M-FISH Image Segmentation Using Fuzzy Logic and Spatial Information

Lijiya A
Department of CSE
National Institute of Technology
Calicut, Kerala, India.
lijiya@nitc.ac.in

Sreejithlal G S
Department of CSE
National Institute of Technology
Calicut, Kerala, India.
sreejithlalgs@gmail.com

Govindan V K
Department of CSE
National Institute of Technology
Calicut, Kerala, India.
vkg@nitc.ac.in

***Abstract* — Karyotype refers to display of the chromosomes of a cell by lining them up. Multicolor fluorescence in-situ hybridization (M-FISH) technique provides color karyotyping that allows simultaneous analysis of numerical and structural abnormalities of whole human chromosomes. In this paper, a modified Fuzzy c-means (FCM) clustering for M-FISH segmentation is presented. A conventional FCM algorithm does not fully utilize the spatial information in the image. Proposed method uses a FCM algorithm that incorporates spatial information into the membership function for segmentation. And also uses a median filter as preprocessing step, which leads to improved accuracy of pixel classification. The algorithm has been tested on an M-FISH database, which demonstrates improved performance in segmentation when compared with standard FCM clustering-based algorithm.**

Keywords — Image segmentation, Fuzzy C-Means (FCM) clustering, Multiplex fluorescence in-situ Hybridization (M-FISH), spatial information.

I. INTRODUCTION

A chromosome is a microscopic structure which contains an individual's genetic information. It is composed of DNA and various proteins. Chromosomes are classified by their size and appearance under the microscope. A normal individual acquires 23 chromosomes from the father and another 23 from the mother, making it a total of 46 chromosomes (22 pairs of non-sex or autosomal chromosomes and a pair of sex chromosomes). The non-sex chromosome pairs are numbered 1 through 22, and the two sex chromosomes are designated X and Y. Females are characterized by an XX chromosome pair, whereas males are characterized by an XY chromosome pair. A typical metaphase chromosome consists of two identical arms called chromatids which touch at a primary constriction called the centromere. The short chromosome arm is designated as *p* and the long arm as *q*. The sister chromatids are often so close to each other, that the whole chromosome appears like a single rod-like structure.

Karyotyping orders and arranges the chromosomes by their lengths, from longest to shortest, followed by the sex chromosomes. These come handy in accurately diagnosing genetic factors behind various diseases. Manual karyotyping is not only time-consuming and expensive, but also needs well trained personnel. During the early periods of chromosome analysis, gray scale images were used. A new staining method called M-FISH [1], developed in 1996, produces color images. Images are captured with a fluorescent microscope with multiple optical filters. M-FISH uses 5 probes to identify all chromosomes in unique colors. A sixth probe, DAPI (4',6-diamidino-2-phenylindole), after counterstaining with DAPI all chromosome are visible in the corresponding image. Using M-FISH technique greatly simplifies karyotyping and detection of subtle chromosome aberrations.

II. RELATED WORKS

A number of different approaches are available for M-FISH classification and segmentation. A few of them are briefly discussed in this section.

Petros *et al.* [2] presented chromosome classification using M-FISH chromosome images. M-FISH used five different dyes and feature vector will be 5 set of pixels. In extracted pixel sets, feature normalization is applied in order to increase accuracy, and then K-means algorithm is used to cluster chromosomes into 24 clusters. This is a semi-supervised method because it uses some prior information from a labeling chart to find initial centroids.

Mehul *et al.* [3] presented automatic pixel-by-pixel classification algorithm for M-FISH images using a Bayesian Classifier. Main objective of this work is to classify human chromosomes. Bayes classifying technique is used to classify images by finding posterior probabilities using prior probabilities. The classification problem is modelled as a 25 class problem. Yu-Ping Wang [4] improves pixel-by-pixel classification algorithm using a multi-resolution image registration, which is applied as a pre-processing step. The image registration is done as follows. First DAPI image is selected as the reference image and other five images are registered to it. The classifier is tested without image registration and with image registration, and results show that image registration improves classification accuracy.

