Walailak J Sci & Tech 2006; 3(2):181-194.

Segmentation of Overlapping Chromosome Images Using Computational Geometry

Wacharapong SRISANG, Krisanadej JAROENSUTASINEE and Mullica JAROENSUTASINEE

CX-KURUE & Computational Science Graduate Program, School of Science, Walailak University, Nakhon Si Thammarat 80161, Thailand

ABSTRACT

Current systems for automatic chromosome classification are interactive and require human intervention for correct separation between touching and overlapping chromosomes. Special separation methods are required to segregate chromosomes because they are non-rigid objects. This study develops a new technique to separate overlapping chromosomes based on computational geometry. This technique requires the identification of all possible cut points from the contour line of overlapping chromosomes, using Voronoi diagrams and Delaunay triangulations to select the four target cut points and cut overlapping chromosomes into two chromosomes. We test our algorithm on 35 overlapping chromosome images and find that 28 out of 35 overlapping chromosomes images can be separated correctly (i.e. 80.0 %). Three out of the 35 images are separate incorrectly (i.e. 8.6 %) and four out of 35 images are not separable by our algorithm (i.e. 11.4 %).

Keywords: Image segmentation, chromosome analysis, overlapping chromosomes, computational geometry

INTRODUCTION

Human chromosome analysis is an essential task in cytogenetics, especially in prenatal screening and genetic syndrome diagnosis, cancer pathology research and environmentally induced mutagen dosimetry [1]. Cells used for chromosome analysis are taken mostly from amniotic fluid or blood samples. One of the aims of chromosome analysis is the creation of a karyotype, which is a layout of chromosome images organised by decreasing size in pairs [2]. The karyotype is obtained by cutting chromosome images from a photograph of a cell, taken using a microscope, and arranging the chromosomes into their appropriate places on the layout according to their visual classification by a cytotechnician. Karyotyping is a useful tool for detecting deviations from normal cell structure. Abnormal cells can have an excess or deficiency of a chromosome and/or structural defects, like breaks, fragments or translocations (i.e. exchange of genetic material between chromosomes). However, even today chromosome analysis and karyotyping are performed manually in most cytogenetic laboratories in a repetitive, time consuming and therefore, expensive procedure.

Great efforts to develop automatic chromosome classification techniques have been made during the last 25 years. However, all have had limited success and have yielded poor classification results compared with those of a trained cytotechnician [3-5]. Some of the reasons for the relatively poor performances are the inadequate use of expert knowledge and experience, and insufficient ability to make comparisons and/or elimination among chromosomes within the same metaphase. In addition, the systems always require operator interaction to separate touching and/or overlapping chromosomes and to verify the classification results [1,3,4].

Automatic separation of overlapping chromosomes is important for the analysis of prophase and pro-metaphase images, but automatic separation has received relatively little attention compared to other aspects of the chromosome analysis problem such as classification. Methods for automatically segmenting both touching and overlapping chromosomes are to decompose a thresholded object into individual components using geometric evidence by reasoning about shapes [6-8], boundary curvature [9,10], chromosome skeleton [11] and banding pattern [12,13]. In this study, we develop a new technique for separating overlapping chromosome images by using computational geometry, which has been widely used for image analysis [14-16].

