

Overlapped Chromosome Segmentation and Separation of Touching Chromosome for Automated Chromosome Classification

Nirmala Madian , Member IEEE and K.B.Jayanthi, Senior Member IEEE

Abstract- Karyotyping is the process of arranging pairs of the chromosomes in an order to examine chromosomal abnormalities which are the causes for Birth Defects. The overlapping and touching chromosomes of input metaphase spread images should be segmented and disentangled for this process. This paper proposes an algorithm to separate the touching chromosomes and the segmentation of overlapping chromosomes from G-Band metaspread images. Segmentation is performed by obtaining the interesting (concave) points on the image contour and constructing proper hypotheses for segmentation and separation.

Index Terms- Birth defect, genetic disorder, karyotyping, automated chromosome classification, concave points, hypotheses

I. INTRODUCTION

Birth defect is a global problem. Birth defects include abnormalities in new born baby's structure, function or body metabolism which lead to physical and mental disabilities and can be fatal sometimes. One of the main causes of this is chromosomal disorders.

Chromosomes are composed of supercoiled DNA and its associated proteins. A chromosome is identified with 24 classes where 22 pairs of chromosomes are autosomes. Each autosome is assigned a number from 1-22 and the 23rd pair is the sex chromosome referred as X or Y. Karyotyping analysis [1] is an important screening and diagnostic procedure for detecting several genetic diseases or chromosomal anomalies. A chromosome anomaly reflects a typical number of chromosomes or a structural abnormality in one or more chromosomes. The numerical anomaly is a variation in number of chromosome. Structural anomaly is the breakage and loss of a portion of chromatid arm or a reunion of the arm at different location on the same chromosome or on a different chromosome. Chromosome anomalies usually occur when there is an error in cell division following meiosis or mitosis operation.

*K.B.Jayanthi, Department of ECE, K.S.Rangasamy College of Technology, Tiruchengode 637 215, India.
Email: jayanthikk@gmail.com. Ph.no: 9443471811*

*M.Nirmala, Department of ECE, K.S.Rangasamy College of Technology, Tiruchengode 637 215, India.
Email: nirmalaooty@gmail.com . Ph.no: 8754854895*

The chromosome analysis is performed on metaspread images as the banding patterns are clearly visible in these images. The metaspread chromosomes may be overlapping or touching with each other. They may be bent or inclined to any angle. All these problems should be addressed when the metaspread chromosomes are analyzed.

There are many methods developed to segment the overlapping and touching chromosomes. Some of the methods are global threshold by OTSU method [2] [3], global threshold with a rethresholding scheme [4] [5], K-means clustering [6] and watershed segmentation [7]. The problem with thresholding and clustering approaches are that small chromosomes and their terminal parts have lighter intensities and are tough to segment. Watershed segmentation is not preferred as it is affected by noise and has contrast variability.

The proposed work is a fully automated method to segment the overlapping and touching chromosomes of G-band metaspread images. To segment the overlapping and touching chromosomes, the image contour is identified. Interesting points (concave points) are calculated by curvature function and plotted on the image and possible hypothesis is constructed for the separation process.

II. SEGMENTATION OF TOUCHING AND OVERLAPPING CHROMOSOMES

The segmentation of touching and overlapping chromosomes is an important issue in image analysis. The shape contour of the image provides information for the segmentation process. This shape contour is obtained by marking the interesting points [13] on the image contours. The interesting points are the middle points of constant curvature curve which help in segmenting the chromosomes by considering the concave points on the contours [14] and convex points if necessary. These concave points are plotted on the image. A polygonal approximation of the shape [15] contour is obtained. Further, the separation lines are drawn with the help of the concave points. Overlapping and touching chromosomes are separated.

Detection of chromosome contour:

The shape is a powerful visual cue for recognizing objects in an image. Contour detection is useful in determining the

shape of an image which further helps in separating the joined images. For obtaining the contour, an image should be perfectly binarized so that the foreground is separated from the background. The contour is plotted which shows the entire shape of the image irrespective of the overlapping or touching chromosomes. The contours are obtained by specifying the coordinates on the images given in Eq (2).

Obtaining curvature function:

The curvature function of a curve is defined as the rate of change of the slope of the curve with respect to its length[10]. The curvature function is related to two concepts called concavity and convexity. Concavity (convexity) relates to unconstrained maximum (minimum) problems. To obtain the concave (convex) points the curvature function[6] is used. The following formulation helps to determine the curvature graph with interesting points.

