Content Based Image Retrieval Using Region Based Shape Descriptor and SVM Algorithm

Surendra Padavala[1], V. Venkata Rao[2]
[1]P.G Student, Dept. of ECE Narasaraopeta Engineering College Narasaraopeta, Andhra Pradesh, India
[2]Professor and HoD, Department. of ECE Narasaraopeta Engineering College Narasaraopeta, Andhra Pradesh, India

Abstract

Content Based Image Retrieval (CBIR) system using Region based shape descriptors is proposed in my work. Further, the image classification efficiency is improved by employing Support Vector Machine (SVM) classifier. In this paper we concentrate on region based shape descriptors. In Region based shape descriptors include Hu moments, Zernike Moments, and exact Legendre Moment. In CBIR system the region based shape descriptors, viz., MI, ZM and ELM in terms of retrieval efficiency and retrieval time are observed. In Exact Legendre Moments (ELM) for gray scale images is proposed in this work. The CBIR system is tested by conducting experiments on Corel shape database,. It consists of 20 classes of images with each class consisting of 72 different orientations resulting in a total of 1440 images. All these gray scale images in the database are of the size 128×128. All images of all the 20 classes are used for experimentation.

Index Terms— Content Based Image Retrieval, Region based shape descriptors, Hu's moments, Zernike Moments, Exact Legendre Moments Support Vector Machines.

I. INTRODUCTION

Shape is one of the primary visual features in CBIR, various type of shape descriptors have been used to extract image patterns in a number of applications. Numerous shape descriptors have been proposed in the literature. Broadly classifies shape descriptors for 2D shapes in two ways:

i. Contour Based where object shape is represented by its boundary and features (e.g. boundary length, curvature, and fourier shape descriptors). The shape description schemes are called *external representations.*
ii. Region Based object shape is described by the region occupied by the object. These description schemes are called *internal representations*.

Contour based shape descriptors make use of only the boundary information, ignoring the shape interior content. Therefore, these descriptors cannot represent shapes for which the complete boundary information is not available. On the other hand, region-based descriptors exploit both boundary and internal pixels, and therefore are applicable to generic shapes. Among the region-based descriptors, moments have been very popular since they were first introduced in the 60's. In this paper we concentrate on region based shape descriptors. Region based shape descriptors include Hu moments, Geometric moments, Legendre Moment, and Zernike Moments. These features have been studied in detail and used extensively in many applications.

The purpose of this paper to experimentally investigate into the combinations of shape descriptors and SVM algorithm that capture the concept of shape similarity and dissimilarity.

II. REGION BASED SHAPE DESCRIPTORS

Shape is one of the most widely used image feature exploited in content-based image retrieval systems, In this section, we describe important region based shape descriptors: Hu's Moments, Zernike Moments, Exact Legendre Moments

Hu's Moments:

Hu moments were proposed in and have the property of being scale, translation and rotation invariant. To compute the Hu moments, first the *central moments* are computed from the geometric moments as:

$$\eta_{ij} = \frac{M_{ij}}{M_{00}^{\gamma}}$$

Where $\gamma = (i+j)/2+1 \; for \, (i+j) \geq 2$

Based on the second and third order central moments, Hu defined the following six absolute orthogonal invariant moments:

$$I_1 = \eta_{20} + \eta_{02}$$

$$I_2 = (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - 3\eta_{03})^2$$

$$I_4 = (\eta_{30} + \eta_{21})^2 + (\eta_{21} + \eta_{03})^2$$

$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + 3\eta_{03})^2$$
$$+ (3\eta_{12} - \eta_{30})(\eta_{12} + \eta_{30})\left[3(\eta_{30} + \eta_{21})^2 - (\eta_{12} + \eta_{30})^2\right]$$

$$I_6 = (\eta_{20} - \eta_{02})\begin{bmatrix}(\eta_{30}+\eta_{12})^2 - (\eta_{21}+\eta_{03})^2 + \\ 4\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})\end{bmatrix}$$

and one skew orthogonal invariant moment:

$$h = (3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{21})\left[(\eta_{30}+\eta_{21})^2 - 3(\eta_{21}+\eta_{03})^2\right] + (\eta_{30}-3\eta_{12})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{21})^2 - (\eta_{21}+\eta_{03})^2\right]$$

The skew invariant is useful in distinguishing mirror images.

This descriptor can be utilized to accomplish pattern identification not only independent of size, position and orientation, but also independent of parallel projection.