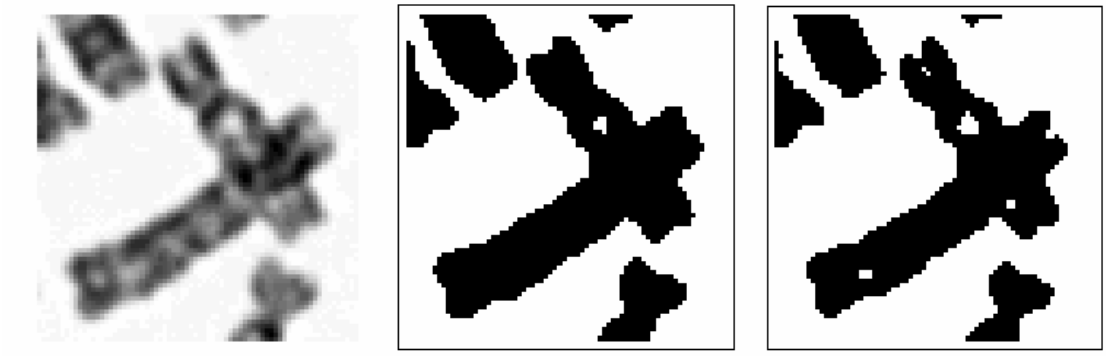
MATERIALS AND METHODS

Thresholding techniques for chromosome images

Thresholding is a well-known technique for chromosome segmentation and the operation of converting a multi-level image into a binary image [17,18]. Each pixel in a binary image value is represented by a single binary digit. In its simplest form, thresholding is a point-based operation that assigns the values of 0 or 1 to each pixel of an image based on a comparison with some global threshold value T as in Eq. (1).

$$f_T(x,y) = \begin{cases} 1, & if\ f(x,y) \geq T \\ 0, & if\ f(x,y) < T \end{cases} \tag{1}$$

Thresholding leads to a significant reduction in data storage and results in binary images that are simpler to analyse. Binary images permit the use of powerful morphological operators for shape and structure-based analysis of image content. A threshold value can be determined based on two selectable minimisation criteria: (1) minimising a mean squared error between the original and the binary image (**Figures 1a,b**) and (2) minimising a weighted sum of group variances, where the groups are formed from the pixels that fall above and below some chosen threshold (**Figure 1c**) [19]. The *Mathematica* Digital Image Processing package is used to generate a histogram of the original image and select threshold values [20].



(a) (b) (c)

Figure 1 Overlapping chromosome images with minimising techniques. (a) an original image of two overlapping chromosomes, (b) threshold image by minimising the mean squared error (T = 218), and (c) threshold image by minimising the weighted sum of group variances (T = 207).

Preprocessing for computational geometry

Further successful information compaction may be achieved by locating interesting points on the shape contours [21]. Interesting points are defined as contour pixels at which an abrupt change occurs. For the purpose of object separation, it is sufficient to consider only high concave points on external contours [22]. These interesting points are possible cut points that are obtained from contour lines by selecting points from a curvature function. The target cut points (i.e. four points used for separating overlapping chromosomes) are cut points around the overlapping area. Voronoi diagrams are used for determining the centre of overlapping area. Delaney triangulation is used for determining the four nearest possible cut points.

Computation of the chromosome outlines

Chromosome contours are obtained from threshold images. They are object pixels that connect to background pixels. These pixels are obtained by forming the convolution of the kernel K in Eq. (2) and Eq. (3) with the threshold image (**Figures 2a,b**). Then contour pixels are obtained from the convolution matrix by Eq. (4) (**Figure 2c**).

$$K = \begin{pmatrix} 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 1 & \frac{1}{4} \\ 0 & \frac{1}{4} & 0 \end{pmatrix} \tag{2}$$

$$f_T(x,y) = \begin{cases} 1, & if \quad 1 < f(x,y) < 2 \\ 0, & if \quad 1 \geq f(x,y) \leq 2 \end{cases} \tag{3}$$

Interpolation of the discrete curvature function

The curvature function γ of a curve *y(x)* is defined as the changing rate of the curve slope φ(*s*) with respect to its length *s*. The exact calculation of the curvature function according to its definition is given by Eq. (4).

$$\gamma \equiv \frac{d\varphi(s)}{ds} \equiv \frac{1}{p} = \frac{\ddot{y}}{(\dot{y}^2 + 1)^{3/2}} \tag{4}$$

Where $\dot{y}$ and $\ddot{y}$ denote the first and second derivatives of y, respectively, and p represents the radius of curvature, which is the radius of a circle that is tangent to the concave side of the curve, and has the same

curvature as the curve at the tangential point. Since the exact calculation of the curvature function requires a second derivative, it is impractical for discrete curves.

The curvature function is a derivative of the contour's slope function, thus the contour's slope function can be evaluated. In order to prevent noise in the slope function caused by fluctuations on the contour, some smoothing must be introduced into the slope evaluation. The *K*-slope at a contour pixel is defined as the slope of a line connecting that pixel with its right neighbour [23]. The selection of a large enough *K* introduces the required smoothing into the slope function evaluation. The *K*-curvature at a contour pixel is defined as the difference between the *K*-slope at that pixel and the *K*-slope of its left neighbour. The values of the *K*-curvature should then be normalised so that a unique representation for each possible angle will be obtained.

