Novel Approach for Color based Comic Image Segmentation for Extraction of Text using Modify Fuzzy Possiblistic C-Means Clustering Algorithm

M.Praneesh
Assistant Professor,
Department of Computer Science,
Sankara College of Science and Commerce,

R.Jaya Kumar
Assistant Professor,
Department of Computer Science,
Sankara College of Science and Commerce,

ABSTRACT

Segmentation in image processing refers to the process of partitioning a digital image into multiple segments. This paper makes an attempt to segment the Comic images for extraction the text. Segmentation of comic images into extract the text is a challenging task primarily because of complexity of the structural Features like color, shape and texture. In this paper we proposed a color based comic image segmentation for extraction of text using Modify Fuzzy Possiblistic C-Means Clustering Algorithm has been tested on different images and obtained better performance than many of the existing methods.

Keywords

Segmentation, comic image, text extraction, Fuzzy Possiblistic C-Means Clustering.

1.INTRODUCTION

The image segmentation is an effort to classify similar colors of image in the same group. It clusters color into several groups based on the closeness of color intensities inside an image. Segmentation refers to the process of partitioning a digital image into multiple regions. All image processing operations generally aim at a better recognition of objects of interest that is finding suitable local features that can be distinguished from the other objects. Color segmentation of the image is an important operation in the image analysis. In many computer vision image interpretation, and pattern recognition plays with vital role in scientific and industrial fields such as medicine, remote sensing and microscopy, content based image retrieval, document analysis etc. in this paper we proposed a color based comic image segmentation for extraction of text using Modify Fuzzy Possiblistic C-Means Clustering Algorithm has been tested on different images and obtained better performance than many of the existing methods.

1.1. Existing Work

In the previous work, unsupervised clustering technique for segmentation a color is based on the K-Means algorithm. The K-Means algorithm is well known for its simplicity and low complexity. However the algorithm has three main drawbacks, first dependency on the initial centroids, Second dependency on the initial centroids, third is degenercency. We can accommodate these three issues by the Modify Fuzzy Possiblistic C-Means algorithm. In this paper Modify Fuzzy Possiblistic C-Means algorithm based text extraction is vital role in image processing and apply this algorithm to make reduce the computation time.

1.2. Proposed Work

The present work is organized as follows Section-A Describes the color transformation of image. Section-B describes Fuzzy Possiblistic C- Means clustering algorithm. Section-C describes the stretch model constructing. Section D-describes the proposed method for segmenting the image based on color with Modify Fuzzy Possiblistic C -Means clustering is presented and discussed.

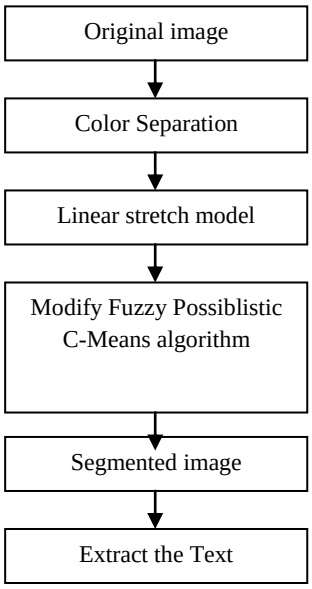

Fig-1 System Architecture

2.COLOR SEPARATION

In this process, we can convert image from RGB colour space to L*a*b* colour spaces. L*a*b*colour spaces is derived from CIE*YZ values, the L*a*b* space consists of a luminously layer 'L*'chromatically layer 'a*' indicating where colour falls along red-green axis and chromatically layer 'b*' indicating where the colour falls along the blue yellow axis. All the colour information is in the 'a*'and 'b*' layers. We can measure the difference between two colours using the Euclidean distance metric and convert the image into L*a*b*. Euclidean distance is used for similarity matching in the present system. The Euclidean distance between two points

