

Content based Image Retrieval and Classification using Support Vector Machine

Spurti Shinde
Department Of Computer
Engineering M.E.S.C.O.E
Pune-01.

Ashwini Lendal
Department Of Computer
Engineering M.E.S.C.O.E
Pune-01.

Nikita Bajaj
Department Of Computer
Engineering M.E.S.C.O.E
Pune-01.

Yogita Shelar
Department Of Computer
Engineering M.E.S.C.O.E
Pune-01.

ABSTRACT

Content Based Image Retrieval (CBIR) is a traditional and developing trend in Digital Image Processing. Therefore the use of CBIR to search and retrieve the query image from wide range of database is increasing. In this paper we are going to explore an efficient image retrieval technique which uses local color, shape and texture features. So, efficient image retrieval algorithms based on RGB histograms, Geometric moment and Co-occurrence Model is proposed for color, shape and texture respectively. Results based on this approach are found encouraging in terms of color, shape and texture image classification accuracy. After the features are selected, an SVM classifier is trained to distinguish between relevant and irrelevant images accordingly.

Keywords

Content Based image retrieval, Support Vector Machine, RGB Color model, Co-occurrence Model, Geometric Moment.

1. INTRODUCTION

Interest in the potential of digital images has increased largely over the last few years. Many opportunities are exploited by professional users offered by the ability to access and manipulate remotely-stored images in all possible new and exciting ways. Therefore, discovering the process of locating a desired image in a large and varied collection can be a source of considerable frustration. An image retrieval system is a very efficient system which allows us to browse, search and retrieve the images. There is a gap between the semantic understandings of users to query image databases and low-level descriptions of image content in the content-based image retrieval. Therefore there is a strong demand for developing an efficient technique from this huge amount of digital information for image retrieval. If user wants to search for many roses images then he can submit an existing rose picture as query so the system will extract image features for this query. Further it will compare these features with that of other images in a database. Relevant results will be displayed to the user.

2. CONTENT BASED IMAGE RETRIEVAL

Content Based Image Retrieval (CBIR) is the process of retrieving the required query image based on the contents of the image from a huge number of databases. In early days the manual entering of keyword approach was more difficult due to large image collections. Content Based Image Retrieval was introduced in order to overcome these difficulties. Content-based image retrieval is the application of computer vision to the image retrieval problem. In this approach images would be indexed using their own visual contents instead of being manually annotated by textual keywords. Color, texture and shape were the visual contents of images. This approach is known to be a general framework of image retrieval. There are three fundamental bases for Content Based Image Retrieval which are visual feature extraction, retrieval system design and multidimensional indexing. Some of the major areas in which CBIR is applicable are Art collections, Medical diagnosis, Crime prevention, Military, Architectural design and Geographical information and Remote sensing systems[3]. The wide application of CBIR has led to increase its efficiency. The areas where the CBIR technique finds its prime importance are Biomedicine, Military, Education, Web image classification and searching. Some of the examples for the current CBIR are Viper which is Visual Information Processing for Enhanced Retrieval, QBIC which is Query by Image Content and Visual seek which is a web tool for searching images and videos. The main purpose of CBIR is, it mainly decreases the heavy workload and overcomes the problem of heavy subjectivity.

3. RELATED WORK

3.1 Image Feature

The initial step of CBIR system is to represent color, texture component and shape regions into feature vector. There are various techniques for extraction and representation of image features like local or global feature, color, gradients, edges, contours, textures, shapes, boundaries & regions, have been reported in the literature[1]. But from the images in the database color, texture, shape features are some of the general techniques used for retrieving a particular image. People restrict themselves to one feature of image because of the classification problem has not been so easy. Content Based Image Retrieval systems work with all the images and the search

is based on comparison of various features with the query image. The geometric shape, colors and the texture of the image are the main components of CBIR. Averaging and histograms are techniques which are the color aspect. The texture aspect can be achieved by using co-occurrence model. Global and local features are used to achieve the shape aspect.