(a) (b) (c)

Figure 2 Overlapping chromosome images. (a) threshold image, (b) convolution image and (c) contour image

Detection of the possible cut points

The possible cut points are contour pixels where an abrupt change occurs. The possible cut points are detected as extremum points of the curvature function (**Figure 3**). Before analysing the curvature function, a low-pass filter for high frequency noise removal is used to filter the noise out in order to stress its main features. The filtered curvature function is segmented, and extremum points are detected within each segment. The result of this process is a list of possible cut points, each having a curvature measure.

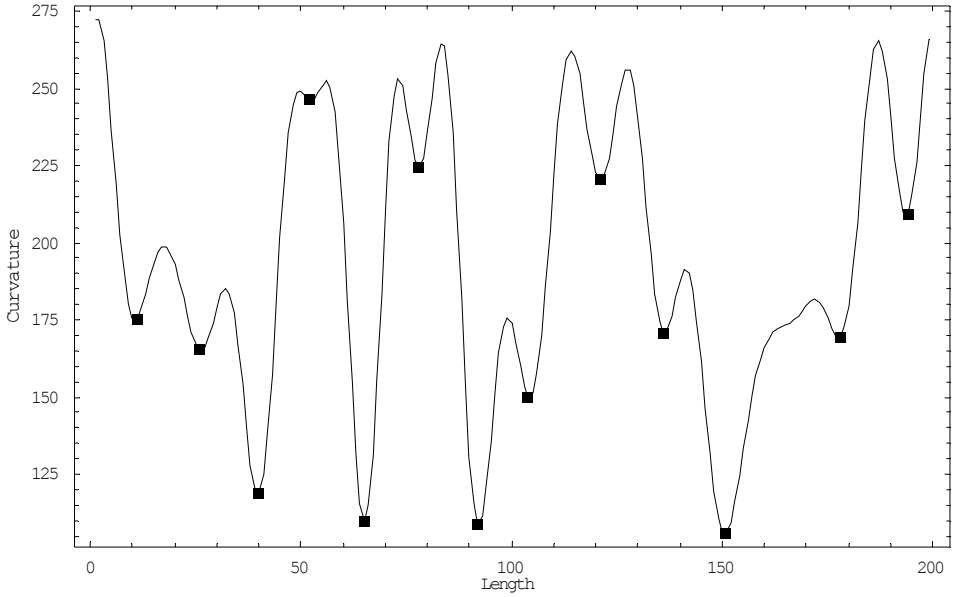
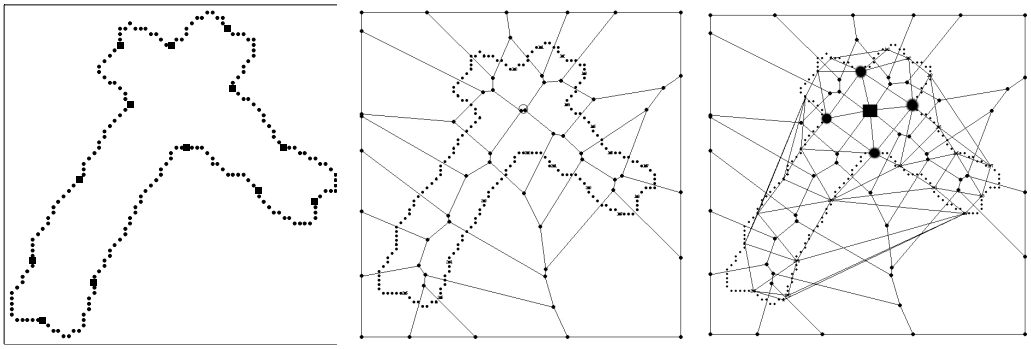


Figure 3 Possible cut points on a filtered curvature graph. ■ represent possible cut points.

