[1]K Naveen Kumar

[2]G V S Raj Kumar

[3]B Raja Koti

[4]M C Jagan Sekhar

[5]P V Pragnatha

[6]P V Sai Manisha

Multivariate Approach for Texture Segmentation using Probabilistic Statistical Model

Journal of Electrical Systems

Abstract: - The analysis of the regions of the image is of the prerogatives in the fields of medical and global systems meant for location identification. This analysis is strongly associated with partitions of regions of interest such as segmentation. For an effective strategy of analyzing the regions of interest, texture of the image plays a major concern. The texture is generally characterized using signal processing methods namely Discrete Cosine Transformation coefficients and their specific insights leading to feature vector selection. Further, to identify regions, a statistical model needs to be identified for the feature matrix vector and thus make use of Gaussian mixture model with extensions. The Expectation Maximization approach is used, and performance is assessed by experimenting with random images from the Brodatz data store domain. Performance measurements for texture segmentation that can be attributed are Global Cons. Error (GC), Prob. Rand. Index (PR) and Variation of data (VA). These are determined alongside the confusion matrix. To assess the improvement, a comparison was made with other existing models and showed better. The algorithms will be exceptionally helpful for clinical analysis and in radio navigation map systems.

Keywords: Image regions, Statistical Models, Signal Processing Methods, Feature matrix, Gaussian approach, Performance assessment

I. INTRODUCTION

Identification of the regions of interest from a textural image is one the primary requirement of the proposed model and this is attributed by a matrix of data items termed as feature vectors. In literature, authors had proposed models for identifying the regions of interest which are based on various approaches by making use of the attributes namely features, regions, edges, histograms, thresholds, neurons, data model and SVM. In these non-parametric approaches, pixel behavior is given very little importance, which is why these techniques have not been able to identify regions more adequately. In parametric modelling, the pixel behavior is considered and to account for, model-based techniques for segmentation i.e., identifying regions of interest are more effective. (Suetens P. et. al. (1991), Srinivasa Rao, K et al. (2007), Dilpreet Kaur et.al. (2014), K. Naveen et.al. (2015), Y.Li. et. al., (2016), Verhoeven et al. (2019). Most of the models relate to the various properties of texture image namely color, shape, size, or depth and could be able to provide better results when the variation in pixels is small in spatial domain. In reality, the disruption in the pixel distribution due to techniques used for capturing images with devices is the major problem and can be addressed with texture-based segmentation to some extent using signal processing methods. Hence the extraction of required data namely features i.e., properties of texture play a major role for effective identification of regions. To aid the process, a suitable Gaussian probability mixture model of multivariate nature is developed and

[1] Associate Professor, GST, GITAM Deemed to be University, A.P, India

[2] Professor, GST, GITAM Deemed to be University, A.P, India

[3] Assistant Professor, GST, GITAM Deemed to be University, A.P, India

[4] Assistant Professor, CSE, Lendi Institute of Engineering & Technology, AP, A.P, India

[5] Assistant Professor, CSIT, Lendi Institute of Engineering & Technology, AP, A.P, India

[6] Assistant Professor, CSIT, Lendi Institute of Engineering & Technology, AP, A.P, India

nkuppili@gitam.edu[1], gganapav@gitam.edu[2], rbadugu@gitam.edu[3],

mchandujagansekhar@gmail.com[4], pragnatha.puvvada@gmail.com[5], manumanisha1313@gmail.com[6]

*Corresponding author: nkuppili@gitam.edu[1]

utilized for effective identification of regions of interest. The article presents the generalized Gaussian model of multivariate nature considering Coefficients of Signal processing methods namely Discrete Cosine Transformations (DT) along with use of Logarithmic domain (LD) and Local binary pattern (LP) approach for compensating light variations and inter pixel dependencies. Further, Section 2 of the article presents the approach for identifying feature vector data and suitable probability model with EM algorithm inclusion for model parameter estimation and revised estimates. Section 3 presents the experimental and comparative performance results and then draws conclusions.

