

Research of Segmentation Algorithms for Overlapping Chromosomes*

Wenzhong Yan, Lei Bai

Department of Computer, North China Institute of Science and Technology, Beijing, China
Email: yanwenzhong@ncist.edu.cn

Received 2013

ABSTRACT

Chromosome segmentation is the most important step in the automatic chromosome analysis system. Since in almost every metaphase image partial touching and overlapping of chromosomes are a common phenomenon, how to separate these chromosomes correctly is a difficult yet vital problem. A number of attempts have been made to deal with this problem. This paper is focused on these attempts. Some algorithms are investigated. The principle and the realization of these algorithms are analyzed. Results of these algorithms are compared and discussed.

Keywords: Chromosome; Segmentation; Overlapping; M-FISH

1. 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]. A computer-assisted system usually includes four processing steps: 1) image enhancement, 2) chromosome segmentation (detection) and alignment, 3) feature computation and selection and 4) chromosome classification [2]. In these steps, chromosome segmentation is the most important one because this step affects the performance of these systems.

Since in almost every metaphase image partial touching and overlapping of chromosomes are a common phenomenon (**Figure 1** shows an example of overlapping chromosomes), finding solutions for automated separation of overlapping chromosomes is difficult yet vital. Early studies found that a number of automated classification systems were somewhat successful in karyotyping the chromosomes under favorable imaging conditions. The typical case error rate was approximately 20% [3]. If the chromosomes were touching, overlapping or deformed, the classification error rate was substantially increased [4].

In order to solve the problems above, a number of attempts have been made to deal with segmentation of overlapping chromosomes. This paper is focused on these attempts. Some segmentation algorithms are investigated. The principle and the realization of these algorithms are

analyzed. Results of these algorithms are compared and discussed.

2. Segmentation Algorithms

A novel recursive searching algorithm was developed and implemented in an automated karyotyping system (AKS). The goal of this system is to label the chromosomes from a metaphase image with minimal human intervention. It is unique in that it is designed to automatically process cells containing overlapping chromosome. It also has four novel algorithmic features: the use of cross section sequence graphs, an over-segmentation-based segmentation strategy, the use of the “pale path” concept for banded chromosomes and the application of mathematical programming for cell-level chromosome classification in cells with overlapping chromosomes. The algorithm was tested using 29 cells (including 397 chromosomes) and achieved 82% accuracy. Three types of skeleton errors were found, which were missing fragments in the skeleton, extra-fragments in the skeleton and hole information [5].

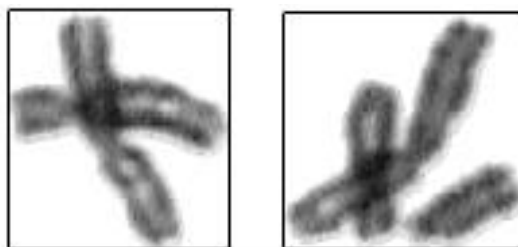


Figure 1. Example of overlapping chromosomes.

*This research is supported by the Fundamental Research Funds for the Central Universities (2011A010).

A new technique to separate overlapping chromosomes based on computational geometry is developed. 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. This algorithm is tested 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 separated incorrectly (*i.e.* 8.6%) and four out of 35 images are not separable by our algorithm (*i.e.* 11.4%). **Figure 2** shows some results of the separation [6].

One research presented an algorithm that is able to automatically identify chromosomes in metaphase images. Researchers take care of a first segmentation step and then of the disentanglement of chromosome clusters by resolving separately adjacencies and overlaps with a greedy approach, which ensures that at each step only the best split of a blob is performed. The performance of this method is better or comparable to the best of other methods reported in the literature. This method to simultaneously tackle segmentation, overlaps and adjacencies with performance on each task higher than 90%, hence providing a tool able to automatically analyze an image and can be handed over with minimal human intervention to a classifier for automatic karyotyping [7].