Determination of the centre of overlapping area using Voronoi diagram technique

Target cut points of overlapping chromosomes are possible cut points at the crossing area. These points can be observed as points that are nearest around to the centre of the crossing area and the intersection point of two chromosome skeletons. From this property, we can obtain possible cut points by skeletonising the chromosome and finding the points. We use a Voronoi diagram to find the centre of this crossing area. The *Mathematica* Computational Geometry package is used to compute both Voronoi diagrams

and Delaunay triangulations [20]. The bounded Voronoi diagram is used due to reduced computational time (**Figure 4a**). From a Voronoi diagram the centre of the overlapping area can be obtained from the diagram as the points that connect to more than two internal nodes. From our previous study [24], we found that if fewer points are used for computing the Voronoi diagram, a unique answer cannot be found. To avoid this problem, we add some contour points for computations. These points are points at a middle position between possible cut points. For two points, we use the middle point for computation in the next step. The algorithm sometimes gives more than two points. Extra points appear which would add more contour points for computation. These points are all the nodes at the end of the skeleton not at overlapping areas. They can be eliminated by cutting the branch end of the chromosome skeleton by using the same criteria for selecting the point again. The usable cut points can be obtained by using Delaunay Triangulation to calculate a set of possible cut points joined to the crossing centre point (**Figure 4b**).



(a) (b) (c)

Figure 4 Overlapping chromosomes with Voronoi diagrams and Delaunay Triangulations. (a) possible cut points on the contour line, (b) a Bounded Voronoi diagram of possible cut points and (c) Delaunay Triangulation of possible cut points and the centre of the overlapping area. Target cut points are connected to the centre of the overlapping area. •••, ▁ , ■ and • represent chromosome contour, Delaunay Triangulation, crossing centre and target cut points, respectively

Separation of chromosomes

Overlapping chromosomes are separated by dividing the contour set into two contour sets using the four cut points. Given a contour set $A \equiv \{a_1, a_2, a_3, ..., a_n\}$, B four cut points set $B \equiv \{a_i, a_j, a_k, a_l\}$ where $i<j<k<l$ set A can be separated into set A_1 and A_2 by Eq. (5) (**Figure 4c**).

$$A_1 \equiv \{a_1, a_2, a_3, ..., a_i\} \cup \{a_j, a_{j+1}, a_{j+2}, ..., a_k\} \cup \{a_l, a_{l+2}, a_{l+3}, ..., a_n\}$$
$$A_2 \equiv \{a_i, a_{i+2}, a_{i+3}, ..., a_j\} \cup \{a_k, a_{k+2}, a_{k+3}, ..., a_l\} \quad (5)$$

Chromosome data set description

Amniotic fluid cells were acquired from the Human Genetics Unit, Faculty of Medicine, Prince of Songkla University. A total of 80 complete cell metaphase images were collected mostly with overlapping chromosomes. Thirty-five sub-images of overlapping chromosomes were randomly selected and cut from the original images. Resulting images were verified by a cytotechnician from karyotype images.

RESULTS AND DISCUSSION

Twenty eight out of the 35 overlapping chromosomes images can be separated correctly (i.e. 80.0 %). Three out of 35 images are separated incorrectly (i.e. 8.6 %). Four out of 35 images are not separable by our algorithm (i.e. 11.4 %). Our results are comparable with Agam and Dinstein's [9] study. They separated overlapping chromosome images as follows: 82 % correctly, 7 % incorrectly, and 11 % were non-separable.

The algorithm works well with long overlapping chromosomes (**Figures 5a-d**). In some cases, the algorithm gives 5 - 6 target-cut points of which 4 out of the 5 - 6 target cut points are the correct target cut points. However, overlapping chromosome images that contain 5 - 6 target-cut points will be separated incorrectly. An Extended algorithm is needed in these cases. Four points nearest to the centre of the overlapping area are selected. These nearest points may be used because the centre of the overlapping area is obtained by the possible cut points at the overlapping area using its geometrical property (i.e. Voronoi diagram).