3.2 Classification Methods

Image classification is a method to label an image with appropriate identifiers. These identifiers are determined by the area of interest, to see whether it is general classification for a specific domain or arbitrary pictures. Consider, for example medical x-ray images or geographical images of terrain and many more are in existence. There are two main methods to classify an image; they are supervised and unsupervised image classification. Supervised classification uses training sets of images to create descriptors for each class. Supervised image classification is a subset of supervised learning. Models of two types can be generated by supervised learning. Mostly, supervised learning generates a global model that input objects to desired outputs. In some of the cases, the map is implemented as a set of local models. The other method of classification is unsupervised image classification. Unsupervised image classification does not depend on a training set, as they use clustering techniques which measure the distance between images and group the images with common image features together[6]. Class-identifiers are then used to label these groups. There have been many ways to classify the image; one of the more frequently used ones is a study where the researches use image features such as edge direction histograms, color histograms, edge direction coherence vector and Bayesian statistical methods in order to classify indoor from outdoor, and cities from landscapes. Another very popular method is the use of SVM (Support Vector Machines) SVMs delivers state-of-the-art performance in real-world applications such as to categorize the text, hand-written character recognition; image classification etc. SVM is now established as one of the standard tools for machine learning and data mining.

3.3 Color Feature Extraction Methods

The Basic technique which is used is based on the technique of color histogram. Color Histogram of each image is calculated and then stored in the database which represents the proportion of pixel of each color within the image. Then matching algorithm will extract those images from the databases whose color histogram matches with the required one. There are various types of histograms such as normal, weighted, dominant, and fuzzy as well as there are various color spaces: HSV, grayscale, HSL, Lab, and HMMD[7].

3.4 Texture Feature Extraction Methods

Texture measures look for visual patterns in images and how they are spatially defined. Textures are represented by pixels which are then placed into a number of sets, depending on how many textures are detected in the image. 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 in it, coarseness and particular directionality may be estimated. Due to this purpose, the problem is in identifying patterns of co-pixel variation and associating

them with particular classes of textures such as silky, or rough. An approach towards the texture analysis is usually divided into statistical, structural, transformed and model based. Textural features can also be used to estimate orientation and depth of object surface. In low level feature, texture co-occurrence matrix is used for retrieval of the images.

3.5 Shape Feature Extraction Methods

Shape is an important visual feature and it is one of the basic features used to describe image content. Searching for an image using shape features are in much attention. Good retrieval accuracy requires a shape descriptor be able to effectively find perceptually similar shapes from a database. Perceptually similar shapes usually mean rotated, translated, scaled shapes and a finely transformed shape. The descriptor should also be able to find noise affected shapes, variously distorted shapes and defective. Using nonlinear combinations of the lower order moments, a set of moment invariants which has the desirable properties of not to be variant under translation, scaling and rotation.

4. PROBLEM DEFINITION

The need to find a desired image from a collection which is shared by many expert groups including journalists, architectural and design engineers and art historians has been increased. The requirements of image users vary; it can be useful to characterize queries into three levels of abstraction: primitive features such as color, texture or shape, other logical features such as the identity of objects shown and the abstract attributes such as the significance of the scenes depicted. CBIR systems operate efficiently and effectively only at the lowest of these levels.

5. IMPLEMENTATION

5.1 Color Feature Extraction Using Average RGB

We use average RGB to calculate color similarity in the image. Average RGB is to calculate the average value in (Red) R, (Green) G, and (Blue) B channel of each pixel in an image, and use further as a descriptor of an image for the purpose of comparison. Averaging color values is always identical to averaging numbers, except with the added initial step of finding the red, green and blue components of the color. To do this we can use **bitwise operators**. RGB Color Space RGB is additive in nature. It is sum of three primary colors red, green and blue. The range of values of each of this components lies within 0 to 255. For obtaining Red pixel bitwise right shift by 16 should be done, similarly for Green right shift by 8 and for Blue bitwise AND is to be done.

Any other color space can be obtained from a linear or non-linear transformation from RGB. It is one of most widely used color space for feature extraction for image retrieval using color spaces and However, high correlation between the red, green and blue colors. The matching process then retrieves those images whose color histograms match those of the query most closely.

R: Number = pixel >> 16 & 0xFF;

G: Number = pixel >> 8 & 0xFF;

B: Number = pixel & 0xFF; (1)

Where R- Red Value, G-Green Value and B-Blue Value Pixel is the point(x, y) in an image whose RGB is extracted

5.2 Texture Feature Extraction Using Color Co-Occurrence

Co-occurrence matrix describes spatial relationships between grey-levels in a texture image. A co-occurrence matrix shows how frequent is every particular pair of grey levels in the pixel pairs, separated by a certain distance say d along a certain direction a as shown in fig. 1. While calculating the co-occurrence matrix, firstly the image is converted into gray scale, and then co-occurrence matrix is calculated called as Gray Level Co-occurrence matrix (GLCM)

Let $g=(g_{x,y}; x=1, \dots, M; y=1, \dots, N)$ be a digital image. Let $Q=\{0, \dots, q_{\max}\}$ be the set of gray levels. The co-occurrence matrix for a given inter-pixel distance d and a directional angle a is defined as

$$COOC_{m=d \cos a, n=d \sin a}(g) = [COOC_{m,n}(q,s|g); q,s=0, \dots, q_{\max}] \quad (2)$$

