SEGMENTATION AND CLASSIFICATION OF MULTISPECTRAL CHROMOSOME IMAGES

Mohammed Alhanjouri [1], Fatma E.Z. Abou-Chadi [2], and Nadder Hamdy [1]
[1]AAST, Alexandria, Egypt, [2]Mansoura University, Mansoura, Egypt. E-mail :-< Alhanjouri@aast.edu>

Abstract

This work presents an automated method for chromosome segmentation and classification. It takes the advantage of the multi-spectral information in M-FISH chromosome images and utilizes the Discrete Wavelet Transform (DWT) and Bayes rule. The results of the suggested method have shown a high performance of 100%, 92%, and 95% correct classification for three different imaging systems.

Keywords: Image Processing, Karyotyping, MFISH Images, Bayes rule, Wavelet Transform.

I. INTRODUCTION

An important part of genetics analysis starts with researchers breaking down the nucleus of a human cell into a jumbled cluster of chromosomes that are then stained with dye so that they can be studied under a microscope (Fig.1). This jumble of stained chromosomes, which carry the genetic code of the host individual, is then photographed, creating a chromosome spread image (see Fig.2(a)). This image is subsequently subject to a procedure called chromosome karyotyping analysis [1, 2]. The result of this procedure is a karyotype image (Fig.2(b)), the standard form used to display chromosomes. In this configuration, the chromosomes are ordered by length from the largest (chromosome 1) to the smallest (chromosome 22 in humans), followed by the sex chromosomes. Karyotype images are used in clinical tests, such as amniocentesis, to determine if all the chromosomes appear normal and are present in the correct number.

Because karyotype images are much more useful for screening and diagnostic purposes than chromosome spread images, chromosome karyotyping analysis is routinely performed in clinical and cancer cytogenetic labs. It involves using modern image processing techniques for individual chromosome segmentation, enhancement, orientation, and classification. The key to chromosome karyotyping analysis is automated chromosome classification, which has been an outstanding object recognition problem for decades [3]. The goal is to automate the laborious and expensive visual recognition of the chromosome images. Image enhancement is typically employed as a preprocessing step prior to chromosome classification.

In the mid-1990's, a technique for staining chromosomes was introduced [4, 5]. It produced an image in which each chromosome type appeared to be a distinct color. This multi-spectral staining technique made analysis of chromosome images easier, not only for visual inspection of the images by humans, but also for computer analysis of the images. The multi-spectral staining technique is called M-FISH (multiplex fluorescence in-situ hybridization). M-FISH uses five color dyes that attach to various chromosomes differently to produce a multi-spectral image, and a sixth dye that attaches to all chromosomes to produce a grayscale image.

In humans, the 46 chromosomes consist of 23 pairs of chromosomes, one of each pair coming from the father and the other from the mother. Of the 46 chromosomes, there are 22 homologous pairs and two sex chromosomes denoted X and Y (Fig. (2)). A normal human female has two X chromosomes, while a normal male has an X and a Y chromosome. By convention, the 22 pairs and the X chromosome and Y chromosome are assigned to 24 distinct classes [6,7].

The most important visible chromosome features are: the relative chromosome size within the cell, the relative position of the centromere (a well-marked constriction in the chromosome shape) within the chromosome body [6,7], and the typical banding pattern associated with each chromosome class. Within a cell there is a range of sizes, morphology and banding pattern [7-9]. Size and shape can discriminate the 46 human chromosomes into 7 groups known as Denever classification [7],[10]

The present work seeks to develop an automated method to take advantage of the color information in M-FISH images to improve on past methods of computer analysis of chromosome images. It utilizes image processing

techniques to chromosome image processing and enhancement and introduces a probabilistic model of M-FISH chromosomes which can be used for simultaneous segmentation and classification.

II. METHODOLOGY

The database is made available from "Courtesy Advanced Digital Imaging Research, LLC, League City, Texas 77573". The database consists of three different groups of data. Each group was collected by special microscope to produce five images by five filters and gray-level scale as sixth image. Each filter has distinct absorption and emission wavelengths as shown in Table 1.