Yu-Ping Wang *et al.* [5] proposed a FCM clustering algorithm for classification of chromosome analysis. The

problem is modeled as a 24 class problem. The main objective of FCM algorithm is to find the cluster centroids. Dissimilarity function is used as the objective function. By minimizing dissimilarity function, cluster centroids are found. Data normalization is done in order to increase the accuracy of classification. This is done by image registration, color compensation etc. FCM classifier yields better accuracy than Bayesian classifier. Classification can be improved by Normalization.

Petros and Dimitrios [6] proposed a region based decorrelation stretching method, which is applied to chromosome image classification. First, the image is decomposed into homogenous regions using the multichannel watershed transform. Then for each segmented area the intensities of the pixels belonging to that region are replaced with the mean color value of the region. Then decorrelation stretch is applied. The decorrelation stretch is a process that is used to enhance the color differences found in a color image.

Petros S. Karvelis *et al.* [7] proposes a new method for the multichannel image segmentation and region classification. This method is a two stage process. In the first step multi-channel image segmentation is carried out by watershed transform. After applying transform, image is decomposed into set of homologous regions. And after this a binary mask is applied to remove artefacts that appear in some channel. In the second step, a Bayes classifier is applied to regions.

Hyohoon Choi *et al.* [8] proposed a method for automatic segmentation and classification for M-FISH chromosome images. First Background correction is done using image subtraction. Color compensation is done in next step; here color spreading is corrected by linear transformation. Then filtering is done to remove shot noise and additive Gaussian noise, median filter is used to remove noise, followed by feature normalization is applied. Classification is carried out using Maximum likely hood function.

Hongbao Cao *et al.* [9] proposed an improved adaptive FCM algorithm for chromosome segmentation and classification. The algorithm improves the classical FCM algorithm by the use of a gain field, which models and corrects intensity inhomogeneities caused by a microscope imaging system, airs of targets (chromosomes), and uneven hybridization of DNA. Improved Adaptive FCM improves error rates and compensates back ground image intensity. This method avoids the computation of large differential equations.

Wade Schwartzkopf *et al.* [10] proposed a minimum entropy segmentation method for M-FISH images. To segment chromosome, first find two cut points. Joining those points may give a cut line, which separates the touching or overlapping chromosome. The method is improved by introducing an objective function as a criterion for selecting cut lines to decompose groups of chromosomes that touch and overlap each other. The objective function is a measure of entropy. This algorithm uses multi-spectral information in chromosome images for more accurate segmentation.

Sreejini K S *et al.* [11] proposed an automated method for M-FISH chromosome segmentation and classification. The method first removes nuclei and debris from M-FISH images. Watershed method [12] is applied to segment the overlapped chromosomes. Mean and standard deviation of each segmented area are the features used for classification. And classification is done using Bayes classifier for 24 class problem. And a post processing step is also included which reclassify segments to most likely class of one of its neighbor's. Results are much better than pixel-by-pixel classification. This method was tested on a small data set. The accuracy of method for larger data sets is unknown.

Nemanja B. Grujic *et al.* [13] proposed a hybrid approach for image segmentation from chromosome images. The main disadvantages of segmentation algorithm are that they are not universally applicable for every image and these techniques are not perfect. The proposed method can be summarized in four phases – Histogram analysis, Threshold modification, Flood fill algorithm and object extraction. In Histogram analysis, background intensity, color and tolerance are found, and each pixel is classified as object or not, according to a threshold value. Then Flood fill algorithm is applied iteratively to find boundaries of important objects. The final stage is object extraction. Based on relative size noise, background and chromosome are extracted. This method gives good results and minimizes user interaction.

Lijiya A *et al.* [14] proposes a new method for M-FISH segmentation and classification. This method includes preprocessing steps for cell removal and noise reduction. Median and low pass filters are used for noise removal. A dilation operation follows, which reduces intensity inhomogeneities. Then three different methods are carried out in order to separate foreground and background from image. Majority voting is used on the three methods to select a segmentation method. After this step, classification of image is done using a fuzzy logic classifier. This classifying technique is very simple. The results show that, this method has high accuracy.

