

## Segmentation of Overlapped and Touching Human Chromosome images

Sri Balaji.V<sup>1</sup>, Pragasam.G<sup>2</sup>, Sowmiya.R<sup>3</sup>, Vijayalakshmi.H<sup>4</sup>,  
Nirmala Madian<sup>5</sup>

<sup>1,2,3,4</sup>B.E. Electronics and Communication, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, India.

<sup>5</sup>Assistant Professor Electronics and Communication, K.S.Rangasamy College of Technology, Tiruchengode, Namakkal, India.

**Abstract:** Chromosomes are essential genomic information carriers. The identification of chromosome abnormalities is an essential part of diagnosis and treatment of genetic disorders such as chromosomal syndromes and many types of cancer. Currently available cytogenetic imaging software is designed to classify only normal chromosomes. The automation of chromosome analysis is involving segmentation of chromosomes and classification into 24 groups. Segmentation of the overlapped chromosomes is a major step toward the realization of homolog classification. Resolving chromosome overlaps is an unsolved problem in automated chromosome analysis. Current systems for automatic chromosome classification are mostly interactive and require human intervention. In this paper, an automatic procedure is proposed to obtain the separated chromosomes. The separations of overlapped and touching chromosomes are obtained by finding the intersecting (concave and convex) points with the help of Novel algorithm. The intersecting points are located on contour of the image and then the curvature function is used to find out concave points. Then the possible separation lines are plotted by using all concave points and finally construct the hypotheses for possible separation lines between concave points. The segmentation is carried out by means of a curvature function scheme, which proved to be successful.

**Keywords:** automatic chromosome classification, diagnosis, genetic disorders, hypothesis, interesting points.

### I. Introduction

Chromosomes are microscopic structures that specify an individual's genetic information. The chromosomes are composed of DNA and various proteins. The DNA contains the actual genetic code of an individual, and the proteins protect the DNA and allow it to duplicate properly when the cell divides.

Chromosome analysis is performed by dividing cells in their metaphase stage. In that stage, the chromosomes may be stained in such a way that creates a typical band-pattern on them. A chromosome in the metaphase stage is constructed from two identical arms (chromatides), which are joined together on a common point (centromere). The centromere actually divides each chromatid into two arms. Chromosomes are the essential genomic information carriers. The chromosomes consist of 22 pairs of autosomes (1–22) and two sex chromosomes (X and Y). Thus, there are 24 classes of chromosomes [1].

A small proportion of the population has cells each one of which has fewer or more than 46 chromosomes, implying a substantial deficit or excess of genetic material. Others have one or more structurally abnormal chromosomes [2], in which a section of a chromosome has moved from its usual to another position.

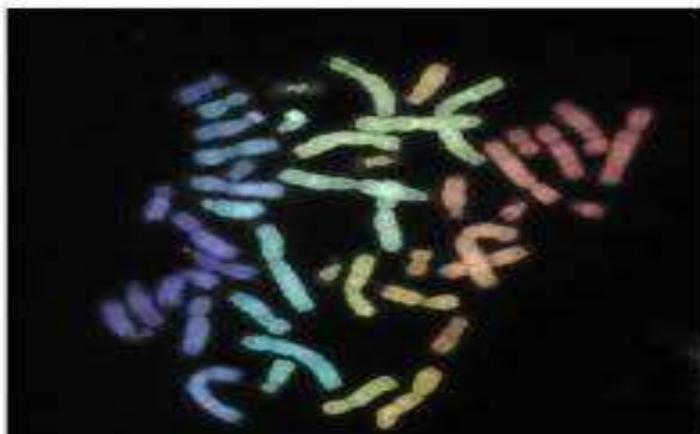


Fig 1: The Q-band metaphase chromosome images

Chromosome analysis is performed on dividing cells in their metaphase stage [2]. The Q-band metaphase chromosome images are shown in fig 1. The chromosomes in a metaphase image may be bent, their arms may be joined (along points other than the centromere), their bands may be spread, and they may touch or partially overlap each other.

Thus, all these problems should be taken into consideration when processing the metaphase image. Current systems for automatic chromosome classification are mostly interactive and require human intervention for correct separation between touching and overlapping chromosomes [3]. The separation is performed by posing a series of hypotheses that fit certain constraints [4], [5].