Table 1 Three Different combinations of Fluorescent labels

GROUPS	Name of Filter	Absorption (nm)	Emission (nm)
First Combination (1st group)	Spectrum Green	495	521
	Spectrum Orange	547	573
	Texas Red	583	603
	Cyan (Cy5)	649	670
	Cyan (Cy5.5)	674	694
Second Combination (2nd group)	DEAC	426	480
	FITC	491	515
	532	532	555
	568	568	571
	Cyan (Cy5)	649	670
Third Combination (3rd group)	Spectrum Aqua	432	472
	Spectrum Green	495	521
	Spectrum Gold	524	550
	Spectrum Red	581	596
	Far Red	590	612

II.1 Image Preprocessing

The first step is the background cancellation. This is based on adaptive thresholding the gray-level image of biological objects on a microscope slide as shown in Figure 3. Valley searching attempts [4] to find a valley of gray values that represent a separation between two chromosomes. This method often works well for finding accurate boundaries, however, it does not handle overlaps.

II.2 Discrete Wavelet Transform

The chromosome images have the size of 645×517 pixels, therefore, the discrete wavelet transform (DWT) was used [12] as a filtration and down-sampling techniques to extract a suitable number of features. The DWT coefficients for two dimensional signals such as images are divided into four types which are approximate, detailed vertical, detailed horizontal and detailed diagonal coefficients. The approximate coefficients are used to reduce the size of images to 325×260 pixels as illustrated in Fig.4.

II.3 Normalization

The approximate coefficients must be normalized to simplify the complex probability calculation in the next step. The basic normalization technique is to determine the maximal pixel value and to divide all pixels over it.

II.4 Pixel by Pixel Classification

The Bayes rule was used to classify each pixel in the normalized approximation image. The probability of occurrence a certain chromosome (C) at a certain pixel (X) is calculated as a posteriori probability [13, 14]:

$$P(C_i \mid X) = \frac{P(X \mid C_i) \cdot P(C_i)}{P(X)} \qquad \text{(1)}$$

where $P(X|C_i)$ is the probability of pixel (X) belongs to the chromosome (C) which can be calculated by convolving the five types of spectra for currently pixel (X) with standard table of the same spectra related to the imaging technique used. Table 2 shows standard values for each chromosome in each spectrum image, for example, using ASI microscope, the chromosome Number 1 appears in spectrum green, red texas and cyan5 images, obviously; while it is unclear or blurred in spectrum orange and cyan5.5. Each modern microscope has a catalog containing a standard table related to the used filter. For simplicity, the minimum error rate was first calculated and then the probability was computed as

$$P(X|C_i) = 1 - \frac{\left|sp_1 - exac_i^1\right| + \left|sp_2 - exac_i^2\right| + \ldots + \left|sp_5 - exac_i^5\right|}{5} \quad (2)$$

where sp_1, sp_2, sp_3, sp_4 and sp_5 are actual spectra of the pixel (X), and $exac^1$, $exac^2$, $exac^3$, $exac^4$ and $exac^5$ are standard values of the same spectra.

P(X) is the summation of probabilities of pixel (X) that can be calculated as follows:

$$P(X) = \sum_i P(X|C_i) \quad (3)$$

where, (i) is the number of chromosomes [13, 14].

$P(C_i)$ is the probability of occurrence a certain chromosome (C_i) as a priori probability, which reflects the prior knowledge about the chromosomes. The relative size of each chromosome can be used as a priori probability.

For each pixel, 24 posteriori probabilities were calculated to make a decision according to maximal posteriori probability. After scanning all the pixels and classifying each pixel, the resultant image has a 325×260 pixels (Fig. 5), therefore zeros must be added between each two neighbored pixels to restore original size of 645×517 pixels.

II.5. Denoising

The added zeros are aliens that contribute noise to the image; therefore, a suitable filter must be used. Using an interpolation approach may cause pixel misclassified. The majority filter was chosen because it removes small segments and maintains the shape and position of large scale edges [3, 4], A majority filter consists of a structuring element H and the image was scanned in raster order, and the class at the center pixel location was replaced by the majority class within spatial extent of the structuring element H. Mathematically,

$$y(m) = \underset{k \in H,\ (m-k) \in \{O_i\}}{Maj} \{x(m-k)\} \quad (4)$$

where x is the input pixel map, y is the output pixel, and Maj denotes the majority operation, notice that only object pixels are used for calculating the majority, not background pixels $\{O_i\}$, H is the structuring element, for example, a 3×3 square window is defined as

$$H = \{(-1,-1),(-1,0),(-1,1),\ldots.(1,0),(1,1)\} \quad (5)$$

A 5×5 window was used to scan the image and to replace a center pixel with the majority values.