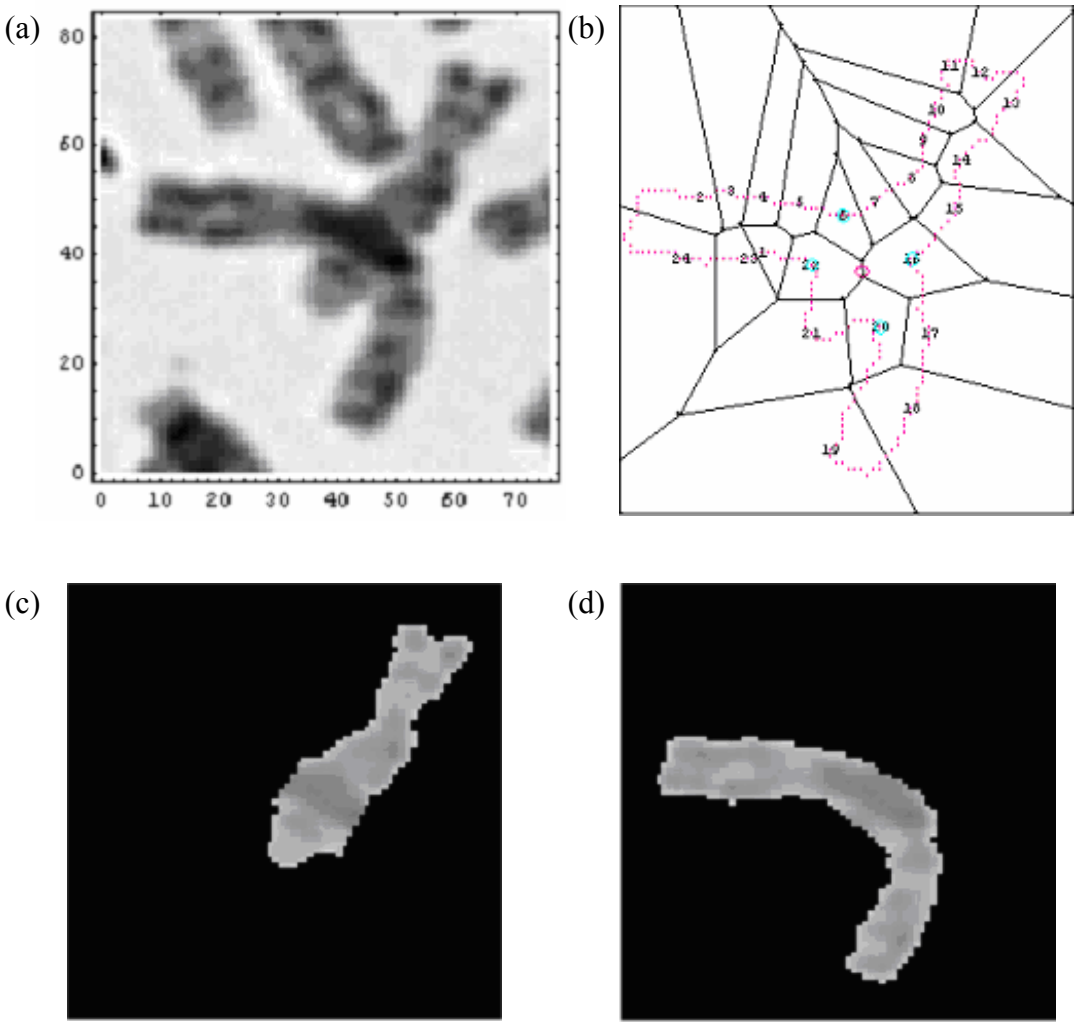