II. IDENTIFICATION OF FEATURE VECTORS AND THEIR EXTRACTION ALONG WITH PROPOSED MODEL

Attributes play an important role during the segmentation process. These attributes, also called features, help to correctly identify the object of interest in a more elegant way (Kather et al. (2016). Many methods for texture-based feature extraction are available in the literature like Principal Comp. Analysis, Linear Disc. Analysis, etc. (Ricciardi, C. et al. (2021), Nagila et. al. (2021). The pixels in regions of interest share some common properties which are very much essential for effective segmentation. Adjacent pixels of image share the spatial relationships. It is clearly stated that signal processing method i.e., DT which attributes to dimensionality reduction yields better results when considered as Feature vector. The extraction is done by looking into each pixel area and decomposing in to signals while maintaining visual abilities. It also transforms from spatial to frequency domain thereby better inherent relationships can be considered and extracted. Another aspect for considering DT is because of minimum quadratic reconstruction error. The model is further extended by making use of applying logarithmic function which facilitate in reducing the dependency of illumination effects. Further, it is to state that these focus on macro information of pixels. However, images also contain micro level information which possess sensitive data points pertaining to neighboring pixels. These data points can be captured by utilizing the local binary pattern approach. The approach is the image is first processed with LP to capture the binary patterns i.e., micro information and then DT is applied for effective segmentation. Accordingly, these feature vectors serve as input to model extracted from images of Brodatz and assessed for performance.

The suitable probability model under consideration is of the form as shown in Eq.(1).

$$p(X/\theta) = \sum_{i=1}^{M} WG\,(X/\theta) \qquad \text{- Eq.(1)}$$

where W, G are weights and Gaussian Prob. Distribution function between 1 to M for feature vector X.

Based on the above equation, the proposed model (GD) of D - dimension is given by

$$\mathrm{GD}(\mathrm{X}/\theta) = \prod_{\mathrm{j}=1}^{\mathrm{D}} \frac{\mathrm{sh_j K(sh_j)}}{2\mathrm{sig_j}} \exp\mathrm{o\,nent}\left\{\left(-\mathrm{A(sh_j)}\left|\frac{\mathrm{x_j - mu_j}}{\mathrm{sig_j}}\right|^{\mathrm{sh_j}}\right\} \quad \text{- Eq. (2)}$$

where, mu, sig and sh are mean, variance and curve shape parameters of distribution that control the function.

The other terms are

$$K(sh_j) = \frac{\Gamma(3/sh_j)^{1/2}}{\Gamma(1/sh_j)^{3/2}} \quad \& \quad A(sh_j) = \left[\frac{\Gamma(3/sh_j)}{\Gamma(1/sh_j)}\right]^{sh_j/2} \text{with gamma function.} \quad \text{- Eq. (3)}$$

Using the above equations, the model parameters can be estimated by the EM algorithm, which maximizes functionality as needed. The DT coefficients represented as X serve as input to the model. In order to arrive at the final estimates of weights, mu and sig., the log. prob. function or expected log likelihood should be maximized and the corresponding equations are listed below.

$$w_i^{(l+1)} = \frac{1}{T}\sum_{r=1}^{T}\left[\frac{w_i^{(l)}.GD(.)}{\sum_{i=1}^{M} w_i^{(l)}.GD(.)}\right] \qquad \text{- Eq. (4)}$$

$$mu_{ij}^{(l+1)} = \frac{\sum_{r=1}^{T} t_i(X,\theta^{(l)})^{A(N,sh_{ij})}(x_{tij})}{\sum_{r=1}^{T} t_i(X,\theta^{(l)})^{A(N,sh_{ij})}} \quad \text{- Eq. (5)}$$

$$sig_{ij}^{(l+1)} = \left[\frac{\sum_{r=1}^{T} t_i(X,\theta^{(l)}) \left(\frac{\Gamma\left(\frac{3}{sh_{ij}}\right)}{sh_{ij}\Gamma\left(\frac{1}{sh_{ij}}\right)} \right) . \left| x_{rij} - mu_{ij} \right|^{\frac{1}{sh_{ij}}}}{\sum_{r=1}^{T} t_i(X,\theta^{(l)})} \right]^{\frac{1}{sh_{ij}}} \quad \text{- Eq. (6)}$$

To start with one needs to have initial estimates and therefore Hierarchical clustering is used as it gives better results (Srinivasa Rao, K et al. (2007). The common approach is to draw sample data at random and assign the values to starting estimates. The final or refined estimates are derived by iterating the equations simultaneously in standard software environment i.e., MATLAB.