Consider a S link chain with sequence of points as shown in Fig 1. A line by line scanning of each points on the chain is performed. This scanning helps in connecting the end points of the sequence.

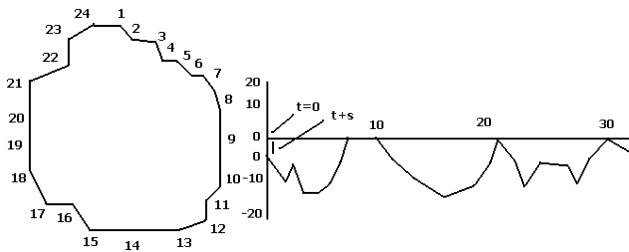


Fig (1) Closed loop chain with curvature graph

The curvature function of the chain is obtained by measuring the angular difference between the successive chain nodes. The initial scan process starts at ‘t’ position which is slightly deviated from zero. This deviation gives the length t+s and helps the curvature to be constant through all non-zero points. The scanning of closed n link chain with pixel points and the coordinates are traced [13] and given as

$$A^n = C_{i=1}^n a_i = a_1 a_2 \dots a_n \dots\dots\dots(1)$$

Where C is the pixel point and

a_i is the ith link in the image.

Let L_j^s be the sub chain link and it is considered as the termination node to which the link a_j should be directed. The

L_j^s is given as

$$L_j^s = C_{i=j-s+1}^j a_i \quad j = 1,2,3,\dots,n$$

X and Y are considered as the image coordinates and X and Y components of L_j^s are calculated as

$$X_j^s = \sum_{i=j-s+1}^j a_i x \quad \text{and} \quad Y_j^s = \sum_{i=j-s+1}^j a_i y \dots\dots\dots(2)$$

where $a_i x$ and $a_i y$ are X and y component and

$$a_i x, a_i y \in \{-1,0,1\}$$

These X and Y components help in determining the length of the chromosome by a simple formulation given as

$$l_j^s = \sqrt{(X_j^s)^2 + (Y_j^s)^2} \dots\dots\dots(3)$$

The length alone is not sufficient for performing curvature function. Angle of inclination is also required for obtaining the curvature function. So the angle for x-axis is given by

$$\begin{aligned} \theta_j^s &= \tan^{-1} Y_j^s / X_j^s \quad \text{if } |X_j^s| \geq |Y_j^s| \\ &= \cot^{-1} Y_j^s / X_j^s \quad \text{if } |X_j^s| < |Y_j^s| \dots\dots\dots(4) \end{aligned}$$

For further smoothing of the curvature function an incremental curvature δ_j^s is defined. The incremental curvature helps in removing the unnecessary points. This curvature is two times the mean over two adjacent angular difference and given as

$$\begin{aligned} \delta_j^s &= 2 \left[\frac{(\theta_{j+1}^s - \theta_j^s) + (\theta_j^s - \theta_{j-1}^s)}{2} \right] \\ &= \theta_{j+1}^s - \theta_{j-1}^s \dots\dots\dots(5) \end{aligned}$$

The increase in s value helps in better smoothing of the curve. The obtained curvature graph is shown in Fig 3. Further analysis of the hypotheses is done only by the concave points. The convex points do not take part in the analysis as the connectivity of the pixel depends only on the concave points

III. CHROMOSOME SEPARATION

Determination of concave property and separation lines:

The concave property of a region is used to find the possible real concave points along a contour. The concavity of a point depends mainly on its neighboring points. The property is explained by the current and previous values. Let the coordinates of the image be X and Y. $p_c(x_c, y_c)$ represent the connectivity of the current point and $p_{pre}(x_{pre}, y_{pre})$ and $p_{next}(x_{next}, y_{next})$ represent the previous and the next points of p_c respectively. Then, the concavity of the contour p_c is calculated as follows[16]:

$$a(p_{pre}, p_c) = \tan^{-1}((y_{pre} - y_c) / (x_{pre} - x_c))$$

$$a(p_{next}, p_c) = \tan^{-1}((y_{next} - y_c) / (x_{next} - x_c))$$

The above equations help to obtain the concave points on the image as shown in Fig 4. Increasing the number of concave points results in increasing the accuracy of separation and segmentation.

The separation lines are obtained by considering all possible pairs of high concave points as shown in Fig 6. The initial set of possible separation lines is obtained by taking the possible combinations of points, excluding the reflexive pairs. The separation lines are traced by pixels with non zero values and zero values as background. The initial point of the image should be considered by specifying the row and column coordinates and the initial search direction should be specified for the next connected pixel. The other parameters to be considered are the connectivity and the direction to trace the connected pixel for obtaining the separation lines. The hypothesis is determined from these separation lines. The hypothesis is based on geometric information which is obtained as the interesting points on the image. This is further used to separate the overlapping and touching portion in the image.

