IJCIT
International Journal of Computational Intelligence Techniques
ISSN: 0976–0466 & E-ISSN: 0976–0474, Vol. 2, Issue 2, 2011, PP-53-55
Online available at http://www.bioinfo.in/contents.php?id=26

GENERALIZED GAUSSIAN MIXTURE MODELS BASED IMAGE RETRIEVALS FROM IMAGE DATABASES

MADHUSUDHANA RAO T.V. [1*] AND SETTY S.P.[2]
[1]Department of Computer Science and Engineering, TPIST, Bobbili(India)
[2]Department of Computer Science and Systems Engineering, Andhra University (India)
*Corresponding Author: Email- madhu11211@gmail.com

Received: December 28, 2011; Accepted: December 31, 2011

Abstract- In this paper we present a novel methodology for image classification based on Generalized Gaussian Mixture models together with PCA. In this paper we consider the texture features in the identification of the images. The developed model is evaluated using various quality metrics such as PSNR, and MSE. The bench mark images are considered for experimentation. The output shows that the developed model outperforms the existing models in terms of image retrievals.
Keywords- Generalized Gaussian Mixer Model, Image Quality metrics, PCA Segmentation, Retrieval

Introduction

Today, in this information era, lots of information is stored in the internet using different media. Among these, information pertaining to images are mostly retrieved and many authors have presented different models for retrieving the images based on segmentation and extracting different features such as Texture, color, shape and other attributes [1] [2] [3] [4] [5]. In many of the models proposed, two variations of the problem are mainly focused: Supervised and unsupervised. In the supervised method of learning some features about the image are known in prior and the segmentation is carried out by using these features. In the later, the unsupervised model, the features inside the image are not known beforehand [6]. Here there should be some methods for effective identification of the number of patterns (Textures) inside the image region.

In order to carryout effective image retrievals, in unsupervised data, it is customary to cluster the images, so that the homogenous data are pooled together from which the required patterns can be easily identified. To cluster the data, various clustering algorithms are discussed in literature and among these K-means algorithm is mostly suggested [7].

For effective image retrievals it is necessary to identify the parameters inside each of the image region [8]. In the parametric methods, the parameters inside the image region are extracted and these parameters are used as the key features for effective classification and identification. Most of the research in this direction is carried out with respect to Gaussian Mixture models, but, due to the limitation of the Gaussian mixture models [9], in this article generalized Gaussian Mixture Models are utilized. The advantage of using Generalized Gaussian Mixture Models is that it can handle the images which are asymmetric and symmetric. The paper is organized as follows, Section- 2 of the paper deals with K-means algorithm. Section-3 presents an insight into Generalized Gaussian Mixture Models. In section-4 the method of extracting the feature vectors using PCA is presented, in Section-5, methodologies along with the classification and retrieval algorithm is presented. Section-6 of the paper deals with the experimentation and the performance of the developed model are presented together with conclusions in section 6.

K-Means Algorithm

To cluster the unsupervised data, it is necessary to use any of the clustering techniques. K-Means algorithm is utilized since it clusters the data most efficiently. [9] The most difficult task in K-means algorithm is determining the finite number of regions (K). To initialize the value of K, histogram technique is utilized. The K-Means algorithm is given below

Step 1- Begin with a decision on the value of K= number of segments

Step 2- Put any initial partition that classifies the pixels into K segments. We can arrange the training samples randomly, or systematically as follows.

1. Take the first K training samples as a single-element Segment.
2. Assign each of the remaining (N-K) training samples to the segment with the nearest centroid. Let there be exactly K segments (c_1, $c_{2,\ldots,}c_k$) and n patterns to be classified such that, each pattern is classified into exactly one segment. After each assignment, recompute the centroid of the gaining segment.

Step 3- Take each sample in sequence and compute its distance from the centroid of each of the segments. If the sample is not currently in the cluster with the closest

centroid switch this sample to that segment and update the centroid of the segment gaining the new sample and cluster losing the sample.

Step 4- Repeat step 3 until convergence is achieved, that is until a pass through the training sample causes no new assignments.

After determining the final value of the K (number of regions), we obtain the estimates the parameters μ_i, σ_i and α_i for the i^{th} region using the segmented region pixel intensities using the Generalized Gaussian Distribution.

Generalized Gaussian distribution

In this section we briefly discuss the probability distribution and its properties used in the image classification algorithm. Let the pixel intensities in the entire image follow a Finite Generalized Gaussian Mixture Distribution. It is also assumed that the entire image s a collection of 'K' image regions, then the pixel intensities in each image region follows a Generalized Gaussian Distribution. The probability density function is

$$f\left(\frac{z}{\mu},\sigma,P\right)=\frac{1}{2\tau\left(1+\frac{1}{p}\right)A(P,\sigma)}e^{-\left|\frac{(z_1-\mu_1)}{A(P,\sigma)}\right|^P}$$

Where

$$\sigma>0,\qquad A(P,\sigma)=\left[\frac{\sigma^2\tau\left(1/P\right)}{\tau\left(3/P\right)}\right]^{\frac{1}{2}}$$

The parameter μ is the mean, the function A (P,σ) is a scaling factor which allows that the Var(Z)= σ^2, and 'P' is the shape parameter. When P=1, the corresponding Generalized Gaussian corresponds to a Laplacian or Doubly Exponential Distribution, when P=2, the corresponding Generalized Gaussian corresponds to a Gaussian distribution. In limiting cases P$\rightarrow+\infty$ equation converges to a uniform distribution in $(\mu-\sqrt{3}\sigma,\mu+\sqrt{3}\sigma)$and when P$\rightarrow$0+, the distribution becomes a degenerate one in Z=μ. The mean value of the Generalized Gaussian distribution is μ. The Variance is σ^2.