## II. Related Works

1) Segmentation methods try to separate touching chromosomes by classifying the chromosome pixels and the background pixels into two different segments. Since segmentation methods are general methods that do not depend on the shape of objects, their link to the specific problem of chromosome separation is weak, and, therefore, they tend to fail in cases of incomplete information. A case of incomplete information occurs; when there is no separating path between chromosomes [5], [6].

2) Heuristic search edge-linking methods try to separate between touching chromosomes by searching for a minimal-cost connected path between the chromosomes. The link lost connection of such methods to the specific problem of chromosome separation leads to inferior results in cases of incomplete information, such as in the case where a separating path does not exist [7].

3) The method for separating the touching chromosomes by using shape decomposition based on fuzzy subset theory [8] is reported as giving inferior results relative to other general methods. The disadvantage of this method is that it yields erroneous decompositions of single chromosomes in cases of bent chromosomes.

## III. Proposed Algorithm

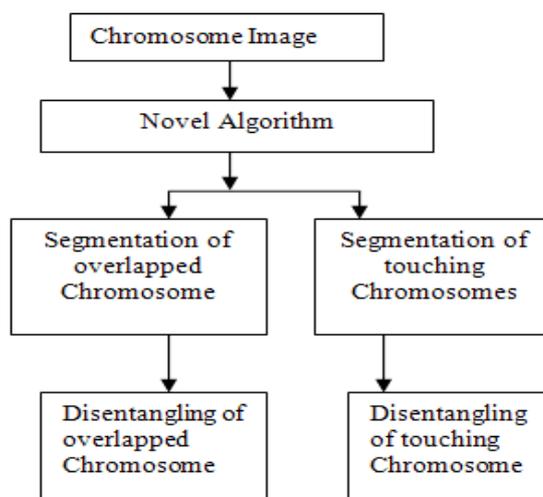


Fig 2: Block diagram

### 3.1 Algorithm Steps

- Get the input RGB chromosome overlapped image.
- RGB chromosome image to gray scale output.
- Applying thresholding to input image.
- Obtain binary and contour map output.
- Obtain the curvature graph for the input image.
- Identify the interesting point of the input image.
- Obtain canny edge detection output.
- Obtain segmentation output.

The segmentation output is obtained by using Novel algorithm. The Novel algorithm is explained in section 3.2.

### 3.2 Novel Algorithm

A novel algorithm based on watershed and concavities is proposed to segment the chromosome images. First, the distance and watershed transform is used to the binary image of chromosome images. Secondly, the watershed post-processing of over segmentation is dealt with by utilizing concavity features of related shapes. Thirdly, the candidate splitting lines of touching images is found by matching the concavities to the un-segmentations left. Finally, the supplementary criterions are applied, such as the shortest distance, the opposite orientation, the splitting path orientation, etc., to determine whether a candidate splitting line can be accepted or not.

### 3.3 Segmentation And Disentangling

The segmentation technique is widely used by the radiologists to segment the input medical image into meaningful region. Segmentation algorithm is widely preferred because of its additional flexibility which allows multiple classes with varying degrees of membership.

This function calculates and plots the intensity values along a line segment or a multiline path in an image. The user defines the line segment (or segments) by specifying their coordinates as input arguments. The user can define the line segments using a mouse. By default, improvise uses nearest-neighbour interpolation, but you can specify a different method. Improvise works best with true color images.

## IV. Preliminary Processing

The preliminary processing of the metaphase image is performed by increasing the effective sampling resolution [9]. This may be achieved by sampling each metaphase image in some overlapping parts. Fig. 3 is the sample overlapped and touching image.



Fig 3: Sample overlapping and Sample touching images

### 4.1 Obtain the Chromosome Contours

The relevant information of a shape is contained in its contour. For separation purposes, only the shape contours are required. The use of shape contours actually represents a successful information compaction that reduces the amount of processed data and stresses the main features of the objects. Chromosome contours are determined by using a common contour-following method to track the contours of the connected segments by row and columns [10]. Fig. 4 shows the binary approximation of the overlapped image.