Since pixel classification is an inherently noisy process, some isolated pixels and small segments would be misclassified in this step. To reduce the effect of this noise, a 3×3 window majority filtering was used as a second stage to produce the image shown in Fig 6.

Table 2. The Spectra of Three Different Combinations

ASI M-FISH (SKY) Kit					
Chromosome Class	Spectrum Green	Spectrum Orange	Texas Red	Cy5	Cy5.5
1	X		X	X	
2					X
3	X	X		X	X
4	X			X	
5	X	X	X		X
6	X		X	X	X
7			X	X	
8	X				
9	X	X			X
10				X	X
11	X	X		X	
12			X		X
13	X	X			
14			X		
15		X	X	X	
16	X		X		
17				X	
18	X	X	X		
19		X		X	
20		X			
21	X				X
22		X	X	X	X
X		X			X
Y	X			X	X

PSI M-FISH Kit					
Chromosome Class	Deac	FITC	532	568	Cy5
1			X		
2				X	
3	X				
4		X		X	
5			X		X
6		X			
7			X	X	
8				X	X
9					X
10	X		X		X
11	X			X	
12		X	X		
13	X	X			
14		X	X	X	
15	X		X	X	
16		X			X
17		X		X	X
18			X	X	X
19		X	X		X
20	X			X	X
21	X	X	X		
22	X	X		X	
X	X				X
Y	X		X		

Vysis M-FISH Kit					
Chromosome Class	Spectrum Aqua	Spectrum Green	Spectrum Gold	Spectrum Red	Far Red
1			X		
2				X	
3	X				
4		X		X	X
5			X		X
6		X			
7					X
8				X	X
9			X	X	
10	X		X		
11	X			X	
12		X	X		
13	X	X			
14		X	X	X	
15	X		X	X	
16		X			X
17		X		X	
18			X	X	X
19		X	X		X
20	X			X	X
21	X	X	X		
22	X	X		X	
X	X				X
Y	X		X		X

II.6 Segmentation

Having applied the previous steps to the chromosome image, each pixel is now classified as a point in a certain chromosome; therefore each chromosome can be segmented by collecting all the pixels belonging to this chromosome, however, overlapping prevents chromosome segmentation.

To solve the overlapping problem, the medial axis transform can be used to find the medial axis of the cluster by measuring the width along a transverse line perpendicular to the tangent of the edges. The upper chromosome can be segmented by extracting the pixels belonging to it to leave the lower chromosome alone, but not complete, therefore, completing the lower chromosome must be done by connecting the medial axis and adding parallel lines to medial axis and repeating until filling all the space between the edges of the chromosome as shown in Fig.7

III. RESULTS

For this work, 230 images (about 10580 individual chromosome) wereused to test the proposed algorithm. The results have shown a high performance of 100%, 92%, and 95% for three different imaging systems as shown in Table 3

Table 3. The recognition rate of chromosome images

Microscope type	Number of images	Number of correct images	Percentage of correct classification
ASI (1st group of images)	90	90	100.00%
PSI (2nd group of images)	50	46	92.00%
Vysis (3rd group of images)	90	85	94.44%

From the resultant karyotype image (Fig. 8), it can be noted that the image sample was taken from a male subject because of the existence of one chromosome X and one chromosome Y. Moreover, Chromosome 15 is abnormal because one of the chromosome pair is missing. In addition, a part from one of the chromosome 12 is missing. Therefore, this can be considered as an abnormal case. Fig. 9 shows another example.

IV. CONCLUSION

Multi-spectral chromosome images (M-FISH) were used to develop an automated method to segment and classify the human chromosomes and to decompose both overlaps and clusters composed of more than two chromosomes. Using wavelet transform to reduce and filter the chromosome images, Bayes decision theory was used to classify the human chromosome image pixel-by-pixel.