Zernike Moments

Zernike Moments (ZM) are orthogonal moments and can be used to represent shape content of an image with minimum amount of information redundancy. Orthogonal moments allow for accurate reconstruction of the image, and makes optimal utilization of shape information. Zernike Moments (ZM) are widely used in CBIR as shape descriptors. ZM have many desirable properties, viz., rotation invariance and robustness to noise. The complex ZM are derived by projecting the image function onto an orthogonal polynomial over the interior of a unit circle $x^2 + y^2 = 1$ as follows.

$$V_{nm}(x,y) = V_{nm}(p,\theta) = R_{nm}(\rho)\exp(jm\theta)$$

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{(n+|m|)}{2}} -1^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!}\rho^{n-2s}$$

where, *n* is non-negative integer, *m* is an integer such than $n-|m|$ is even and $|m| < n$,

$\rho = \sqrt{x^2+y^2}$, $\theta = \tan^{-1}\left(\frac{x}{2}\right)$. Projecting the image function onto the basis set, results Zernike moments of order *n* with repetition *m* given by

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)V_{nm}(\rho,\theta)$$

where, $x^2 + y^2 \le 1$.

Exact Legendre Moments:

Legendre Moments (LM) are continuous and orthogonal moments, they can be used to represent an image with minimum amount of information redundancy. Many algorithms are developed for the computation of LM , but these methods focus mainly on 2D geometric moments. When they are applied to a digital image, a numerical approximation is necessary. Error due to approximation increases as the order of the moment increases. An accurate method for computing the Exact Legendre Moments (ELM) proposed by Hosney is as follows.

Legendre moments of order $g = (p+q)$ for an image with intensity function $f(x,y)$ are defined as

$$L_{pq} = \frac{(2p+1)(2q+1)}{4}\int_{-1}^{1}\int_{-1}^{1} p_p(x)p_p(y)f(x,y)dxdy$$

Where, $p_p(x)$ is the P^{th} order Legendre polynomial defined as

$$p_p(x) = \sum_{k=0}^{p} a_{kp}x^k = \frac{1}{2^p p!}\left(\frac{d}{dx}\right)^p\left[(x^2-1)\right]^p$$

where, $x \in [-1,1]$ and $p_p(x)$ obeys the following recursive relation

$$p_{p+1}(x) = \frac{2p+1}{(p+1)}xp_p(x) - \frac{p}{p+1}p_{p+1}(x)$$

With $p_0(x) = 1$, $p_1(x) = x$ and P>1.

The set of Legendre polynomials $p_p(x)$ forms a complete orthogonal basis set on the interval [-1, 1]. A digital image of size $N \times N$ is an array of pixels. Centres of these pixels are the points (x_i, y_j).

In order to improve accuracy, it is proposed to use the following approximated form

$$L_{pq} = \frac{(2p+1)(2q+1)}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} h_{pq}(x_i,y_j)f(x,y)$$

Where, $x_i = -1 + \left(i - \frac{1}{2}\right)\Delta x$ and

$y_i = -1 + \left(j - \frac{1}{2}\right)\Delta y$ with i,j=1,2,3…N

$$h_{pq}(x_i,y_j) = \int_{x_i - \frac{\Delta x_i}{2}}^{x_i + \frac{\Delta x_i}{2}}\int_{y_j - \Delta y_j/2}^{y_j + \Delta y_j/2} p_p(x)p_q(y)dxdy$$

This double integration is required to be evaluated exactly to remove the approximation error in computation of Legendre moments. A special polynomial is given as follows.

$$\int p_p(x)dx = \frac{p_{p+1}(x) - p_{p-1}(x)}{2p+1}$$

where, $p \geq 1$, The set of Legendre moments can thus be computed exactly by

$$\tilde{L}_{pq} = \sum_{i=1}^{N}\sum_{j=1}^{N} I_p(x_i) I_q(x_j) f(x,y)$$

Where,

$$I_p(x_i) = \left(\frac{(2q+1)}{(2q+2)}\right)\left[xp_p(x) - p_{p-1}(x)\right]_{u_i}^{u_{i+1}}$$

$$I_q(x_j) = \left(\frac{(2q+1)}{(2q+2)}\right)\left[yp_p(y) - p_{p-1}(y)\right]_{V_j}^{V_{j+1}}$$

Where,

$$U_{i+1} = x_i + \frac{\Delta x_i}{2} = -1 + i\Delta x$$

$$U_i = x_i - \frac{\Delta x_i}{2} = -1 + (i-1)\Delta x$$

Similarly,

$$V_{j+1} = y_i + \frac{\Delta y_j}{2} = -1 + j\Delta y$$

$$V_j = y_i + \frac{\Delta y_j}{2} = -1 + (j-1)\Delta y$$

Equation $\tilde{L}_{pq}$ is valid only for $p \geq 1, q \geq 1$.Further, moment kernels can be generated using $I_p(x_i)$ and $I_q(x_j)$. Computation of ELM using $\tilde{L}_{pq}$ is time consuming.