III. RESULTS, EXPERIMENTS AND COMPARATIVE STUDY

The segmentation algorithm includes the following steps.

Step 1: Use the method described in Section 2 to obtain the feature vector.

Step 2: Use hierarchical clustering algorithm to divide samples into M groups.

Step 3: Mean vector and variance vector) of each multivariate data category is calculated.

Step 4: consider weights $w_i = 1/M$, i iterating from 1 to M.

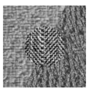
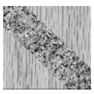
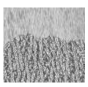
Step 5: Use the update equation of the EM algorithm to obtain a final estimate of w_i, $mu_{ij} and sig_{ij}$ for each category.

Step 6: The assignment of the feature vector is based on the Eq. (7)

$$L_j = \max \left\{ \prod_{j=1}^{D} \frac{\exp onent\left(-\left| \frac{x_{ij} - mu_{ij}}{A(\rho_{ij}, sig_{ij})} \right|^{\beta_{ij}} \right)}{2\Gamma\left(1 + \frac{1}{sh_{ij}}\right) A(sh_{ij}, sig_{ij})} \right\} \quad \text{- Eq. (7)}$$

In order to assess the model with proposed three scenarios, texture image quality metrics namely GC, PR and VA (Sokolova, M et.al. (2006)) are calculated for check of improvement and shown in Table I.

TABLE I. SEGMENTATION METRICS ASSESSMENT

RESULTS OF THE EVALUATION PROCESS					
Picture Description		**Scenario**	**S1**	**S2**	**S3**
	Picture 1	PMH	0.732	0.222	1.060
		PMHLDT	0.858	0.195	1.038
		PMHLDTLP	0.821	0.354	1.085
	Picture 2	PMH	0.721	0.380	1.923
		PMHLDT	0.657	0.150	1.720
		PMHLDTLP	0.825	0.138	1.648
	Picture 3	PMH	0.848	0.152	1.242
		PMHLDT	0.781	0.146	1.206
		PMHLDTLP	0.899	0.128	0.824
	Picture 4	PMH	0.625	0.552	1.301
		PMHLDT	0.729	0.521	1.221
		PMHLDTLP	0.789	0.105	1.251
	Picture 5	PMH	0.628	0.558	1.231

		PMHLDT	0.636	0.519	1.189
		PMHLDTLP	0.788	0.438	1.184

*PMH – PROB. MODEL WITH H, PMHLDT – PROB. MODEL WITH H AND LOG DT AND PMHLDTLP – PROB. MODEL WITH H AND LOGARITHMIC DT+ LP & *S1 – PR, S2 – GC, S3 – VA

Using the values of confusion matrix of the segmented regions, we calculated the accuracy, sensitivity, specificity, accuracy, recall and F-measure as shown in Table II.

TABLE II. ILLUSTRATIVE COMPARISON OF MODEL(S) WITH VARIANTS OF DT, LP AND LOG DT

Description	Scenario	M1	M2	M3	M4	M5	M6
Picture 1	PMH	0.81	0.81	0.16	0.85	0.81	0.83
	PMHLDT	0.89	0.89	0.12	0.89	0.89	0.89
	PMHLDTLP	0.92	0.88	0.1	0.82	0.88	0.89
Picture 2	PMH	0.74	0.81	0.15	0.82	0.81	0.81
	PMHLDT	0.82	0.88	0.11	0.74	0.88	0.84
	PMHLDTLP	0.86	0.92	0.12	0.75	0.92	0.86
Picture 3	PMH	0.89	0.84	0.21	0.89	0.84	0.86
	PMHLDT	0.92	0.87	0.16	0.94	0.87	0.90
	PMHLDTLP	0.95	0.93	0.14	0.96	0.93	0.94
Picture 4	PMH	0.69	0.78	0.20	0.72	0.78	0.74
	PMHLDT	0.82	0.84	0.15	0.91	0.84	0.87
	PMHLDTLP	0.85	0.9	0.08	0.88	0.9	0.89
Picture 5	PMH	0.78	0.78	0.24	0.68	0.78	0.73
	PMHLDT	0.82	0.82	0.16	0.74	0.82	0.78
	PMHLDTLP	0.85	0.88	0.13	0.78	0.88	0.83

* Metrics: M1- Accuracy, M2 – Sensitivity, M3 -1-Specificity, M4 -Precision, M5 -Recall, M6 - F-Measure.