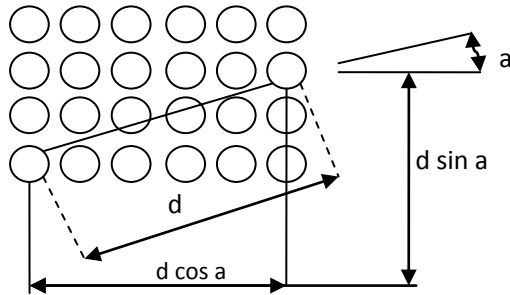


Fig 1: Grey levels in the pixel pairs

Where $COOC_{m,n}(q,s|g)$ is the cardinality of the set $C_{m,n}$ of pixel pairs $[(x,y), (x+m, y+n)]$ such that $g_{x,y}=q$ and $g_{x+m, y+n}=s$.

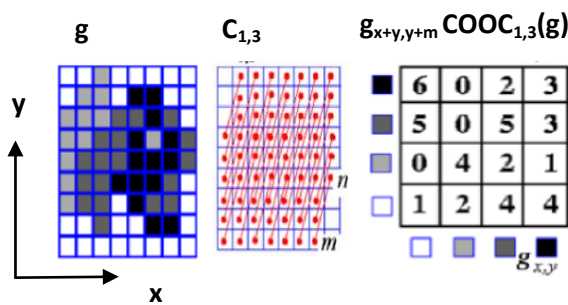


Fig 2: Gray Level Co-occurrence matrix

5.3 Shape Feature Extraction Using Geometric Moment

In content-based image retrieval shape descriptions are an important task. It is a mapping that converts the shape space into a vector space and satisfies the requirement that two similar shapes will also have close-to-identical shape descriptors. Fourier descriptors (GD) prove to be more advantageous than other techniques in terms of computation complexity, robustness, easy normalization

and retrieval performance[4]. The two-dimensional moment (for short 2-D moment) of a 2-D object R is defined as:

$$m_{pq} = \iint_R x^p y^q f(x,y) dx dy \quad (3)$$

Where $f(x,y)$ is the characteristic function describing the intensity of R , and $p+q$ is the order of the moment. In the discrete case, the double integral is often replaced by a double sum giving as a result:

$$m_{pq} = \sum \sum_R x^p y^q f(x,y) \quad (4)$$

5.4 Support Vector Machine(SVM)

Classification performed by SVM is by generating a hyper plane. The optimal hyper plane separates the required data in two categories. SVM are close cousin multilayer perception neural network. Generally support vector machine can be defined as the training method if polynomial radial basis function where the weights are calculated by solving QP problem[2]. According to the support vector machine predictor value is called an attribute and transformed attributes which is used to define the hyper plane is known as feature. The method of selecting the appropriate representation is called feature selection. Now the vector can be defined by set of features that describes predicted values. The aim of support vector machine is to find the optimal hyper plane that is used to distinguish group of vectors in a way that one category of the required variables is on one side of hyper plane and the other class of variables are on the other side of plane[5]. Vectors nearer to the hyper plane are called as support vectors. The optimum hyper plane can be defined as the linear classifier with the maximum margin for a given set of variables.

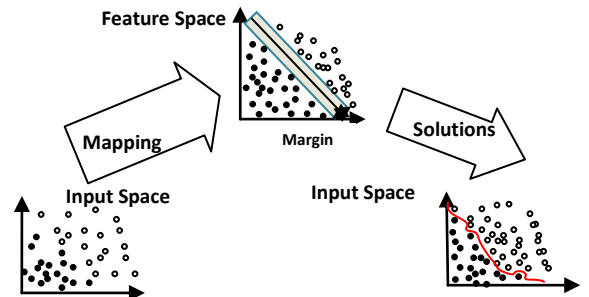


Fig 3: SVM Algorithm

5.5 Euclidean Algorithm

Distances and metrics are the most common problem in information retrieval, machine learning and character recognition. The performance of algorithms for data classification and clustering often depends largely on the availability of a good metric. Many real world problem use distances and metrics such as face recognition, visual object recognition, and automated speech recognition etc. In CBIR, the space of features is a vector space. Euclidean or the distance measure are used basically for comparing the similarities between descriptors. A Euclidean distance matrix is one in which the (x, y) entry specifies the squared distance between particle x and particle y . Euclidean distance measures the similarity between the two different feature vectors. Euclidean distance can be shown by the below formula. Q and D are feature vectors of the Query image and database image.

$$\text{Euclidean Distance} = \sqrt{\sum_{i=1}^n [Q_i - D_i]^2} \quad (5)$$

6. EXPERIMENTAL RESULTS

After applying all necessary algorithms like RGB, geometric moment and co-occurrence matrix for number of images in the database, the features are extracted. Then by classifying the images with the help of Support Vector Machine the results are calculated by using the Euclidean distance. According to their Euclidean distances the result is shown as below.