Karyotyping needs very high accuracy, and it is observed that fuzzy logic algorithms have higher accuracy compared to other methods. Also, fuzzy classifiers are unsupervised. This paper uses the concept of fuzzy logic classification. The proposed algorithm uses FCM incorporating spatial information into the membership function for segmentation. The algorithm has been tested on an M-FISH database.

The rest of the paper is organized as follows: Standard FCM is discussed in Section III, and Spatial FCM is presented in Section IV. The Section V deals Results. And finally the approach is concluded in Section VI.

III. FUZZY C-MEANS ALGORITHM

FCM [15, 16] is a clustering method that allows a datum to belong to more than one cluster. Based on the distance between each cluster center and a data point, FCM algorithm assigns each data point, membership to each cluster center. Nearer the data is to the cluster center, higher is its membership towards the particular cluster center. Sum of membership-value of each data point to all cluster centers must be equal to one. Following

each iteration, membership and cluster centers are updated [17] according to the formula

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}} \quad (1)$$

$$v_j = \frac{\sum_{i=1}^{n} (\mu_{ij})^m x_i}{\sum_{i=1}^{n} (\mu_{ij})^m}, \quad \forall j = 1, 2, \ldots, c \quad (2)$$

where

- n is the number of data points;
- v_j represents the number of clusters;
- m is the fuzziness index, $m \in [1, \infty]$;
- c represents the number of cluster centers;
- μ_{ij} is the membership of i^{th} data to j^{th} cluster center; and
- d_{ij} is the Euclidean distance between i^{th} data and j^{th} cluster center.

Main objective of FCM algorithm is to minimize

$$J(U, V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (\mu_{ij})^m \|x_i - v_j\|^2 \quad (3)$$

where

- $U = (\mu_{ij})_{n \times c}$ is the fuzzy membership matrix
- V is set of cluster centroids
- $\|x_i - v_j\|$ is the Euclidean distance between i^{th} data and j^{th} cluster center.

A. FCM Algorithm

Let $X = \{x_1, x_2, x_3, \ldots, x_n\}$ be the set of data points and $V = \{v_1, v_2, v_3, \ldots, v_c\}$ be the set of centers.

1) Randomly select c cluster centers.
2) Calculate the fuzzy membership μ_{ij} using (1).
3) Compute the fuzzy centers v_j using (2).
4) Repeat steps 2 and 3 until the value of J is minimized or $\|U^{(k+1)} - U^{(k)}\| < \beta$.

where

- k is the iteration step;
- β is the termination criterion between [0, 1];
- $U = (\mu_{ij})_{n \times c}$ is the fuzzy membership matrix; and
- J is the objective function.

The M-FISH image contains nuclei and debris along with chromosomes. First nuclei and debris are removed. Then FCM algorithm is used to segment the image, which means background and foreground are separated. Here, the algorithm can be used as a 2 class problem.

IV. Spatial Fuzzy C-Means Algorithm

A Spatial FCM algorithm [17, 18] proposed for MRI image segmentation is incorporated in this work for M-FISH image segmentation. An important characteristic of an image is that neighboring pixels possess similar feature values, and there is a greater probability that they belong to the same cluster. This spatial relationship is important in clustering, but standard FCM algorithm does not utilize this. The spatial function exploits spatial information [18, 19] and is defined as

$$h_{ij} = \sum_{k \in NB(x_j)} \mu_{ik} \quad (4)$$

where $NB(x_j)$ represents a square window, whose center coincides with pixel x_j in the spatial domain. A 3 × 3 window was used in this paper. The spatial function of a pixel for a cluster is large if the majority of its neighborhood belongs to the same cluster. The spatial function is incorporated into membership function [18, 19] as

$$\mu'_{ij} = \frac{\mu_{ij}^p h_{ij}^q}{\sum_{k=1}^{c} \mu_{ik}^p h_{ik}^q} \quad (5)$$

where p and q are parameters to control the relative importance of both functions. The initial clustering is similar to standard FCM algorithm, and later, the new spatial function (4) is used and the FCM iteration proceeds with the new membership that is incorporated with the spatial function. The iteration stops when the maximum difference between two membership matrices at two successive iterations becomes less than a threshold.