p = (p1, p2 …pn) and Q= (q1, q2….qn) is defined as

$$\sqrt{(p1-q1)^2+(p2-q2)^2+\cdots..(pn-qn)^2}$$

3. STRETCH MODEL CONSTRUCTING

After the clustering process the image is transformed into membership space from gray level space. A simple linear sketch model is designed to smoothly enhance each cluster sketching in one of the image enhancement techniques which can be used to improving or locally adjusting the image picture elements. So that it is better to view in dark and bright portion of the image and linear sketch model is performed the equations is given as

$$mi(g) = \frac{g-bi,L}{bi,u-bi,L} * (L-1)$$

where mi (g)= stretched gray value, g = original gray value, L = number of gray level, bi,u= upper boundary for stretching of i^th cluster, bi,L = lower boundary for stretching of i^th cluster.

4. MODIFY FUZZY POSSIBLISTIC C-MEANS CLUSTERING ALGORITHM

The choice of an appropriate objective function is the key to the success of the cluster analysis and to obtain better quality clustering results; so the clustering optimization is based on objective function. To meet a suitable objective function, we have started from the following set of requirements. The distance between clusters and the data points assigned to them should be maximized and the distance between clusters is modeled by term; it is the formula of the objective function [11]. Also Wen-Liang Hung proposed a new algorithm called Modified Suppressed Fuzzy C-Means (MS-FCM), which significantly ameliorates the performance of FCM due to a prototype-driven learning of parameter α. The learning process of α is based on an exponential separation strength between clusters and is updated at each

$$\alpha = \exp\left(-min\frac{||vi-vj||^2}{\beta}\right)$$

Where β is a normalized term so that we choose β as a sample variance. But the remark which must be mentioned here is the common value used for this parameter by all the data at each iteration, which may induce in error. We propose a new parameter suppresses this common value of α and replaces it by a new parameter like a weight to each vector. Or every point of the data set has a weight in relation to every cluster. Therefore this weight permits to have a better classification especially in the case of noise data.Where w_{ji} is weight of the point j in relation to the class i.

This weight is used to modify the fuzzy and typical partition. All update methods that were discussed in section II are iterative in nature, because it is not possible to optimize any of the objective functions reviewed directly. Or to classify a data point, cluster centroid has to be closest to the data point, it is membership; and for estimating the centroids, the typically is used for alleviating the undesirable effect of outliers.

The objective function is composed of two expressions: the first is the fuzzy function uses a fuzziness weighting exponent; but the two coefficients in the objective function are only used as exhibitor of membership and typicality. A new relation, lightly different, enabling a more rapid decrease in the function and increase in the membership and the typically when they tend towards 1 and decrease this degree when they tend towards 0. This relation is to add Weighting exponent as exhibitor of distance in the two under objective functions. The objective function of the MFPCM can be formulated as follows:

$J_{MFPCM}=$

$$\sum_{i=1}^{c}\sum_{j=1}^{n}\left(\bigcup_{ij}^{m} w_{ji}^{m} d^{2m}(x_j,v) + t_{ij}^{n} w_{ji}^{n} d^{2n}(x_j,v_i)\right)$$

$U = \{u_{ij}\}$ represents a fuzzy partition matrix, is defined as:

$$u_{ij} = \left[\sum_{k=1}^{c}\left(\frac{d}{d}\right)^{-1}\right.$$

$T = \{t_{ij}\}$ represents a fuzzy partition matrix, is defined as:

$$t_{ij} = \left[\sum_{k=1}^{c}\left(\frac{d?x_j v_i}{d?x_j v_k}\right)\frac{2n}{(n-1)}\right]^{-1}$$

$V = \{v_{ij}\}$ represents a fuzzy partition matrix, is defined as:

$$v_i = \frac{\sum_{j=1}^{n}\left(u_{ij}^{m} \quad w_{ji}^{m} +\right.}{\sum_{j=1}^{n}\left(u_{ij}^{m} \quad w_{ji}^{m}\right.}.$$

5. RESULTS

The various experiment carried out on the above said imagery in MATLAB vs 7.10. The complete process and the standard results are summarized in subsequent figures.

