTHOR BIOGRAPHY

Neetesh Gupta completed B. E. (Computer Science & Engg.) from Oriental Institute of Science and Technology / Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal(M.P.)-India in year 2003 and M. Tech. (Computer Science & Engg.) from Bansal Institute of Science and Technology/ Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (M.P.)-India in year 2010. Now he is Pursuing Ph.D. (Computer Science & Engineering Branch).He worked as a lecturer in the department of C.S.E. / I.T. at Shri Satya Sai Institute of Science and Technology (SSSIST), Sehore (M.P.) India between July 2003 to Dec 2004. Now he is working as an Asst. Professor in the Department of Information Technology at Technocrat Institute of Technology (TIT)-Bhopal (M.P.) India since 06 Jan 2005.

His research area interests include Data Mining and Warehousing, Natural Language Processing,Wireless Network, Image Processing, He is having 10 papers in International Journals (IEEE Xplore Digital Library, IEEE CS Digital Library, ACM Digital Library ,IJCTE, CiiT International Journal of Digital Image Processing, CiiT International Journal of Biometrics and Bioinformatics, CiiT International Journal of Wireless Network, International Journal of Computer Applications), 07 papers in International Conferences, 03 papers in National Conferences, 07 International / National Seminar/Workshops. He is reviewer for 08 International Journal (International Journal of Computer Applications (IJCA),International Journal of Engineering and Technology (IJET),International Journal of Computer Science and Information Security (IJCSIS), International Journal of Computer Science Issues (IJCSI),International Journal of Advanced Computational Technologies (IJACT),Journal of Convergence Information Technology (JCIT),Journal of Communications and Sciences (JCIS), Journal of Electrical and Engineering and Research (JEEER),.He is member of International Association of Computer Science &Information Technology (IACSIT) & International Association of Engineers (IAENG) .He is very hardworking and believes in GOD.