PARAMETER ESTIMATION AND CLASSIFICATION USING GAUSSIANMIXTURE MODEL

[1]P.Prabhakaran, [2]S.Saravanan
[1]PG Scholar, Department of ECE, Surya Engineering College, Erode, India.
[2]Assistant Professor, Department of ECE, Surya Engineering College, Erode, India.
[1]prabhasp.sp@gmail.com, [2]saravananece@gmail.com

Abstract: Gaussian mixture based parameter estimation and classification has recently received great attention in modeling and processing data. Gaussian Mixture Model (GMM) is the probabilistic model for representing the presence of subpopulations and it works well with the classification and parameter estimation strategy. Here in this work Maximum Likelihood Estimation (MLE) based on Expectation Maximization EM) is being used for the parameter estimation approach. With the mean and the variance calculated they are used in the Gaussian Mixture Model (GMM) estimation for the parameters to be estimated and the classification is being done based on parameters estimated.

Keywords: Maximum Likelihood Estimation (MLE), Expectation Maximization (EM), Gaussian Mixture Model (GMM).

1. INTRODUCTION

Classification, as a big part of supervised learning problem, has always attracted lots of attention for its various applications. Also, many methods are brought forward to tackle this problem. Statistical methods manipulate probabilities and try to model the entire hypothesis space and data distribution using probabilities distribution density, thus providing more complete description of the actual problems; however, it also asks for huge complexity to achieve the goal. Mixture models are very preferable in areas where the statistical modeling of data is needed, for instance in signal and image processing, pattern recognition, bioinformatics, and machine learning. Gaussian mixture modelling performs classification by extracting global statistics from Gaussian distributions of pixel intensity in image data set. GMM is especially well-suited for parameter estimation and classification because of implementation facility and efficiency in the representation of data. The histogram gives a main idea about the probability density function (pdf) of pixel values [1]. GMM is a distribution, which consists of finite number of Gaussian distributions in the linear way. The Gaussian distribution is common used for its high level of realistic applicability, similar data behaviors together. In the statistical approaches it has been assumed that pixel values follow a particular distribution and hence mixture model approach is applied on those values. The common practice is to assume that intensity values follow Gaussian distribution with two parameters, mean and variance[2]. The well-known approach for estimating parameters of model is either by Maximum Likelihood Estimation (ML) via Expectation-Maximization algorithm (EM) or Maximum a Posterior Estimation (MAP).

1.1 Methodology

In this paper statistical approach, Gaussian Mixture Model is being used for the parameter estimation and the classification. Since it is the most commonly used method for the estimation of the parameters and better classification and here it is used for the medical images. In this Expectation Maximization algorithm is being used in the Maximum Likelihood Estimation method for the estimation of the parameters. Mean and the Variance are calculated and they are used in the E-step and the M-step in the Expectation Maximization algorithm.

2. Gaussian Mixture Model:

A Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as vocal-tract related spectral features in a speaker recognition system. GMM parameters are estimated from training data using the iterative Expectation-Maximization (EM) algorithm or Maximum *A Posteriori* (MAP) estimation from a well-trained prior model.

A Gaussian mixture model is a weighted sum of M component Gaussian densities as given by the equation,

$$p(X\ \lambda) = \sum_{i=1}^{M} w_i\ g\ X\ \mu_i \Sigma_i \qquad (1)$$

where x is a D-dimensional continuous-valued data vector (i.e. measurement or features), wi, i = 1, . . . ,M, are the mixture weights, and g(x|μi,_i), i = 1, . . . ,M, are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form.

$$g\ X\ \mu_i \Sigma_i\ = \frac{1}{2\pi^{\frac{D}{2}}\ \Sigma_i^{\frac{1}{2}}} exp[-\frac{1}{2}\ x - \mu_i\ '\Sigma_i^{-1}\ x - \mu_i\ ] \qquad (2)$$

with mean vector μi and covariance matrix i. The mixture weights satisfy the constraint that

Equation

The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities. These parameters are collectively represented by the notation,

$$\lambda = \ w_i, \mu_i, \Sigma_i \quad i = 1 \ldots\ldots\ldots.. M \qquad (3)$$

2.1 Maximum Likelihood Parameter Estimation:

Given training vectors and a GMM configuration, we wish to estimate the parameters of the GMM, , which in some sense best matches the distribution of the training feature vectors. There are several techniques available for estimating the parameters of a GMM [3]. By far the most popular and well-established method is maximum likelihood (ML) estimation.