IV. HYPOTHESIS DETERMINATION OF OVERLAPPING AND TOUCHING IMAGES

The hypothesis is obtained by determining the interesting points on the chromosomes. The entire chromosome image is separated into subset of pixels. A chromosome separation problem is generally considered as the partition problem. A minset partition is performed for the separation. Minset is nothing but the interconnection of some subsets and complement of others. Finding the minset partition helps in reducing the problems involved in separation hypothesis [10]. One minset is available for each object and the union of minset is not considered for the hypothesis. The separation lines by different combination pairs are analyzed for obtaining the partition. For reducing the number of partitions, the spatial distance between the vertices of separation lines of touching objects should be small, compared to the distance between the vertices along the shape contour. The pairs are obtained by a simple distance formulation given as

$$D = \sqrt{(bx - ex)^2 + (by - ey)^2}$$

where (bx, by) are the branch points and (ex, ey) end points of the image. In some cases, where the obtained partitions are not obtained from minset then the partition subset are obtained by union of several minsets. This hypothesis is applied only for the touching chromosome as shown in Fig 7.

When an overlapping chromosome is taken for the analysis, the possible partition and possible minset of the object must be considered. The distance constraint does not help the reduction of possible pairs. To construct the hypotheses for the separation of overlapping objects[10], an overlap minset should be located. An overlap zone is created as shown in Fig 6a with two parallel separation lines. The hypothesis is obtained by considering the subsets that have more than one element. A pair of parallel separation lines fall in overlap zone if the distance between the separation lines meets the given threshold value. The lines are known as permitted separation lines and helps in segmenting the overlapping portion as shown in Fig 6.

V. RESULTS AND DISCUSSION

In the proposed work, the chromosome images are analyzed for automated image classification. This classification helps in detecting various chromosomal anomalies. For performing

segmentation, initial step is the contour determination. The contour helps in determining the shape of the image irrespective of the overlapping or touching portion. Fig 2 shows the output of the obtained contour as explained in section III .



Fig (2) Contour Image

The curvature function helps in determining the concave (convex) points. The interesting points are considered as the extreme points on the curvature. Fig 3 presents the output of the curvature function with interesting points and Fig 4 shows the output of concave (green triangle) and convex (red dots) points plotted on the image as discussed in section III and IV

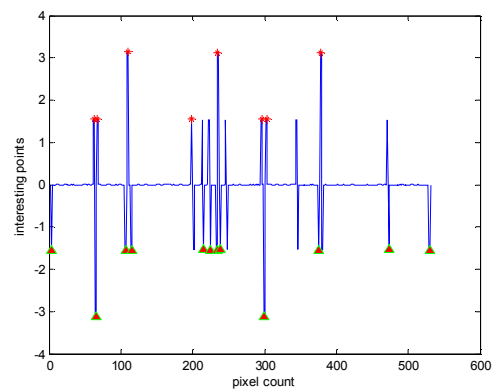


Fig (3) Curvature Graph

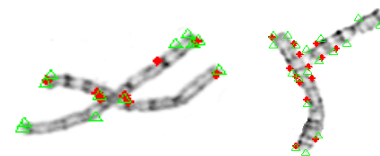


Fig (4) Interesting Points on the Image

The separation lines are obtained using the concave points. The connectivity and the direction are traced by the connected pixel for obtaining the separation lines as discussed in section IV. Fig 5 shows the separation lines on the image where the reflexive points and the pairs with identical elements are removed.



Fig (5) Separation Lines on the Image

The hypothesis is determined for segmenting the overlapping and touching chromosomes. For obtaining the original shape of the chromosome image the bounding polygon is constructed covering the corners of the image. The outputs related for overlapping chromosome is given in Fig 6 and Fig 6(a) gives

the separation zone region for overlapped chromosome, which further helps in separating the overlapping chromosomes. Fig 7 shows the segmentation output for the touching chromosome.