Another research is aimed at designing a classifier that its input is an image and its output is decision of being

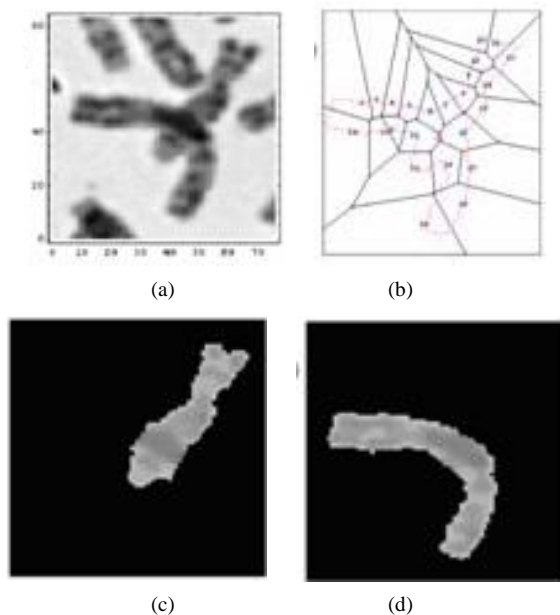


Figure 2. Separation overlapping chromosome images. (a) Original image of two overlapping chromosomes; (b) a Bounded Voronoi diagram of possible cut points; (c) chromosome 1; and (d) chromosome 2.

multi-chromosome or one-chromosome image. Since the gray-scale level of image has no effect on decision of human, researchers convert the images to binary. The output of classifier has two conditions that can be showed with one bit. Therefore, if the classifier output is “zero”, the image is one-chromosome and if it is “one”, the image is multi-chromosome image. If one wants to use a classifier except neural network classifier, he should define a procedure for feature extraction. But if one uses neural network classifier, he doesn’t need to define features for neural network. When a neural network is learning, in fact it learns the features. In this condition, decision space is divided in two regions. Decision boundary is specified during learning process. The best experiment result is obtained in second simulation with 73% true separation. **Figure 3** shows that the best result is 73% and obtained with 8, 16, 20 and 23 neurons in hidden layer. The specifications of this neural network are: multi-layer feedforward perceptron (MLP) neural network with one hidden layer that has 9 neurons in input layer, 8 neurons in hidden layer and one neuron in output layer. Input vector of neural network includes: surface of image, surface of chromosome, number of pixels of boundary of chromosomes, and six momentums [8].

One group proposed a novel approach to image completion for overlapping chromosomes. In their system, only given missing regions, the task can be performed automatically without human intervention. They address the problem of image completion for overlapping chromosomes in the context of a discrete global optimization problem. In order to reconstruct the original chromosome image as faithfully as possible, the objective cost function of this problem is defined under constraint conditions of band patterns in chromosomes image, and cor-

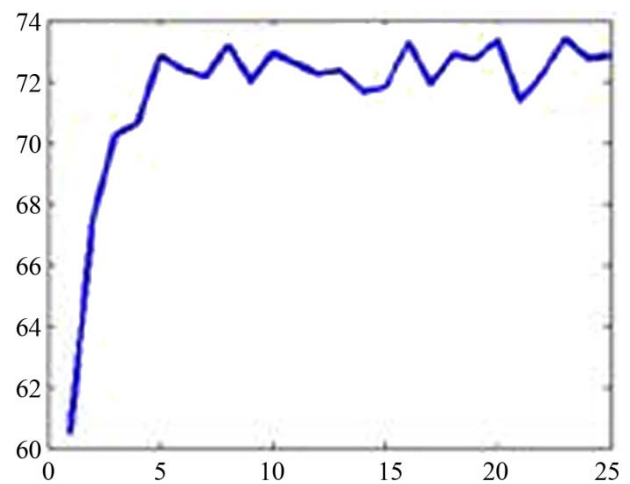


Figure 3. The results of neural network simulation with 9 features at input. Horizontal axis is the number of neurons in hidden layer and vertical axis represents the average percentage of true classification.

responds to the energy of a discrete Markov random fields. For efficiently optimizing this MRF, a loopy Belief Propagation algorithm is utilized to perform it. By using an efficient optimization algorithm, this approach integrates the high-level knowledge of chromosomes image into image completion process to yield good results in comparison to conventional methods. Moreover, this approach can automatically perform without human intervention. **Figure 4** shows some results obtained by this approach to image completion for overlapping chromosomes [9].

3. Algorithms for M-FISH

1) Basic Theory

In the mid-1990s, a new technique for staining chromosomes was introduced. It produced an image in which each chromosome type appeared as a distinct color [10]. This multispectral staining technique is called multiplex fluorescence *in-situ* hybridization, or MFISH, which made analysis of chromosome images easier, not only for visual inspection of the images by humans, but also for computer analysis of the images. M-FISH uses five color dyes that attach to various chromosomes differently to produce a multispectral image, and a sixth dye that attaches to all chromosomes to produce a grayscale image.