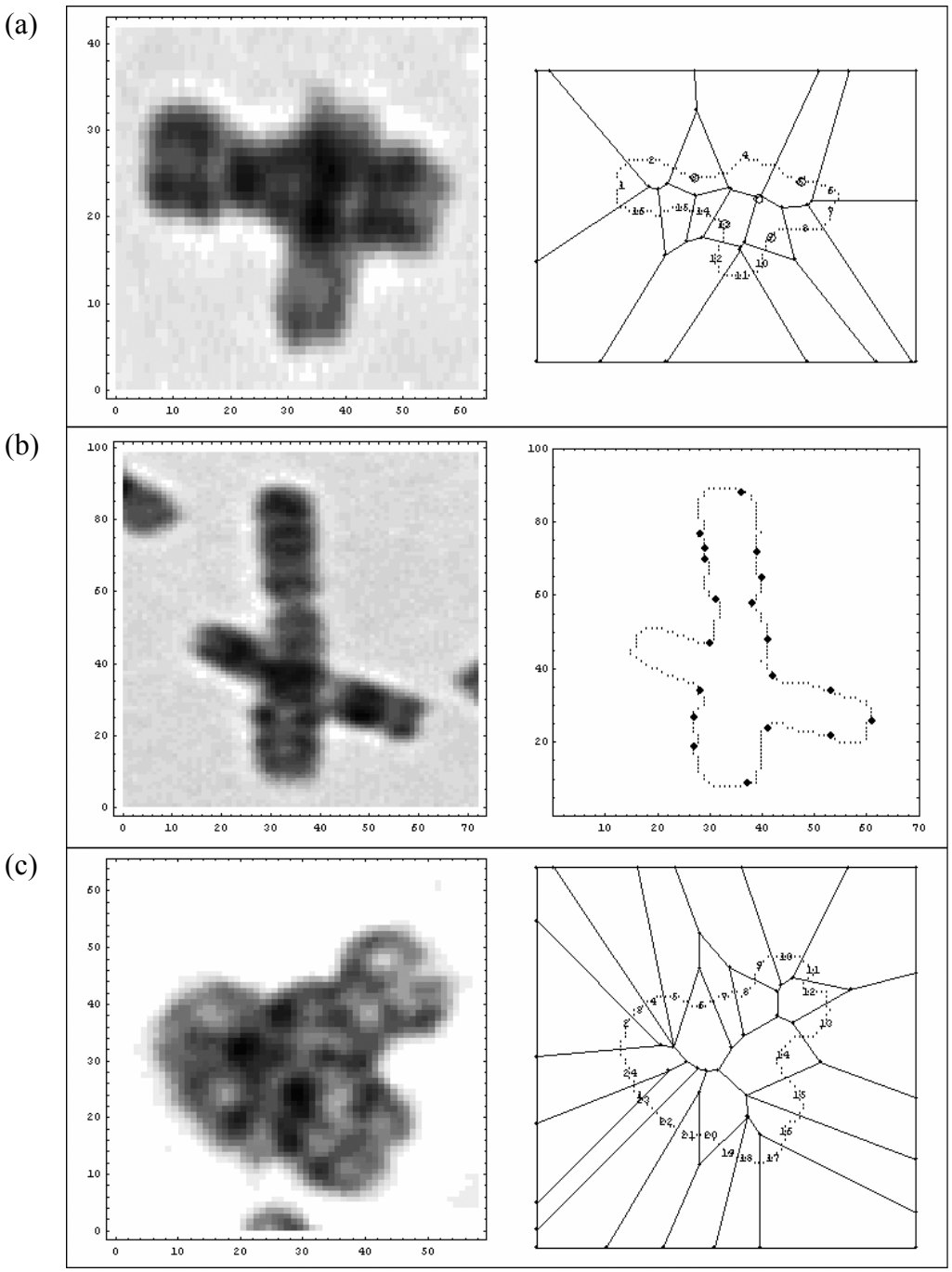