Fig (6) Segmentation Output for Overlapped Images

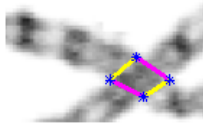


Fig (6a) Separation Zone for Overlapped Images

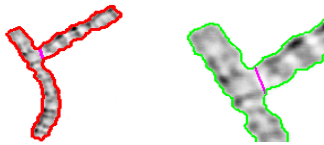


Fig (7) Segmentation Output for Touching Images

The segmentation output of the overlapping image is shown in Fig (6). The yellow color line in the image indicates the separation line across overlap part on other image. Fig (7) shows the segmentation of touching images. The separation lines are indicated on the images which further helps in separation of two touching images.

The above outputs discuss about the segmentation of the overlapping and touching images. For the separation, the hypothesis performs well on the touching chromosome. The output of the separated chromosome image is shown in Fig (8) and the extended view of the original and the separated images are shown.

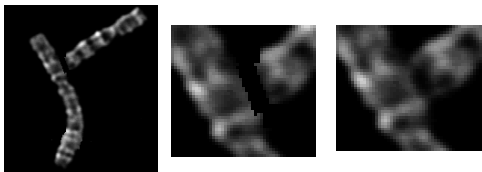


Fig (8) Separation of Touching Images

VI. CONCLUSION

The proposed algorithm helps in the automated chromosome classification, taking care of the first step of segmentation of overlapping image and disentangling of touching image. The proposed work achieves a good result compared to the other methods. This work helps in the automated segmentation which minimizes the human intervention for the classification of chromosome images.

ACKNOWLEDGEMENT

The authors wish to thank Dr.Suresh, Director, Birth Registry of India for providing the G-band chromosome images for the work.

REFERENCES

- [1] J.Graham and J.Piper, Automatic karyotype analysis, Totowa, NJ: Humana 1994, pp. 144-185.
- [2] L. Ji, "Intelligent splitting in the chromosome domain," *Pattern Recognit.*, vol. 22, no. 5, pp. 519-532, 1989.
- [3] L. V. Guimarães, A. Schuck, Jr., and A. Elbern, "Chromosome classification for karyotype composing applying shape representation on wavelet packet transform," in *Proc. 25th Silver Anniversary Int. Conf. IEEE Eng. Med. Biol. Soc.*, Cancun, Mexico, Sep. 2003, pp. 941-943.
- [4] L. Ji, "Fully automatic chromosome segmentation," *Cytometry*, vol. 17, pp. 196-208, 1994.
- [5] R. J. Stanley, J. M. Keller, P. Gader, and C. W. Caldwell, "Data-driven homologue matching for chromosome identification," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 451-462, Jun. 1998.
- [6] B. Lerner, "Toward a completely automatic neural-network-based human chromosome analysis," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 28, no. 4, pp. 544-552, Aug. 1998.
- [7] P. S. Karvelis, D. I. Fotiadis, M. Syrrou, and I. Georgiou, "Segmentation of chromosome images based on a recursive watershed transform," presented at the 3rd Eur. Med. Biol. Eng. Conf., EMEBC 2005, Prague, Czech Republic, Nov.
- [8] B. Lerner, H. Guterman, and I. Dinstein, "A classification-driven partially occluded object segmentation (CPOOS) method with application to chromosome analysis," *IEEE Trans. Signal Process.*, vol. 46, no. 10, pp. 2841-2847, Oct. 1998.
- [9] X. Shunren, X. Weidong, and S. Yutang, "Two intelligent algorithms applied to automatic chromosome incision," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, '03. (ICASSP '03), Apr., pp. 697-700.
- [10] G. Agam and I. Dinstein, "Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1212-1222, Nov. 1997.
- [11] C. Urdales Garcia, A. Bandera Rubio, F. Arrebola Perez and F. Sandoval Hernandez, "A curvature-based multiresolution automatic karyotyping system", *Mach. Vis. Appl.*, vol. 14, pp. 145-156, 2003
- [12] G. C. Charters and J. Graham, "Disentangling chromosome overlaps by combining trainable shape models with classification evidence," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 2080-2085, Aug. 2002.
- [13] H. Freeman and L.S. Davis, "A corner finding algorithm for chain code curves", *IEEE Trans. Computer*, vol. 26, no. 3, pp. 297-303, Mar. 1977
- [14] T. Pavlidis, "Algorithms for shape analysis of contours and waveform", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 2, no. 4, pp. 301-312, 1980
- [15] A. Pikaz and I. Dinstein, "Optical polygonal approximation of digital curves", *Pattern Recognition*, vol. 28, no. 3, pp. 373-379, 1995
- [16] Xiangzhi Bai, Changming Sun, Fugen Zhou, "Splitting touching cells based on concave points and ellipse fitting", *Pattern Recognition* 42(2009)2434-2446