Figure 4. Some results of image completion for overlapping chromosomes. (a) Images of overlapping chromosomes; (b) Background chromosome with occluded regions (in green) after segmentation process; (c) Image completion based on the approach.

Thus, it is possible to envision new and improved methods for the location, segmentation and classification of chromosome images by exploiting the color information in M-FISH images.

2) Segmentation Algorithms

One study introduces entropy as a criterion for selecting cut lines to decompose groups of chromosomes that touch and overlap each other. This algorithm uses multi-spectral information in chromosome images for more accurate segmentation. They locate the chromosome material of a sixth dye, DAPI (4,6-diamidino-2-phenylindole nucleic acid stain), which binds to all chromosomes, and then find the connected components. Next, they label every connected component as a separate object and calculate its entropy. If its entropy is above a given threshold, then they examine that object for possible cut lines. They consider only points whose cut line was contained within the chromosome clusters and had an 8-connected neighbor whose class differed from theirs. Candidate cut lines are made by straight lines between all combinations of candidate cut points. Once all entropy-reducing divisions are made, they recombine objects that may have been labeled as separate objects due to overlap. Experiment results show that this algorithm is able to decompose clusters of touching and overlapping chromosomes [11].

An ML method to segment and classify M-FISH chromosomes images using multispectral data was introduced. This method is developed for automatic chromosome identification by exploiting the multispectral information in M-FISH chromosome images and by jointly performing chromosome segmentation and classification. Researchers develop a maximum-likelihood hypothesis test that uses multispectral information, together with conventional criteria, to select the best segmentation possibility. Then this likelihood function is used to combine chromosome segmentation and classification into a robust chromosome identification system. Experiment results show that the proposed likelihood function can also be used as a reliable indicator of errors in segmentation, errors in classification, and chromosome anomalies, which can be indicators of radiation damage, cancer, and a wide variety of inherited diseases. **Figure 5** shows that one single flagged segment can correct a whole cluster [12].

A new decomposition method for overlapping and touching M-FISH chromosomes is also presented. To overcome the limited success of previous decomposition methods that use partial information about a chromosome cluster, researchers have incorporated more knowledge about the clusters into a maximum-likelihood framework. A cluster was better decomposed by incorporating more knowledge. Multiple hypotheses were formed based on color and the geometry defined by the basic elements of

a cluster, and then evaluated based on the pixel classification results and chromosome sizes. A hypothesis that has a maximum-likelihood is chosen as the best decomposition of a given cluster. About 90% of accuracy was obtained for two or three chromosome clusters, which consist about 95% of all clusters with two or more chromosomes, and 100% accuracy was obtained for clusters with a single chromosome. **Figure 6** shows some segmentation results [13].

4. Conclusion

Since in almost every metaphase image partial touching and overlapping of chromosomes are a common phenomenon, finding solutions for automated separation of overlapped chromosomes is difficult yet vital. In this paper, some segmentation algorithms for overlapped chromosomes are investigated. The principle and the realization of these algorithms are analyzed. Results of these algorithms are compared and discussed.

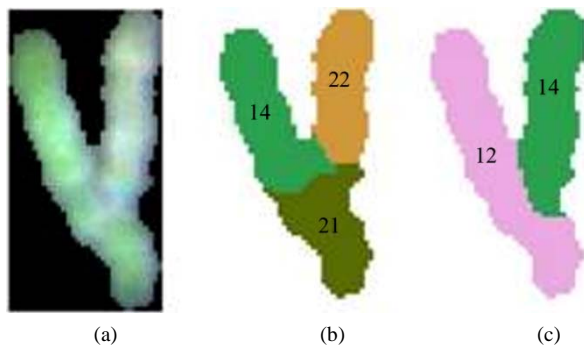


Figure 5. Single flagged segment can correct a whole cluster. (a) M-FISH cluster; (b) Incorrect segmentation and classification; (c) Correct segmentation and classification.