Hence, ELM can be obtained in two steps by successive computation of 1D q^{th} order moments for each row as follows. By rewriting in separable form

$$\tilde{L}_{pq} = \sum_{i=1}^{N} I_p(x_i) Y_{iq}$$

Where,

$$Y_{iq} = \sum_{i=1}^{N} I_p(y_j) f(x_i, y_j)$$

Where, Y_{iq} is the q^{th} order moment of i^{th} row

Since, $I_o(x_i) = \frac{1}{N}$. Substituting this in $\tilde{L}_{pq}$ results the following

$$\tilde{L}_{oq} = \frac{1}{N}\sum_{i-1}^{N} Y_{iq}$$

The number of ELM of order g is given by $N_{total} = \frac{(g+1)(g+2)}{2}$.

These ELM features are used for CBIR in this work.

III. SVM ALGORITHM

Support Vector Machines (SVMs) are supervised learning methods used for image classification. It views the given image database as two sets of vectors in an ‘ n ’ dimensional space and constructs a separating hyper plane that maximizes the margin between the images relevant to query and the images not relevant to the query. SVM is a kernel method and the kernel function used in SVM is very crucial in determining the performance.

The basic principle of SVMs is a maximum margin classifier. Using the kernel methods, the data can be first implicitly mapped to a high dimensional kernel space. The maximum margin classifier is determined in the kernel space and the corresponding decision function in the original space can be non-linear. The non-linear data in the feature space is classified into linear data in kernel space by the SVMs. This is illustrated in Figure 1 as follows.

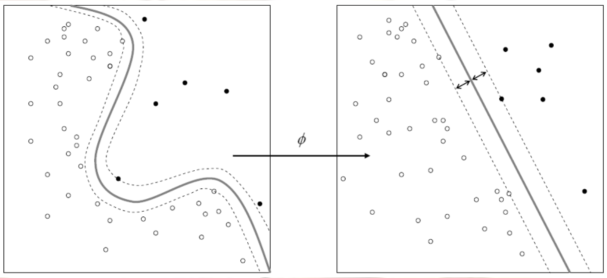

Fig. 1. The function ‘f’ embeds the data in the original space(a) kernel space (b) where the nonlinear pattern now becomes linear.

The aim of SVM classification method is to find an optimal hyper plane separating relevant and irrelevant vectors by maximizing the size of the margin (between both classes). Image classification or categorization is a machine learning approach and can be treated as a step for speeding-up image retrieval in large databases and to improve retrieval accuracy. Similarly, in the absence of labeled data, unsupervised clustering is also found useful for increasing the retrieval speed as well as to improve retrieval accuracy. Image clustering inherently depends on a similarity measure, while image classification has been performed by different methods that neither require nor make use of similarity measures.

Faster and accurate CBIR algorithms are required for real time applications. This can be achieved by employing a classifier such as Support Vector Machine (SVM). SVM is a supervised learning method used for image classification. It views the given image database as two sets of vectors in an ‘ n ’ dimensional space and constructs a separating hyper plane that maximizes the margin between the images relevant to query and the images non relevant to the query. A CBIR system using ELM features and ELM features with SVM as classifier is proposed in this work.

The basic procedure involved in the proposed CBIR system is as follows.

- Computation of ELM for the given image to form the feature vector.
- Calculation of distance measure between the feature vectors of query and data base images.
- Retrieval of similar images based on minimum distance.
- Employ SVM classifier to classify the images in the database
- Increase the number of training samples to improve the classification efficiency.

The steps of proposed CBIR algorithm are as follows.

Exact Legendre moments for the database images are computed by using to form the feature database.