Median filtering is widely used in digital image processing because, it has the peculiarity of preserving edges while removing noise. It is a non-linear spatial filtering technique. Each pixel value is replaced by the median of intensities in the window (usually of size 3 × 3) centered at that pixel. Median filter is applied to M-FISH images in order to remove the noise in images. This method is used as a preprocessing technique and improves the results of spatial FCM algorithm.

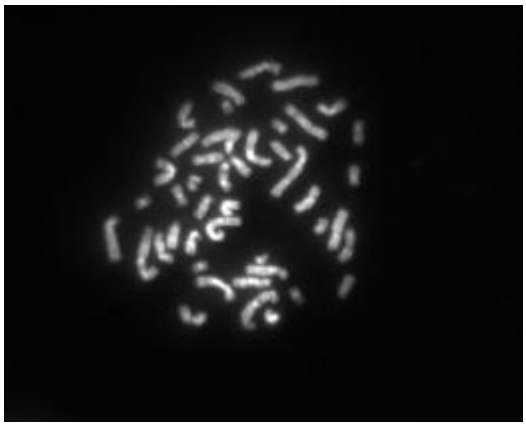
V. Results

M-FISH images from M-FISH database [19] of 40 cells were tested, and the results of image segmentation were compared over the proposed Spatial FCM algorithm and standard FCM algorithm. The value of fuzziness index (m) was taken as 2. The SFCM control parameters (p and q) were taken as $p = 1$, $q = 1$ to ensure equal relative importance to the membership function (μ_{ij}) and the spatial function (h_{ij}) in the new membership function (μ'_{ij}).

The performance of the segmentation was evaluated with the correct detection rate (*CDR*), which is given by

$$CDR = \frac{\#\text{ chromosome pixels correctly segmented}}{\#\text{ total chromosome pixels}} \quad (6)$$

Table I shows segmentation CDR and the results indicate that SFCM segmentation yields 5% more correct detection rate. And SFCM with Median filter improves the accuracy. Fig. 1 shows a DAPI image and Segmented Image using Spatial FCM with median filtering.



(a) Original DAPI Image

(b) Segmented Image

Figure 1. Segmentation Results

TABLE I. SEGMENTATION RESULTS USING FCM, SFCM AND SFCM WITH MEDIAN FILTER

Dataset #	FCM	SFCM	SFCM + M
1	91.7806	95.1923	95.0654
2	96.4028	97.8148	97.8354
3	93.8313	96.2083	96.6518
4	84.1407	92.6429	93.12
5	86.5508	92.9434	93.0624
6	93.0220	95.9436	96.0437
7	78.934	93.0772	93.3727
8	90.8215	95.167	95.8745
9	87.7386	91.4384	92.2633
10	89.5666	95.6486	95.7905
Avg.	**89.27889**	**94.60765**	**94.90797**

VI. CONCLUSION

M-FISH technique is used for color karyotyping permitting simultaneous analysis of numerical and structural abnormalities of whole human chromosomes. Though there are a number of attempts to improve the accuracy of karyotyping, the performances are still constrained by the chromosome image quality. In this paper, a Spatial FCM algorithm is proposed, which incorporates the spatial information into the membership function to improve the segmentation results. This method has a higher correct detection rate (*CDR*) compared to standard FCM algorithm. Accuracy is further improved using the median filter for preprocessing. The future work is to run and test the algorithm on large data sets to ensure better segmentation CDR.

REFERENCES

[1] T Ried, A Baldini, T C Rand, and D C Ward. "*Simultaneous visualization of seven different DNA probes by in situ hybridization using combinatorial fluorescence and digital imaging microscopy*", Proc Natl Acad Sci U S A, 89(4):1388 – 1392, 1992.

[2] P. Karvelis, A. Likas, and D.I. Fotiadis. "*Semi unsupervised M-FISH chromosome image classification*", In Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on, pages 1 – 4, November 2010.

[3] M.P. Sampat, K.R. Castleman, and A.C. Bovik. "*Pixel-bypixel classification of MFISH images*", In Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint, volume 2, pages 999 – 1000, 2002.