Feature Vector identification

The steps to be followed for feature vector identification are:

1. Normalize the images
2. Calculate the Eigen values from the training set
3. Calculate the corresponding distribution in M-dimensional weight space for each image

Methodology

In order to demonstrate our methodology, we have developed an image dataset with 50 different gray level images, each of fixed size 320 X 320. The images are first segmented using K-means algorithm to group the related images. For the initialization of K-means algorithm the value of K is assumed by constructing the histogram for each of these images. The parameters μ, σ inside each of these images are identified. The PDF of the images are obtained by using the Probability Density function of the Generalized Gaussian Mixture Model given in section 3 using PDF's and the segmentation is carried out using the Likelihood of the pixels, the feature vectors are identified using PCA and basing on these feature vectors, the images are classified. The classified images are to be evaluated. The image quality metrics are utilized for this purpose. The steps to carry out for effective image retrievals are given as follows.

1. Consider an image from the image Database.
2. Cluster the image using K-Means algorithm
3. Segment the images by using the Generalized Gaussian Mixture Model algorithm.
4. Identify its feature vectors using PCA
5. Evaluate the performance using image quality metrics.

Experimental Results

In order to demonstrate the methodology we have developed the program in the MATLAB environment. The image database is generated with different images and the most relevant images are grouped by clustering. The feature vectors are obtained and after classification the retrieval is performed. The performance evaluation is carried out using quality metrics such as signal to noise ratio, mean square error, image fidelity and structural quality index. The results obtained are tabulated in table 1. From the values of the table 1, it can be clearly seen that the classification algorithm retrieves the images effectively.

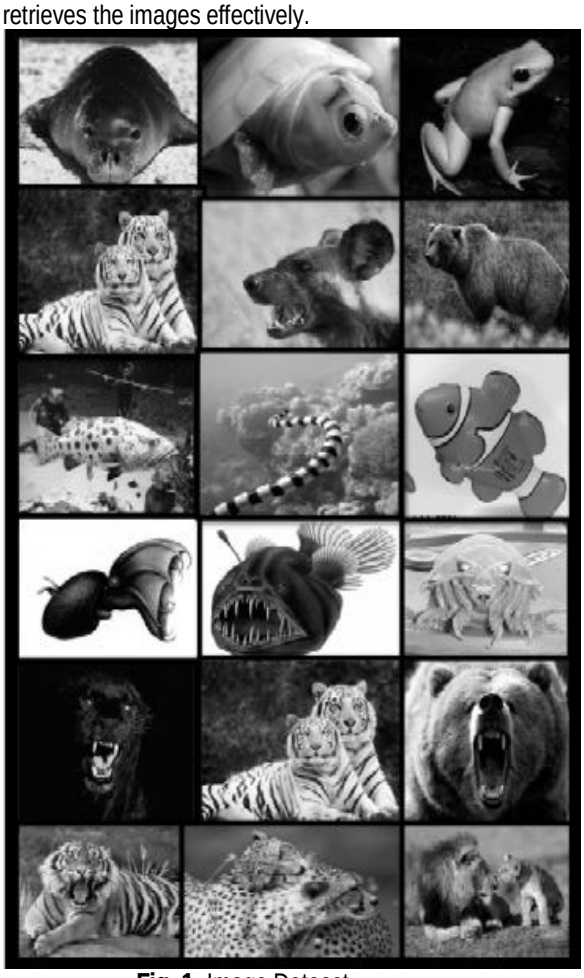

Fig. 1- Image Dataset

Fig. 2- Query Image

Fig. 3- Retrieved Image

Table 1- Performance Evaluation

Metric	Value	Standard Limits
PSNR	35.5	-∞ to +∞
MSE	0.322	0-1

Conclusion

In this paper a novel methodology for effective image classification and retrievals is presented. This model is developed using Generalized Gaussian Mixture Model together with PCA. The advantage of this model is that Generalized Gaussian Distribution includes Gaussian distribution as a particular case. The outputs derived shows that the developed method outperforms the existing methods in classification of the images.

References

[1] Reed T. and du Buf J.M.H. (1993) *CVGIP Image Understanding,* vol. 57, pp. 359–372.

[2] Li J., Najmi A., and Gray R.M. (2000) *IEEE Trans. Sig. Proc.,* vol. 48, no. 2, pp. 517–533.

[3] Povlow B.R. and Dunn S.M. (1995) *IEEE Trans. Patt. Anal. Mach.* Intell., vol. 17, no. 10, pp. 1010–1014.

[4] G. Fan and X.-G. Xia (2000) *Proc. 34th Asilomar Conf. Signals, Systems and Computers*.

[5] Choi H. and Baraniuk R.G. (1999) *Proc. SPIE Tech.Conf. Mathematical Modeling, Bayesian Estimation and Inverse Problems,* pp. 306–320.

[6] Prasad Reddy PVGD et al (2007) *International Journal of Computer Science and Network Security,* VOL.7 No.4.

[7] Sharifi K et al (1995) *IEEE Trans. On Circuits and Systems for Video Technology* Vol 5, No.1, pp 52-56

[8] Yiming WU et al (2003) "*Unsupervised Color Image Segmentation based on Gaussian Mixture Models" Proceedings of 2003 Joint Conference.*

[9] Armando Dominguez J et al (2003) *IEEE Transactions on Image Processing.*