- Feature database is created by feature vector $f_{database} = (f_1, f_2, .. f_M)$ for the image database consisting of M images. Each feature vector f_i for $i = 1,2,..M$, is a set of ELM of order $(p+q) = g$
  $f_i = \{L_{00} L_{01} .. L_{pq}\}_{database}$
- A feature vector comprising of ELM of order $(p+q) = g$ for the query image is formulated.
  $f_q = \{L_{00} L_{01} .. L_{pq}\}_{query}$
- Distance measure between the feature vector f_b of the query image and each feature vector of the database images f_i is calculated by using Canberra distance d_{qi}^c.
  $d_{qi}^c = \sum_{i=1}^{M} \frac{|f_q|}{|f_q| f_i|}$ where, g is the order of moments.
- Retrieve all the relevant images to the query image based on minimum distance d_{qi}^c.
- Train the SVM by selecting proper samples of the database from each class. All the classes of the image database are labeled.
- Pass the class labels with their features to the SVM classifier with the chosen kernel. The Gaussian Radial Basis Function kernel is considered as defined in.
- Classify all the images from the database by considering each image in the database as the query image.

A query image may be any one of the database images. This query image is then processed to compute the feature vector in equation for f_q.The distance d_{qi}^c is computed between the query image $('q')$ and image from database $('i')$. The distances are then sorted in increasing order and the closest sets of images are then retrieved. The top "*N*" retrieved images are used for computing the performance of the proposed algorithm. The retrieval efficiency is measured by counting the number of matches.

IV. RESULTS

Retrieval performance of the proposed CBIR system is tested by conducting experiments on Corel shape database, COIL-20. It consists of 20 classes of images with each class consisting of 72 different orientations resulting in a total of 1440 images. All these gray scale images in the database are of the size 128×128. All images of all the 20 classes are used for experimentation. Experiments are conducted using MATLAB 7.2.0 with Pentium-IV, 3.00 GHz computer and osusvm toolbox. Fig.2 shows the result for order 4, 5, 6, 7, 8, and 9 are considered. It results in feature vectors of dimension 9, 12, 16, 20, 25, and 30 respectively for ZM. Whereas it is 15, 21, 28, 36, 45, and 55 respectively for ELM and Comparative average retrieval efficiency of the CBIR system for various moments and moment orders is presented in Table I. The Fig. 3 shows the classification efficacy of the CBIR system for various moments and moment orders. As the number of training samples increases, the classification efficiency also increases. This is presented in Table II.

TABLE I.

Methods	Moment order						
	4	*5*	*6*	*7*	*8*	*9*	*10*
Hu's Moments	45.20	45.20	45.20	45.20	45.20	45.20	45.20
Zernike Moments	49.09	52.31	52.46	52.57	53.62	54.36	54.26
Proposed(ELM)	68.47	69.30	78.68	82.77	87.50	89.23	92.75

As SVM is a kernel method, the kernel function used in SVM is very crucial in determining the performance. A kernel function needs to be chosen with appropriate parameters. The kernel is tuned with a pre-defined ideal kernel matrix. As a kernel method, SVMs can efficiently handle nonlinear patterns. However, the choice of kernel and tuning of appropriate parameters, adapting SVMs for specific requirements of CBIR such as learning with small sample is a challenging problem. Average retrieval times for order 9 for the CBIR systems based on MI, ZM, and ELM are 0.49, 1.25, and 0.549 seconds respectively. It is observed that MIs are faster but inefficient for CBIR. ELM based CBIR system is faster and efficient compared to other moment based CBIR systems.

TABLE II.

Methods	**Number of Training Samples**			
	4	***5***	***6***	***7***
Hu's Moment	47.71	48.93	50.64	51.41
Zernike Moments	82.01	83.26	84.72	86.74
Proposed(ELM)	82.70	84.37	84.65	87.22

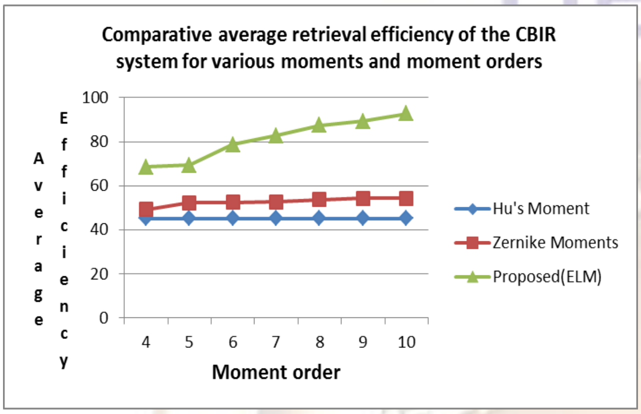

Fig. 2. Comparative average retrieval efficiency of the CBIR system for various moments and moment orders