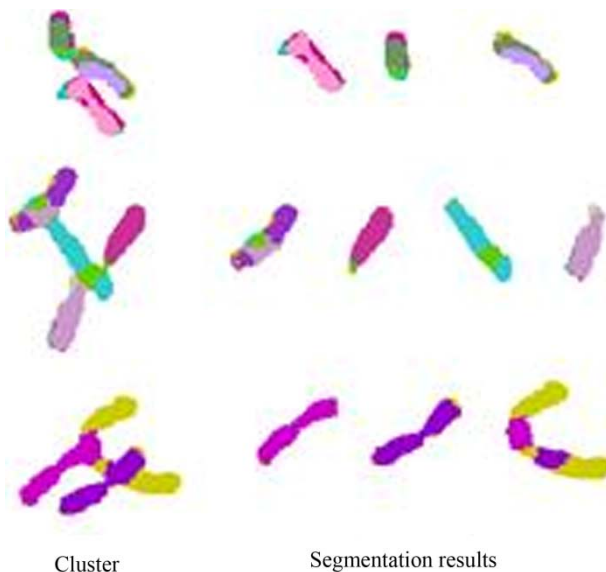


Figure 6. Segmentation results.

REFERENCES

- [1] C. Lundsteen and J. Piper, "Automation of Cytogenetics," Springer-Verlag, Berlin, 1989.
<http://dx.doi.org/10.1007/978-3-642-74738-0>
- [2] J. Liang, "Intelligent Splitting in the Chromosome Domain," *Pattern Recognition*, Vol. 22, No. 5, 1989, pp. 519-532.
[http://dx.doi.org/10.1016/0031-3203\(89\)90021-6](http://dx.doi.org/10.1016/0031-3203(89)90021-6)
- [3] G. Agam and I. Dinstein, "Geometric Separation of Partially Overlapping Nonrigid Objects Applied to Automatic Chromosome Classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 11, 1997, pp. 1212-1222.
<http://dx.doi.org/10.1109/34.632981>
- [4] X. W. Wang, B. Zheng, M. Wood, *et al.*, "Development and Evaluation of Automated Systems for Detection and Classification of Handed Chromosomes: Current Status and Future Perspectives," *Journal of Physics D: Applied Physics*, Vol.38, No.15, 2005, pp. 2536-2542.
<http://dx.doi.org/10.1088/0022-3727/38/15/003>
- [5] M. Popescu, *et al.*, "Automatic Karyotyping of Metaphase Cells with Overlapping Chromosomes," *Computers in Biology and Medicine*, Vol.29, 1999, pp. 61-82.
[http://dx.doi.org/10.1016/S0010-4825\(98\)00040-7](http://dx.doi.org/10.1016/S0010-4825(98)00040-7)
- [6] S. Wacharapong, J. Krisanadej and J. Mullica, "Segmentation of Overlapping Chromosome Images Using Computational Geometry," *Walailak Journal of Science and Technology*, Vol. 2, No. 3, 2006, pp. 181-194.
- [7] E. Grisan, E. Poletti, *et al.*, "Automatic Segmentation of Chromosomes in Q-Band Images," *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 2007, pp. 5513-5516.
- [8] Y. Rahimi, R. Amirfattahi and R. Ghaderi, "Design of a Neural Network Classifier for Separation of Images with One Chromosome from Images with Several Chromosomes," *3rd International Conference on Broadband Communications, Information Technology & Biomedical Applications*, 2008, pp. 186-190.
- [9] X. Z. Zhou, Y. Lu and Z. Y. Yan, "Image Completion for Overlapping Chromosomes," *2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application*, 2008, pp. 413-417.
<http://dx.doi.org/10.1109/PACIIA.2008.271>
- [10] M. R. Speicher, S. G. Ballard and D. C. Ward, "Karyotyping Human Chromosomes by Combinatorial Multi-Fluor FISH," *Nature Genetics*, Vol. 12, 1996, pp. 368-375.
<http://dx.doi.org/10.1038/ng0496-368>
- [11] W. Schwartzkopf, B. L. Evans and A. C. Bovik, "Minimum Entropy Segmentation Applied to Multi-Spectral Chromosome Images," *2001 International Conference on Image Processing*, 2001, pp. 865-868.
- [12] C. S. Wade, C. B. Alan and L. E. Brian, "Maximum-Likelihood Techniques for Joint Segmentation-Classification of Multispectral Chromosome Images," *IEEE Transaction on Medical Imaging*, Vol. 24, No. 12, 2005, pp. 1593-1610.
<http://dx.doi.org/10.1109/TMI.2005.859207>
- [13] H. Choi, A. C. Bovik and K. R. Castleman, "Maximum-

Likelihood Decomposition of Overlapping and Touching M-FISH Chromosomes Using Geometry, Size and Color Information," *Proceedings of the 28th IEEE EMBS An-*

nual International Conference, 2006, pp. 3130-3133.