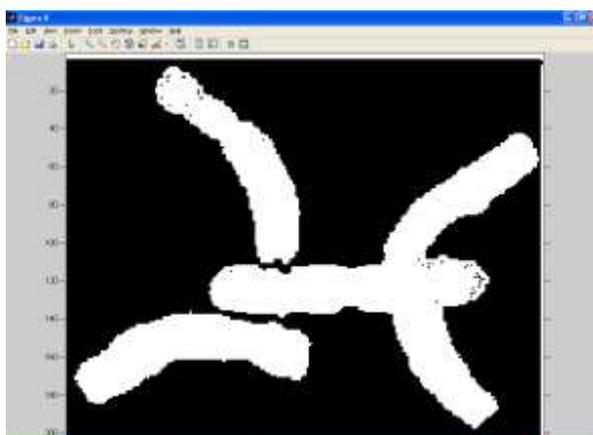


Fig 4: Binary approximation of the given overlapped image

### 4.2 Obtain the discrete curvature function

The curvature function of a curve is defined as the rate of change of the curve slope with respect to its length. The radius of a circle is tangent to the concave side of the curve, and has the same curvature as the curve at the tangential point. The curvature function is a derivative of the contour's slope function. To reduce the computational requirements of the K-slope evaluation, an extended novel method is used. Using the extended view, it is possible to determine the K-slope at a given pixel directly from a look-up table where the extended

novel algorithm serves as the entry point. This method used for detecting and locating concave or convex point. The detection of a corner is a function of the magnitude of the discontinuity. The concave and convex points of the object are detected by using novel algorithm method. Fig 5 shows the curvature graph for the given image.

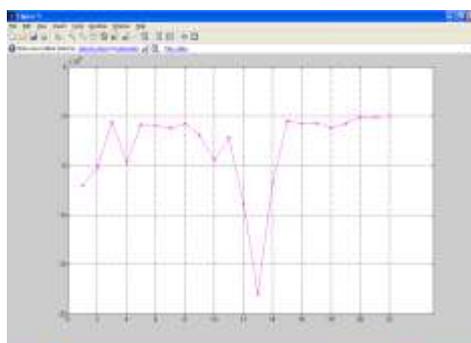


Fig 5: Curvature graph

#### 4.3 Obtaining the Interesting Points

The interesting points are detected as extreme points of the curvature function, or as middle points of constant curvature curve segments. The required filter should combine a low-pass filter for high frequency noise removal, and a high-pass filter for curvature peaks enhancement [11].

Therefore, it results in a band-pass filter. The filtered curvature function is segmented, and extreme points are detected within each segment. From that list, only concave points with a curvature measure above a certain threshold are selected. The interesting points (concave and convex points) are located on the chromosomes. Fig. 6 shows the location of concave points on the overlap chromosome.

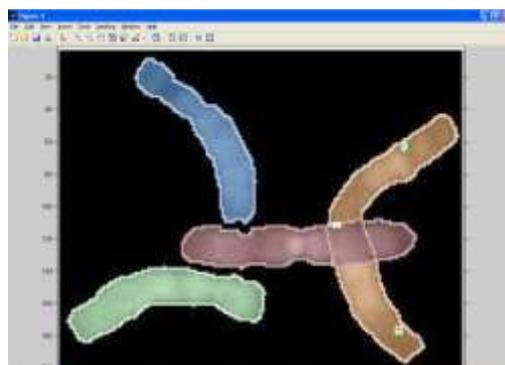


Fig 6: Concave points on the overlap chromosome

### V. Chromosome Separation

When each chromosome is observed as a separate subset of pixels, it is possible to view the chromosome separation easily. The minset is obtained by taking the intersection between some subsets and the complement of the other subsets. Each minset of the chromosomes represents either a part unique to one chromosome or a part common to more than one chromosome. The minset normal form representation of the chromosomes can be constructed by a union of such parts. The required partition is obtained by splitting the object along the lines that connect high concave points.

#### 5.1 Hypothesis for Overlapping Objects

The constraint requires that the spatial distance between the vertices of a separation line of touching objects [12] should be small relative to the distance between the vertices along the shape's contour.