Figure 5 Separation overlapping chromosome images. (a) original image of two overlapping chromosomes, (b) a Bounded Voronoi diagram of possible cut points, (c) chromosome 1, and (d) chromosome 2.

Incorrect chromosome segmentation occurred in four cases. First, the algorithm provides the wrong four possible cut points that are around the overlapping area (**Figure 6a**). Second, the algorithm cannot find a Voronoi diagram of possible cut points (**Figure 6b**). Third, the algorithm cannot find a centre of overlapping area (**Figure 6c**). Finally, this algorithm can be used only with X-shape overlapping chromosomes. Therefore, with T-shape overlapping chromosomes, the algorithm will give incorrect chromosome separation (**Figure 6d**). T-shape Chromosomes cannot be separated because possible cut

points are not at high concave pixels. For future work, this technique could be used to segment three and four overlapping chromosome images and T-shape chromosomes by extending our algorithm.



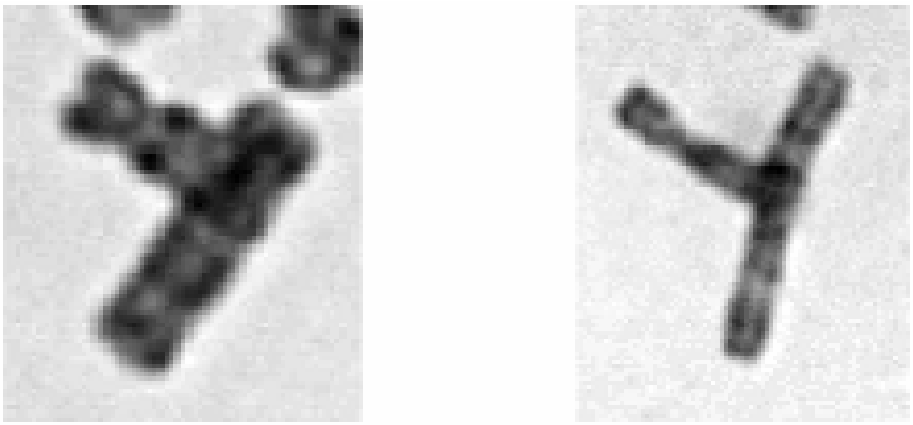
(d) 

Figure 6 Incorrect separation. (a) the algorithm gives the wrong four points, (b) the algorithm cannot find a Voronoi diagram of possible cut points, (c) the algorithm cannot find a centre of overlapping areas and (d) the algorithm cannot be used with T-shape overlapping chromosomes.

ACKNOWLEDGEMENTS

We thank Asst. Prof. Dr. David Harding for comments on our previous versions of this manuscript. We would like to thank Asst. Prof. Dr. Pornprot Limprasert and Mrs. Somkhae Puangpech, Human Genetic Unit, Department of Pathology, Faculty of Medicine, Prince of Songkla University for both data and advisement. This work was supported in part by the Development and Promotion of Science and Technology Talents Project, Thailand and CX-KURUE, the Institute of Research and Development, Walailak University.

REFERENCES

[1] A Carothers and J Piper. Computer-aided classification of human chromosomes: A review. *Statist. Computing* 1994; **4**, 161-71.

[2] B Lerner. Toward a completely automatic neural-network-based human chromosome analysis. *IEEE Trans. Syst. Man Cybern. B: Cybern.* 1998; **28**, 544-52.

[3] FCA Groen, TK Kate, AWM Smeulders and IT Young. Human chromosome classification based on local band descriptors. *Patt. Recog. Lett.* 1989; **9**, 211-22.

[4] J Piper, E Granum, D Rutovitz and H Ruttledge. Automation of chromosome analysis. *Sig. Proc.* 1980; **2**, 203-21.

[5] M Moradi and SK Setarehdan. New features for automatic classification of human chromosomes: A feasibility study. *Patt. Recog. Lett.* 2006; **27**, 19-28.

[6] J Liang. Intelligent splitting in the chromosome domain. *Patt. Recog.* 1989a; **22**, 519-32.

[7] J Liang. *Decomposition of overlapping chromosomes*. *In*: C Lundsteen and J Piper (ed). Automation of Cytogenetics, Springer, Berlin, 1989b, p. 177-90.

[8] J Liang. Fully automatic chromosome segmentation. *Cytometry* 1994; **17**, 196-208.

[9] G Agam and I Dinstein. Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification. *IEEE Trans. Patt. Anal. Mach. Intell.* 1997; **19**, 1212-22.

[10] B Lerner, H Guterman and I Dinstein. A classification-driven partially occluded object segmentation (CPOOS) method with application to chromosome analysis. *IEEE Trans. Sig. Proc.* 1998; **46**, 2841-47.

[11] M Popescu, P Gader, J Keller, C Klein, J Stanley and C Caldwell. Automatic karyotyping of metaphase cells with overlapping chromosomes. *Comp. Biol. Med.* 1999; **29**, 61-82.

[12] GC Charters and J Graham. Trainable grey level models for disentangling overlapping chromosomes. *Patt. Recog.* 1999; **32**, 1335-49.

[13] GC Charters and J Graham. J. Disentangling chromosome overlaps by combining trainable shape models with classification evidence. *IEEE Trans. Sig. Proc.* 2002; **50**, 2080-85.