It clearly indicates that the best segmentation result is obtained over previous models. Especially the F-dimensional value of the proposed classifier on comparison to the previous model, shows good sign of improvement.

IV. CONCLUSION

This article presents some algorithms for identification of texture regions using multidimensional generalized Gaussian mixture models. Some other variations of feature vectors with more general distribution families (such as Pearson distribution hybrid system) can be used to design better segmentation algorithms. A hierarchical clustering algorithm is used to define M components which is the basis for model parameter estimation. The work can be further extended by considering MDL, MML and AIC criteria that define the number of components for effective segmentation. These models with several other types of feature vectors help in identifying regions of interest in medical imaging techniques like CT scan, resonance imaging and surveillance systems with location identification. The experimental results showcased a better outcome. As new technologies are evolving in imaging, a more focus on model-based techniques is needed and be explored further for developing more effective systems with augmented results.

REFERENCES

[1] Ricciardi, C. et al. (2020) ‘Linear discriminant analysis and principal component analysis to predict coronary artery disease’, Health Informatics Journal, pp. 2181–2192.

[2] K. Naveen Kumar, K. Srinivasa Rao, Y. Srinivas, Ch. Satyanarayana (2015), “Texture Segmentation based on Multivariate Generalized Gaussian Mixture Model under log DCT domain”, International Journal of Applied Engineering Research, Vol.10 Issue No.22, pp. 43045-43051.

[3] Dilpreet Kaur et.al. (2014), Various Image Segmentation Techniques :A review",International Journal of Computer Science and Mobile Computing, Vol.3 Issue.5, May- 2014, pg. 809-814

[4] Sokolova, M, Japkowicz, N, Szpakowicz, S. Beyond accuracy F-score and ROC: a family of discriminant measures for performance evaluation. In: Australasian joint conference on artificial intelligence, Hobart, TAS, Australia, 4–8 December 2006, pp. 1015–1021. Berlin: Springer.

[5] Srinivasa Rao.K and Srinivas. Y (2007), "Unsupervised image segmentation using finite doubly truncated Gaussian mixture model and hierarchical clustering", Journal of Current Science, Vol.93(4), pp.507-514.

[6] Verhoeven, J.; Miller, N.R.; Daems, L.; Reyes-Aldasoro, C.C. Visualisation and Analysis of Speech Production with Electropalatography. J. Imaging 2019, 5, 40

[7] Kather, J.N.; Weis, C.A.; Bianconi, F.; Melchers, S.M.; Schad, L.R.; Gaiser, T.; Marx, A.; Zollner, F. Multi-class Texture Analysis in Colorectal Cancer Histology. Sci. Rep. 2016, 6, 27988

[8] Suetens P., Verbeeck R., Delaere D., Nuyts J., Bijnens B. (1991) Model-Based Image Segmentation: Methods and Applications. In: Stefanelli M., Hasman A., Fieschi M., Talmon J. (eds) AIME 91. Lecture Notes in Medical Informatics, vol 44. Springer, Berlin, Heidelberg.

[9] Y. Li, F. Jia, and J. Qin, Brain tumor segmentation from multimodal magnetic resonance images via sparse representation, Artif Intell Med 73, (2016), 1– 13.

[10] Nagila, Ashish and Nagila, Ritu, ADVANCED FACE RECOGNITION FOR CONTROLLING CRIME USING PCA (April 29, 2021).

[11] Ashish Nagila, Ritu Nagila and Shelly Bhardwaj, Advanced Face Recognition for Controlling Crime Using PCA, International Journal of Advanced Research in Engineering and Technology (IJARET), 12(2), 2021, pp. 657-663