The use of the majority filter was appropriate to reduce misclassified pixels without creating interpolation values that may belong to another chromosome. The majority filter applied to segmented chromosome to fill any pixel surrounded with chromosome pixels.

The proposed method was applied to images of three different imaging systems to gives 100%, 92% and 95% for ASI M_FISH Kit, PSI M_FISH Kit, and Vysis M_FISH Kit, respectively.

REFERENCES

[1] M. Speicher, S. Ballard, and D. Ward, "Karyotyping human chromosomes by combinatorial multi-fluor FISH", Nature Genet. Vol. 12, 368-375, 1996.

[2] S. Müller , M. Rocchi, M. Ferguson-Smith and J. Wienberg, "Toward a multicolor chromosome bar code for the entire human karyotype by fluorescence in situ hybridization", Hum Genet. Vol. 100, 271–278, 1997.

[3] J. Graham and J. Piper, "Automatic karyotype analysis," Methods in Molecular Biology, vol.29, pp.141-185, 1994.

[4] W. Schwartzkopf, B.L. Evans, and C. Bovik, "Minimum entropy segmentation applied to multi-spectral chromosome images" Proc. IEEE Int. Conf. on Image Processing, Vol. II, pp. 865-868, Thessaloniki, Greece. Oct. 7-10, 2001

[5] W. Schwartzkopf, B.L. Evans, and C. Bovik, "Entropy estimation for segmentation of multi-spectral chromosome images", IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 234 238, Santa Fa, NM. April 7-9, 2002

[6] P. Mousavi, R. Ward, S. Fels, M. Sameti, P. Lansdrop, "Feature Analysis and centromere segmentation of human chromosome images using an iterative fuzzy algorithm", IEEE Trans. on Biomed. Eng., Vol.49, No.4, April 2002, pp. 363-371

[7] M. P. Sampat, K. Castleman, and C. Bovik, "Pixel-by-Pixel classification of M-FISH Images", 2nd Joint Conf. of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, Houston, Oct. 23-26, 2002.

[8] E. Decheve, "Heuristic knowledge based chromosome classification", Proc. Of the 1st national Conf. INFORMATION'94, Sofia, 1994, pp. 85-89

[9] E. Decveva, "Object-Oriented approach to the design of CAIS- an intelligent system for chromosome analysis and classification", Mathematics and Education in Mathematics, Proc. of the 21st spring conf. of union of Bulgarian mathematicians, Sofia, 1992, pp. 226-230

[10] E. Decheva, "CAIS- An intelligent system for chromosome analysis and classification", int. workshop "Artificial intelligent and Humanities", Sozopol, Sept. 1996, pp. 49-54.

-driven partially occluded object segmentation
cessing, Vol.46, No. 10,

pp. 2841-2847,

[12] J. C. Goswami, and A. K. Chan. Fundamentals of wavelets thory, algorithms and applications. John Wiley & Sons, Inc. New York: 1999

[13] R.O., Duda, and P.E. Hart.. Pattern Classification and Scene Analysis. New York: Wiley, 1973.

[14] R. Schulkoff. Pattern Recognition: Statistical, Structural, and Neural Approaches. John Wiley and Sons, NY, 1992.

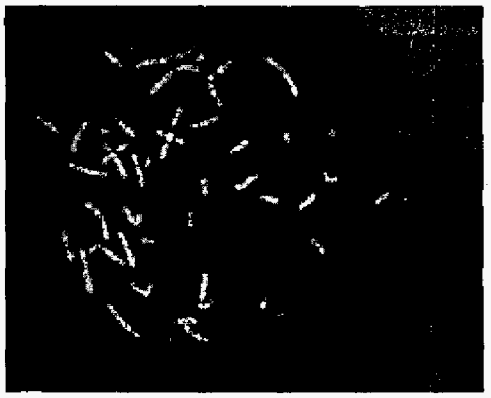
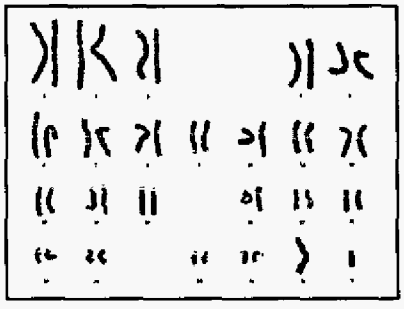