On the basis of Euclidean distance it is understood that how much relevant the retrieved image is to the query image. So with the output image Euclidean distance is tagged with the help of which user get to know that how much nearer the output image is to Query image. If any irrelevant query image is given by user, of which category is not trained by the system then large Euclidean distance is tagged with output which is up to the threshold is displayed to user.



Fig 4: Query Image



Fig 5: Retrieved Images

The system is executed with different number of images from different categories. The time required for training the database is increased as the no of images in database increases. The graph of time required for training image versus number of images is as shown in Fig. 6. But if we train large no of images in the dataset then accuracy of system increases. And the accuracy increased because of using all primary features together instead of using single feature for each image is shown in Fig. 7.

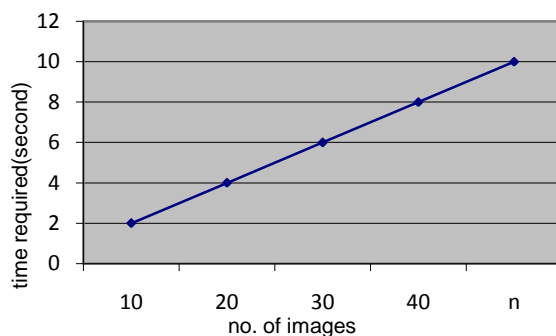


Fig 6: Time required for training images

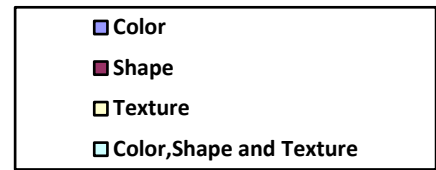
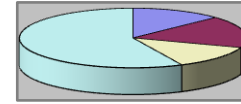


Fig 7: Distribution of accuracy using image features

7. CONCLUSION

The algorithm used for the retrieval is presented in this paper basically to reduce the computational time as well as increases the user interaction. CBIR often performs better than keyword indexing, because many of the images cannot adequately be described by linguistic cues. CBIR is desirable because most web based image search engines rely purely on metadata and this produces a lot of garbage in the results. Also having humans manually enter keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes image. Thus a system that can filter images based on their content would provide better indexing and return more accurate results. The retrieval accuracy is increased greatly as the images are retrieved on the basis of and color feature, texture feature and the shape feature. Since this idea is implemented in high level language like Java, it can be used readily in many real time applications. One of the applications is crime prevention, in which police uses visual information from people to identify the criminals. Future scope for this is the efficiency can be further increased by taking into account several features at a time as well as the quality of image can be improved and compute the similarity to obtain more accuracy.

8. ACKNOWLEDGMENTS

We feel great pleasure in publishing this journal. We wish to express true sense of gratitude toward our guide and all the people, who have helped us directly and indirectly.

9. REFERENCES

- [1] Darshak G. Thakore, A. I. Trivedi, "Content based image retrieval techniques – Issues, analysis and the state of the art".
- [2] J. Weston and C. Watkins, "Multi-class support vector machines," Technical Report CSD-TR-98-04, Royal Holloway, University of London, 1998.
- [3] Ms. K. Arthi 1, Mr. J. Vijayaraghavan, "Content Based Image Retrieval Algorithm Using Color Models," International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 3, March 2013
- [4] Aman Chadha, Sushmit Mallik, Ravdeep Johar, "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval," International Journal of Computer Applications (0975 – 8887) Volume 52– No.20, August 2012

- [5] Saurabh Agrawal, Nishchal K Verma, Prateek Tamrakar, Pradip Sircar, “Content Based Color Image Classification using SVM” 2011 Eighth International Conference on Information Technology: New Generations.
- [6] Philippe H. Gosselin Matthieu Cord, “A Comparison of Active Classification Methods for ContentBased Image Retrieval”.
- [7] Ajay Kumar Bansal, Swati Mathur, “Feature Extraction in Content Based Image Retrieval: A Review,” International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 International Conference on Advancement in Information Technology (ICAIT- 23 February 2013)