To construct a hypothesis for the separation of overlapping objects, an overlap minset should be located. The distance between the separation lines should be above a given threshold.

#### 5.2 Hypothesis verification

The verification of hypotheses is based on an evaluation of the fit of the obtained parts to the prototype shapes of the chromosomes [13]. In general, the shape of a chromosome is not fixed. A chromosome may have various heights and widths. The evaluation of the fit of a chromosome to its prototype shape is done by fitting a

bounding polygon to the chromosome [13], [14]. The bounding polygon is constructed by locating the shape's minimal bounding box, contracting it in the middle, and rounding its corners. In order to retain the uniqueness of the bounding polygon as a chromosome prototype, constraints are posed on possible locations of its vertices. The minimal bounding box of a shape is found by computing the principal component transformation of that shape. Fig. 7 shows the touching region in the chromosome, fig 8 shows the separation of first chromosome from second chromosome and the second chromosome from first chromosome, fig 9 shows canny edge detection output and fig. 10 a, b, c shows the segmentation of overlapped and touching chromosome.

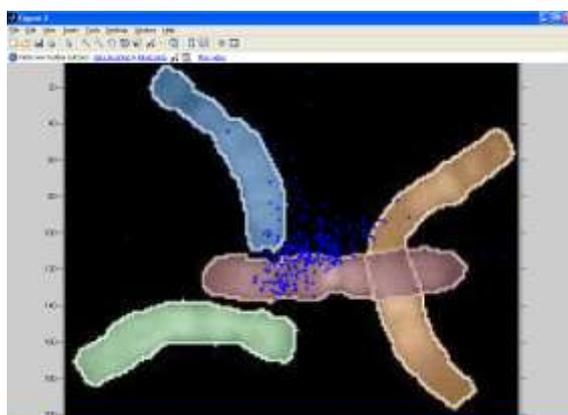


Fig 7: Touching region in the chromosome

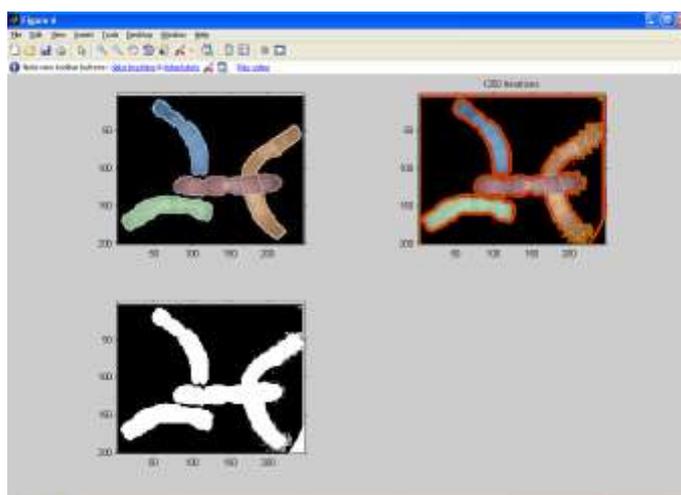


Fig 8: The separation of first chromosome from second chromosome and the second chromosome from first chromosome



Fig 9: Canny edge detection output.

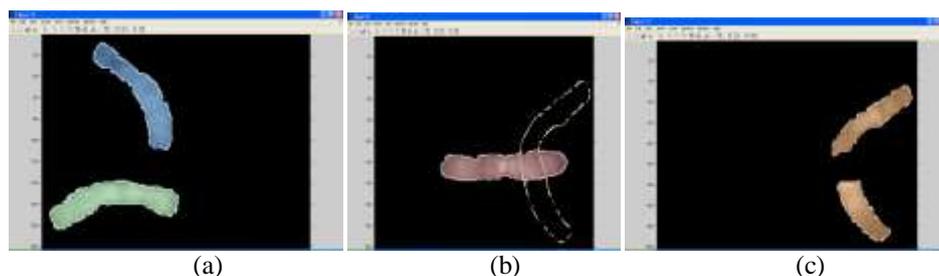


Fig10: a, b, c Shows the segmentation of overlapped and touching chromosome

## VI. Conclusion

Karyotype analysis is a widespread procedure in cytogenetics to assess the possible presence of genetics defects. The procedure is lengthy and repetitive, so that an automatic analysis would greatly help the cytogeneticist routine work. Still, automatic segmentation and full disentangling of chromosomes are open issues.