Fig.1 A typical chromosome image

Fig. 2 (a) A chromosome spread image



(b) Karyotype of image (a)

Fig. 3 Foreground Extraction for the image of Fig. 1

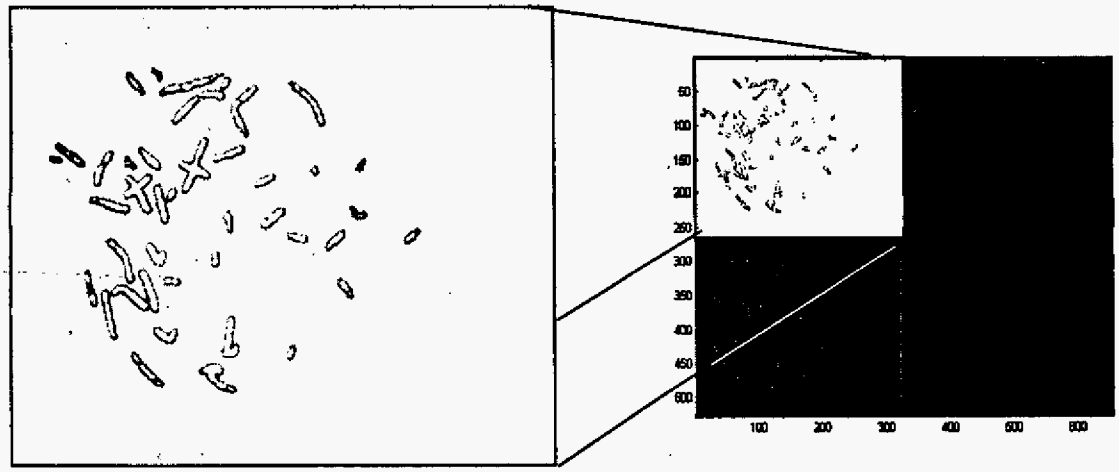


Fig.4 The Approximation image

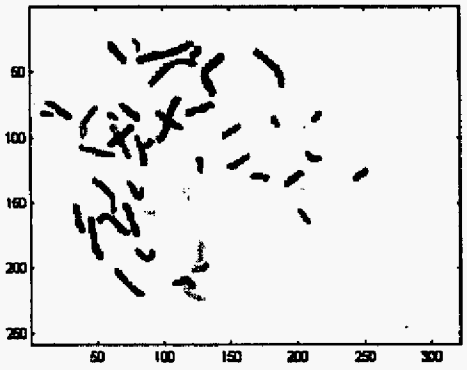


Fig. 5. The resultant image from the pixel by pixel classification of size 325×260 pixels