[14] R Ogniewicz and M Ilg. Voronoi skeletons: Theory and applications. *In*: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Illinois. 1992, p. 63-9.

[15] M Styner, G Gerig, J Lieberman, D Jones and D Weinberger. Statistical shape analysis of neuroanatomical structures based on medial models. *Med. Image Anal.* 2003; **7**, 207-20.
[16] CM Boesse, MK Henry, TW Hyde and LS Matthews. Digital imaging and analysis of dusty plasmas. *Adv. Space Res.* 2004; **34**, 2374-78.
[17] PK Sahoo, S Soltani and AKC Wong. A survey of thresholding techniques. *Comp. Vis. Graph. Image Proc.* 1988; **41**, 233-60.
[18] RM Haralick and LG Shapiro. *Computer and Robot Vision*, Vol. I, Addison-Wesley, Reading, MA, 1992, p. 1-608.
[19] N Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* 1979; **9**, 62-6.
[20] S Wolfram. *The Mathematica Book*, 5th ed., Book News, Oregon, 2004, p. 1-1200.
[21] H Freeman and LS Davis. A Corner Finding Algorithm for Chain Coded Curves. *IEEE Trans. Comp.* 1977; **26**, 297-303.
[22] T Pavlidis. Algorithms for Shape Analysis of Contours and Waveforms. *IEEE Trans. Patt. Anal. Mach. Intell.* 1980; **2**, 301-12.
[23] A Rosenfeld and AC Kak. *Digital Picture Processing*. Academic Press, California, 1982, p. 1-349.
[24] W Srisang, K Jaroensutasinee and M Jaroensutasinee. Segmentation of overlapping chromosome images. *In*: Abstracts of the 30th Congress on Science and Technology of Thailand, Bangkok, Thailand. 2004, p. 40.

บทคัดย่อ

วัชรพงศ์ ศรีแสง[1] กฤษณะเดช เจริญสุธาสินี[1] และ มัลลิกา เจริญสุธาสินี[1]

การแบ่งภาพโครโมโซมที่ซ้อนทับกันโดยใช้วิธีการคำนวณแบบเรขาคณิต

การจำแนกโครโมโซมแบบอัตโนมัติในปัจจุบันยังคงเป็นแบบที่ต้องการมนุษย์เข้าไปช่วย โดยเฉพาะในกรณีที่โครโมโซมสัมผัส หรือ โครโมโซมซ้อนทับกัน โครโมโซมเหล่านี้ต้องอาศัยเทคนิคพิเศษในการจำแนกเพราะว่า โครโมโซมเป็นวัตถุที่ไม่มีรูปร่างแน่นอน การศึกษาครั้งนี้ได้พัฒนาเทคนิคใหม่ขึ้นเพื่อแยกโครโมโซมซ้อนทับกันโดยอาศัยหลักการของการคำนวณแบบเรขาคณิต เทคนิคนี้ต้องจำแนกจุดที่สามารถตัดได้ทั้งหมดจากเส้นขอบโครโมโซมที่ซ้อนทับกัน ใช้โวโรนอยไดอะแกรม (Voronoi diagram) และ ดีเลานี ไตรแองกูเลชั่น (Delaunay triangulation) เพื่อเลือกจุดตัด 4 จุด และตัดโครโมโซมซ้อนทับกันออกเป็น 2 โครโมโซม เราทำการทดสอบโปรแกรมที่เราได้เขียนขึ้นโดยใช้โครโมโซมซ้อนทับกันจำนวน 35 รูป และพบว่าสามารถจำแนกโครโมโซมซ้อนทับกันได้ถูกต้องจำนวน 28 รูป คิดเป็น 80.0 % จำแนกโครโมโซมซ้อนทับกันผิดจำนวน 3 รูป คิดเป็น 8.6 % และไม่สามารถจำแนกโครโมโซมซ้อนทับกันจำนวน 4 รูป คิดเป็น 11.4 %

[1] สำนักวิชาวิทยาศาสตร์ มหาวิทยาลัยวลัยลักษณ์ อำเภอท่าศาลา จังหวัดนครศรีธรรมราช 80161