Fig-2 original and cluster index image

Fig-3 extraction of text from the cluster1 image

Fig-4 extraction of text from the cluster2 image

6.CONCLUSION

In this paper color based image segmentation of text extraction. It is possible to reduce computation time and improve the precision of image. The main objective of this paper is to get a non-overlapping and reliable output. Color based segmentation of image is an analysis such as image interpretation, pattern recognition system with application in scientific and industrial fields. Furthermore we analyse various clustering algorithms with text blob extraction methods like ballon detection.

7.REFERENCES

[1] Li, J., Wang, J. Z. and Wiederhold, G.,"Integrated Region Matching for ImageRetrieval," ACM Multimedia, 2000, p. 147-156.

[2] Flickner, M., Sawhney, H., Niblack, W.,Ashley, J., Huang, Q., Dom, B., Gorkani, M.,Hafner, J., Lee, D., Petkovic, D., Steele, D.and Yanker, P., "Query by image and videocontent: The QBIC system," IEEE Computer,**28(9)**, 1995,pp.23-32

[3] Pentland, A., Picard, R. and Sclaroff S.,"Photo book: Content based manipulation of image databases", International Journal ofComputer Vision, **18(3)**, 1996, pp.233–254

[4] Smith, J.R., and Chang, S.F., "Single color extraction and image query," In Proceeding IEEE International Conference on ImageProcessing, 1997, pp. 528–531

[5] Gupta, A., and Jain, R., "Visual information retrieval," Comm. Assoc. Comp. Mach., **40(5)**, 1997, pp. 70–79.

[6] M. Saadatmand-Tarzjan and H. A. Moghaddam, "A Novel Evolutionary Approach for Optimizing Content-Based Image Indexing Algorithms", IEEE Transactions On Systems, Man, And Cybernetics—Part B: Cybernetics, Vol. 37, No. 1, February 2007, pp. 139 153.

[7] N. Vasconcelos, "From Pixels to Semantic Spaces: Advances in Content-Based ImageRetrieval",Computer Volume: 40, Issue: 7, 2007, pp. 20-26.

[8] N. Rasiwasia and N. Vasconcelos, "A Study of Query by Semantic Example", 3rd International Workshop on Semantic Learning and Applications in Multimedia, Anchorage, June 2008, pp. 1-8.

[9] N. Rasiwasia, P. J. Moreno and N. Vasconcelos, "Bridging the Gap: Query by Semantic Example", IEEE Transactions On Multimedia, Vol. 9, No. 5, August 2007, pp. 923-938.

[10] S. Cheng, W. Huang, Y. Liao and D. Wu, "A Parallel CBIR Implementation Using Perceptual Grouping Of Block-based Visual Patterns", IEEE International Conference on Image Processing – ICIP, 2007, pp. V -161 - V - 164.

[11] Ricardo Dutra da Silva et al; Satellite image segmentation using wavelet transforms based on color and texture features, ISVC 2008, part II, LNCS 5359, 2008

[12] Robert A Schowengerdt, Remote sensing- models and Methods for Image Processing, IIIrd edition, Elsevier Inc.

[13] M.Swain and D. Ballard,"Color indexing", International Journal of Computer Vision, Vol.7, no. 1,1991.

[14] W. Hsu, T.S. Chua, and H. K. Pung, "An Integrated color-spatial approach to Content-Based Image Retrieval", ACM Multimedia Conference

[15] Ka-Man Wong, Chun-Ho Chey, tak-Shing Liu, Lai-Man Po, "Dominant color image retrieval using merged histogram", Circuits and Systems,ISCAS'03 Proceedings of 2003 International Symposium, Vol. 2, 2003

[16] Ju Han and Kai-Kuang Ma, "Fuzzy color Histogram and its use in color image retrieval", IEEE Transactions on Image Processing, Vol. 11, No. 8, 2002.

[17] Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J., "Color image segmentation: Advances and prospects", Pattern Recognition 34, 2259–2281, 2001.

[18] Liew, A.W., Yan, H., Law, N.F., "Image segmentation based on adaptive cluster prototype estimation", IEEE Trans. Fuzzy Syst. 13 (4), 444–453, 2005.