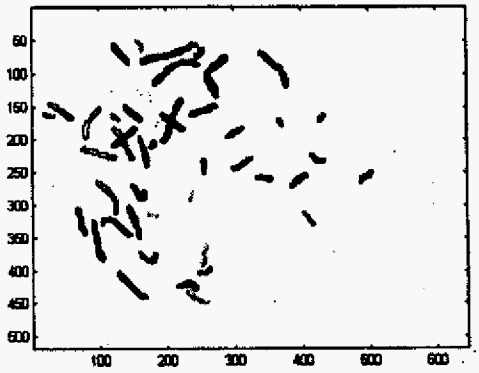
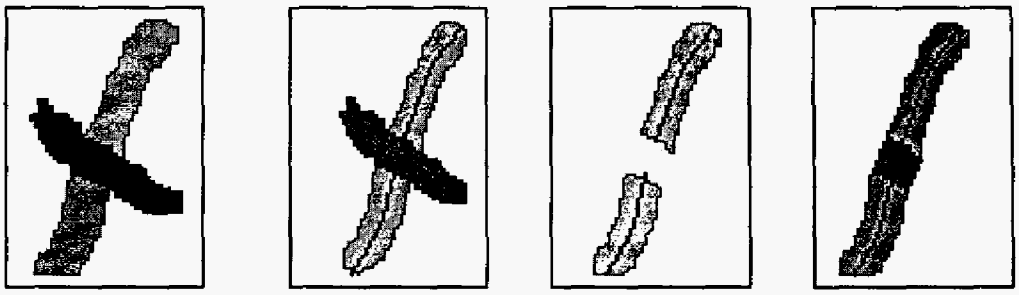
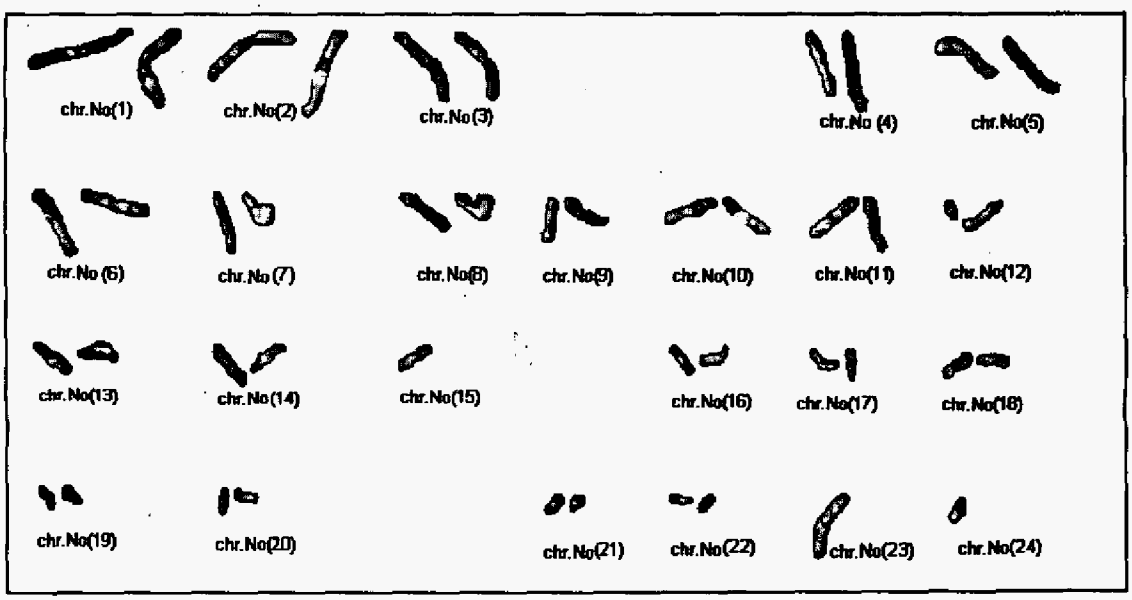


Fig.6. The resized image



(a) (b) (c) (d)

Fig.7 Separation of overlapped chromosomes
(a) Overlapped chromosomes (b) Medial axis
(c) Separated upper chromosome and grow it (d) Connect the medial axis



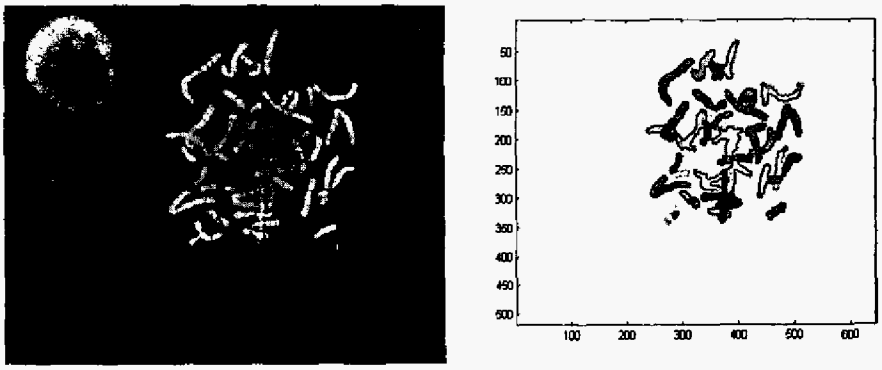


Figure 9. Typical Chromosome Image and its resultant classified version