The aim of ML estimation is to find the model parameters which maximize the likelihood of the GMM given the training data. For a sequence of T training vectors X = {x1, . . . , xT }, the GMM likelihood, assuming independence between the vectors1, can be written as,

$$p\ X\ \lambda\ = \prod_{t=1}^{T} p\ X_t\ \lambda \qquad (4)$$

Unfortunately, this expression is a non-linear function of the parameters and direct maximization is not possible. However, ML parameter estimates can be obtained iteratively using a special case of the expectation-maximization (EM) algorithm [4]. The new model then becomes the initial model for the next iteration and the process is repeated until some convergence threshold is reached. The initial model is typically derived by using some form of binary vector quantisation.

On each EMiteration, the following re-estimation formulas are used which guarantee a monotonic increase in the model's likelihood value,

Mixture weights:

$$W_\iota = \frac{1}{T} \sum_{t=1}^{T} p_r\ \ i\ X_t, \lambda \qquad (5)$$

Mean,

$$\mu_\iota = \frac{\sum_{t=1}^{T} p_r\ \ i\ X_t, \lambda\ X_t}{\sum_{t=1}^{T} p_r\ \ i\ X_t, \lambda} \qquad (6)$$

Variance (diagonal Covariance)

$$\sigma_{\iota^2} = \frac{\sum_{t=1}^{T} P_r\ \ X_t, \lambda\ x^2 t}{\sum_{t=1}^{T} p_r\ \ i\ X_t, \lambda} \qquad (7)$$

Where σ_{i^2}, xt, and μi refer to arbitrary elements of the vectors ,xt, and μi, respectively[5].

2.2 Expectation Maximisation Map Algorithm:

In general, EM iterates through two steps to obtain estimates. The first step is an Expectation (E) step, in which missing values are filled-in with a guess, that is, an estimate of the missing value, given the observed values in the data. The second step is a Maximization (M) step, in which the completed data from the E step are processed using ML estimation as though they were complete data, and the mean and covariance estimates are updated. Using the newly updated mean and covariance matrix, the E step is repeated to find new estimates of the missing values.

These two steps (E and M) are repeated until the maximum change in the estimates from one iteration to the next does not exceed a convergence criterion. The result of this process is a mean vector and covariance matrix that uses all available information. In [6]the EM estimates of the mean vector and covariance matrix can then be used in multivariate analyses to obtain estimates of model parameters and standard errors, to test hypotheses, and to score or predict values for observations using the model selected. [7]There are popular algorithm for GMM in several papers and this is modified here as the name of EM-MAP algorithm. The algorithm is described as follows:

Input

Observed Image in a Vector $X_j \, j = 1,2, \dots \dots, n$ and $i \in \{1,2, \dots . k\}$ labels sets

Initialize:

$$\theta^{0} = p_1^{\ 0}, \dots p_k^{\ 0}, \dots \mu_1^{\ 0}, \dots . \mu_k^{\ 0}, \dots \sigma_1^{2\ 0}, \dots \sigma_k^{2\ 0} \quad (8)$$

E-Step:

$$p_{ij}^{(r+1)} = p^{\ r+1}\ \ i\ X_j \ = \frac{p_i^{(r)} N\ \ x_j\ \mu_i^{(r)}, \sigma_i^{2(r)}}{f\ \ x_j} \quad (9)$$

M-Step:

$$p_i^{\ r+1} = \frac{1}{n} \sum_{j=1}^{n} p_{ij}^{\ r} \quad (10)$$

$$\mu_i^{(r+1)} = \frac{\sum_{j=1}^{n} p_{ij}^{(r+1)}}{n p_i^{(r+1)}} X_j \quad (11)$$

$$\sigma_i^{2(r+1)} = \frac{\sum_{j=1}^{n} p_{ij}^{(r+1)}}{n p_i^{(r+1)}} \ \ x_j - \mu_i^{(r+1)}$$

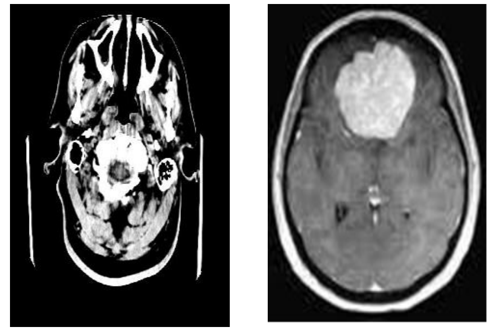
Iterate steps 3 and 4 until an specify error (i.e) $\sum_i e_i^2 < \varepsilon$.