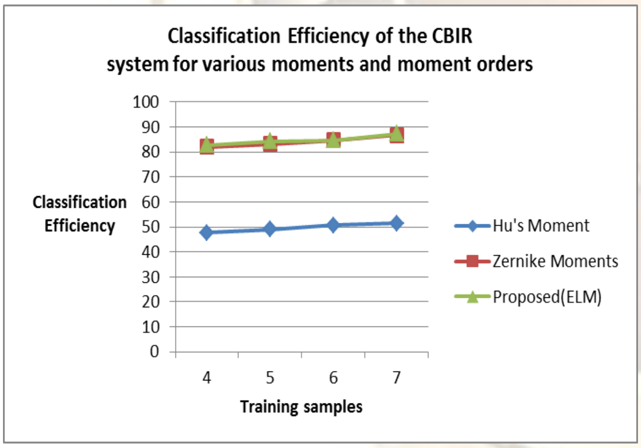

Fig. 3. Comparative classification efficiency of the CBIR system for various moments and moment orders

V. CONCLUSION

A CBIR system using region based shape descriptors is proposed in this work. Performance of the proposed CBIR system is superior compared to Bishnu et al., method on COIL-20 database in terms of average retrieval efficiency and average retrieval time. It is also shown that region based shape descriptors are mainly ELMs perform better compared to other image moments, viz., Hu's, ZM and LM for CBIR applications. Further, improved classification efficiency is also obtained by employing SVM classifier. It is observed that the average retrieval efficiency is increased as the moment order increases. It is also observed that the classification efficiency of the proposed CBIR system increased with the increase in the number of training samples.

REFERENCES

[1] C. Wi06 and J. W. Davis, "A Perceptual Study of Shape Metrics."

[2] C. Rao, S. S. Kumar, and B. C. Mohan, "Content based image retrieval using exact legendre moments and support vector machine," arXiv preprint arXiv:1005.5437, 2010.

[3] N. Santhosh and M. T. M. Kumari, "Content Based Image Retrieval Using Multiwavelet."

[4] K. A. Kumar and Y. B. Reddy, "Content Based Image Retrieval Using SVM Algorithm," International Journal of Electrical and Electronics Engineering (IJEEE), ISSN (PRINT), pp. 2231–5284.

[5] D. N. F. Awang Iskandar, J. A. Thom, and S. M. Tahaghoghi, "Content-based image retrieval using image regions as query examples," in Proceedings of the nineteenth conference on Australasian database-Volume 75, 2008, pp. 38–46.

[6] A. Jain, R. Muthuganapathy, and K. Ramani, "Content-based image retrieval using shape and depth from an engineering database," in Advances in Visual Computing, Springer, 2007, pp. 255–264.

[7] R. R. Eamani and G. H. Prasad, "Content-Based Image Retrieval Using Support Vector Machine in digital image processing techniques."

[8] D. Zhang and G. Lu, "Content-Based Shape Retrieval Using Different Shape Descriptors: A Comparative Study.," in ICME, 2001.

[9] K. Rieck, P. Laskov, and K.-R. Müller, "Efficient algorithms for similarity measures over sequential data: A look beyond kernels," in Pattern Recognition, Springer, 2006, pp. 374–383.

[10] Y. Aytar and A. Zisserman, "Enhancing Exemplar SVMs using Part Level Transfer Regularization," 2012.

[11] D. N. M. Cardoso, D. J. Muller, F. Alexandre, L. A. P. Neves, P. M. G. Trevisani, and G. A. Giraldi, "Iterative Technique for Content-Based Image Retrieval using Multiple SVM Ensembles." J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[12] A. Marakakis, N. Galatsanos, A. Likas, and A. Stafylopatis, "Relevance Feedback for Content-Based Image Retrieval Using Support Vector Machines and Feature

Selection," in Artificial Neural Networks–ICANN 2009, Springer, 2009, pp. 942–951.
[13] M. Vorobyov, "Shape Classification Using Zernike Moments," 2011.

Dr. V.Venkata Rao has his M.E in Microwave and Radar Engineering from Osmania University Hyderabad. He obtained his PHD from JNTU Hyderabad. Presently he is working as Professor and HOD of ECE (Electronics and Communication Engineering) in Narasaraopeta Engineering College, Narasaraopeta, Guntur Dt, Andhra Pradesh, India.
He has over 19 years of experience in industry R&D and teaching. He has published and presented 40 research papers in reputed international /national journals and conferences .He is a life member of ISTE. His research area includes GPS image processing and embedded systems.

Surendra padavala received Bachelor's Degree in Electronics and Communication Engineering from JNTU, Hyderabad. He is Pursuing his M.Tech in Digital Electronics and Communication System in JNTU, Kakinada.
His current research interests are Digital image Processing, Embedded systems and Communication systems