The problem of extracting individual objects from overlapping configurations is one that occurs in a number of applications in machine vision. It is a difficult problem in general and can be thought of in two parts: 1) identifying overlapping objects in the first place and 2) identifying the individual components. In the case of chromosomes, it is usually fairly easy to identify an overlap on straightforward shape criteria.

This proposed method for Automatic Chromosome segmentation of touching and overlapping images are more accurate than other previous methods. This method of segmentation is applicable for large number of bends present in the chromosome structures and more number of clusters by optimized the location of each separation point, yielding the minimal possible distance between the smoothed approximation and the original curve. The experiment proved that this algorithm can segment the overlapping and touching chromosomes successfully.

## REFERENCES

- [1] A. Carothers and J. Piper, "Computer-Aided Classification of Human Chromosomes: A Review," *Statistics and Computing*, vol. 4, no. 3, pp. 161-171, 1994.
- [2] W. C. Schwartzkopf, A. C. Bovik, and B. L. Evans, "Maximum-likelihood techniques for joint segmentation-classification of multispectral chromosome images," *IEEE Trans. Med. Imag.*, vol. 24, no. 12, pp. 1593-1610, Dec. 2005.
- [4] P. Mousavi, R. K. Ward, and P. M. Lansdorp, "Feature analysis and classification of chromosome 16 homologs using fluorescence microscopy image," *IEEE Can. J. Elect. Comput. Eng.*, vol. 23, no. 4, pp. 95-98, 1999.
- [5] M. Moradi and S. K. Staredhan, "New features for automatic classification of human chromosomes: A feasibility study," *Pattern Recognit. Lett.*, vol. 27, pp. 19-28, 2006.
- [6] K. Z. Mao, P. Zhao, P. Tan, "Supervised learning-based cell image segmentation for P53 immunohistochemistry," *IEEE Transactions on Biomedical Engineering* 53 (6) (2006) 1153-1163.
- [7] Carsten Garnica, Frank Boochs, Marek Twardochlib, "A new approach to edge-preserving smoothing for edge extraction and Image segmentation," *IAPRS*, Vol. XXXIII, Amsterdam, 2000.
- [8] T. Law, K. Yamada, D. Shibata, T. Nakamura, L. He, and H. Itoh, "Edge extraction using fuzzy reasoning," in *Soft Computing for Image Processing*, S. K. Pal, Ed. New York: Physica-Verlag, 2000.
- [9] K. Saracoglu, J. Brown, L. Kearney, S. Uhrig, J. Azofeifa, C. Fauth, M. Speicher, and R. Eils, "New concepts to improve resolution and sensitivity of molecular cytogenetic diagnostics by multicolor fluorescence in situ hybridization," *Cytometry*, vol. 44, no. 1, pp. 7 May 2001.
- [10] E. Grisan, A. Pesce, A. Giani, M. Foracchia, and A. Ruggeri, "A new tracking system for the robust extraction of retinal vessel structure," in *Proc. 26th Annu. Int. Conf. IEEE-EMBS*. New York: IEEE Press, 2004, pp. 1620-1623.
- [11] G. Agam and I. Dinstein, "Geometric separation of partially overlapping non-rigid objects applied to automatic chromosome segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 11, pp. 1212-1222, Nov. 1997.
- [12] Xiangzhi Baia, Changming Sun, Fugen Zhou, "Splitting touching cells based on concave points and ellipse fitting," *Pattern Recognition* 42, 2009.
- [13] Kolesnikov, A., Fränti, P.: A fast near-optimal algorithm for approximation of polygonal curves. *Proc. Int. Conf. Pattern Recognition-ICPR'02*. 4 (2002) 335-338.
- [14] Chen, D.Z., Daescu, O.: Space efficient algorithm for polygonal curves in two dimensional space, *Proc. 4th Int. Conf. Computing and Combinatorics*. (1998) 55-64.