[4] Yu-Ping Wang. "*M-FISH image registration and classification*", In Biomedical Imaging: Nano to Macro, 2004. IEEE International Symposium on, pages 57 – 60, April 2004.

[5] Y.-P. Wang and Ashok Kumar Dandpat. "*Classification of M-FISH images using fuzzy C-means clustering algorithm and normalization approaches*", In Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on, volume 1, pages 41 – 44, November 2004.

[6] P.S. Karvelis and D.I. Fotiadis. "*A region based decorrelation stretching method: Application to multispectral chromosome image classification*", In Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, pages 1456 – 1459, October 2008.

[7] P.S. Karvelis, D.I. Fotiadis, A. Tzallas, and I. Georgiou. "*Region Based Segmentation and Classification of Multispectral Chromosome Images*", In Computer-Based Medical Systems, 2007. CBMS '07. Twentieth IEEE International Symposium on, pages 251 – 256, June 2007.

[8] Hyohoon Choi, K.R. Castleman, and A.C. Bovik. "*Joint segmentation and classification of M-FISH chromosome images*", In Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE, volume 1, pages 1636 – 1639, September 2004.

[9] Hongbao Cao and Yu-Ping Wang. "*Segmentation of M-FISH Images for improved classification of chromosomes with an adaptive fuzzy c-means clustering algorithm*", In Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, pages 1442 – 1445, 30 2011-april 2 2011.

[10] W. Schwartzkopf, B.L. Evans, and A.C. Bovik. "*Minimum entropy segmentation applied to multi-spectral chromosome images*", In Image Processing, 2001. Proceedings. 2001 International Conference on, volume 2, pages 865 – 868, October 2001.

[11] Sreejini K S, Lijiya A, and V K Govindan. "*Minimum entropy segmentation applied to multi-spectral chromosome images*", In International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), volume 2, pages 105 – 117, April 2012.

[12] Lijiya A, Sreejini K.S., and V.K. Govindan. "*M-FISH Chromosome Images Classification by Watershed Based Segmentation Approach*", In Proceeding of Futuristic of Computer Science Engineering and Information Technology-ICCT2012, volume 2, pages 501 – 505, March 2012.

[13] N.B. Grujic, N.V. Ilijic, and D. Jankovic. "*Hybrid approach to image segmentation applied on human karyotype determination*", In Telecommunications in Modern Satellite, Cable and Broadcasting Services, 2005. 7th International Conference on, volume 1, pages 193 – 196, September 2005.

[14] Lijiya A, Sangeetha M.K., and V.K. Govindan. "*Segmentation and Classification of M-FISH Human Chromosome Images*", In The second International Conference on Advances in Computing and Communications (ACC-2012), August 2012.

[15] J.C. Dunn. "*Some recent investigations of a new fuzzy partition algorithm and its application to pattern classification problems*", In Cybernetics, pages 1 – 15, 1974.

[16] J.C. Bezdek. "*Pattern Recognition with Fuzzy Objective Function Algorithms*", Plenum, New York, 1981.

[17] Indah Soesanti, Adhi Susanto, SriWidodo Chen, Jeng-Chang Thomas, and Maesadji Tjokronagoro. "*MRI Brain Images Segmentation Based on Optimized Fuzzy Logic and Spatial Information*", International Journal of Video & Image Processing and Network Security IJVIPNS-IJENS, 11(4):6 – 11, August 2011.

[18] Tzong-Jer Chen, Keh-Shih Chuang, Sharon Chen, Jeng-Chang Lu, and Ya-Hui Shiao. "*A novel image smoothing filter using membership function*", J Digit Imaging, 20(4):381 – 392, 2007.

[19] Keh-Shih Chuang, Hong-Long Tzeng, Sharon Chen, Jay Wu, Tzong-Jer Chen. "*Fuzzy c-means clustering with spatial information for image segmentation*", Computerized Medical Imaging and Graphics, vol-30, 2006, pp9-15.

[20] The ADIR M FISH Image Database. "*http://www.adires.com/05/Project/MFISHDB/MFISHDB.shtml*", Accessed : 21/07/2012.