Compute:

$$p_{ij} = ArgMax_i \quad p_{ij}^{(final)}, j = 1,2, \dots, n \quad (12)$$

3. Experimental Results:

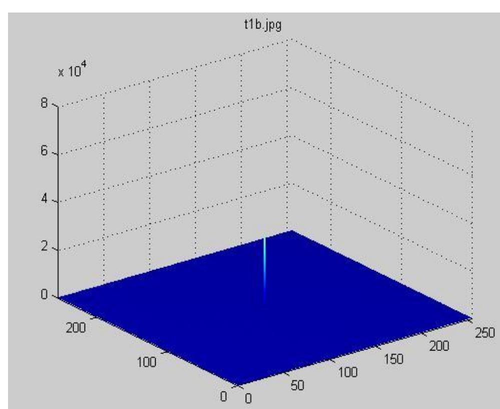
The calculation for the mean, probability and the variance are being carried out here with respect to the medical images as the input. For each part taken with different scanning methods have the corresponding change in their mean and then each iteration the corresponding values of the number of classes, mean, probability and the variance are being calculated and are used in esc iteration. the variance values. The parameters estimated here are the mean and the variance and according to this the classification can be done with respect to the classes. The classified image for the respective images is also obtained with the labels.

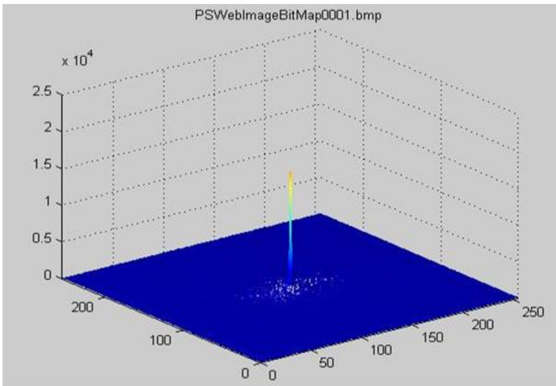


(a) Original image

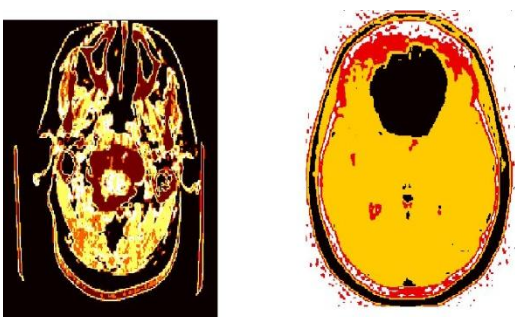
image	k	Convergence steps	mean	sigma
1	5	132	0.876	0.6754
	6	91	0.9887	1.5670
	7	74	0.6573	0.9865
2	5	364	0.8765	1.6574
	6	120	1.6785	0.7865
	7	105	0.8923	0.4653

(b)Table for the parameters estimated using the expectation maximization in maximum likelihood





(c)Gaussian plot for the classified images by the expectation maximization



(d)classification of images by the expectation maximization algorithm

4. Conclusion

The parameters have been estimated and the classifications of the respective images are obtained with the respective labels. In further work the statistical classifiers must be used with the expectation maximization and the maximum likelihood estimation as an algorithm for the training and the testing of the image datasets with the Gaussian based mode classifier.

References:

[1] T. Elguebaly, N. Bouguila, Bayesian learning of finite generalized Gaussian mixture models on images, Signal Processing 91 (2011) 801–820.

[2] J. Liu, Y.B. Ku, S. Leung, Expectation-maximization algorithm with total variation regularization for vector-valued image segmentation, Journal of Visual Communication and Image Representation 23 (2012) 1234–1244.

[3] McLachlan, G., ed.: Mixture Models. Marcel Dekker, New York, NY (1988)

[4] Dempster, A., Laird, N., Rubin, D.: Maximum Likelihood from Incomplete Data via the EMAlgorithm. Journal of the Royal Statistical Society 39(1) (1977) 1– 38

[5] H. Permuter, J. Francos, I.H. Jermyn, Gaussian mixture models of texture and colour for image database retrieval, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Hong Kong, China, 2003

[6] Buck, S.F. (1960). ―A method of estimation of missing values in multivariate data suitable for use with an electronic computer‖. Journal of the Royal Statistical Society, B22, 302-306.

[7] M.A.T.Figueiredo, Bayesian Image Segmentation Using Gaussian Field Prior, EMMCVPR 2005, LNCS 3757, pp. 74-89.

[8] J. Tohka, I. D. Dinov, D. W. Shattuck, and A. W. Toga,"Brain MRI tissue classification based on local Markov random fields", Magn. Reson. Imag., vol. 28, no. 4, pp. 557-573, 2010.

[9] C. Nikou, N. Galatsanos, and A. Likas, "A class-adaptive spatially variant mixture model for image segmentation", IEEE Trans. Image Process., vol. 16, no. 4, pp. 1121-1130, Apr. 2007.

[10] D. Titterington, A. Smith, and U. Makov, "Statistical Analysis of Finite Mixture Distributions", New York: Wiley, 1985.

[11] T. Kanungo, D. M. Mount, N. Netanyahu, C. Piatko, R. Silverman, & A. Y.Wu (2002) An efficient k-means clustering algorithm: Analysis and implementation Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp.881-892.