ORIGINAL RESEARCH PAPER

Computer Science

HUMAN METAPHASE CHROMOSOME ANALYSIS AND FINDING OVERLAPPING CHROMOSOME USING IMAGE PROCESSING

KEY WORDS: chromosome, Centromere , Overlapping

Dr. S. Prabhu Post Doctoral Research Fellow ,Srinivas University, Mangalore, Karnataka, India

Mr. V. Suresh Kumar Assistant Professor, Department of Computer Science ,Shree Venkateshwara Arts and Science (Co-education) College, Gobi, Tamilnadu, India

Mr. E.Sathishkumar Assistant Professor, Department of Computer Science, Park's College (Autonomous), Tiruppur, Tamilnadu, India

ABSTRACT

This approach prevents the possibility of boundary irregularities adversely affecting the centerline and therefore making the width profile measurements noisy. Once the contour is segmented, we then utilized a Laplacian based thickness measurement algorithm where intensity was integrated through a weighting scheme to bias the thickness measurement trace lines into homogenous intensity regions known as chromosome bands. The algorithm is capable of partitioning the telomere region, detecting evidence of premature sister chromatid separation and then correcting for the artifact. Finally, a classifier was trained where the distance from the separating hyper plane was then used as measurement of goodness of fit in order to find the best overlapping and touching chromosomes.

1 . INTRODUCTION

A human chromosome is comprised of DeoxyriboNucleic Acid (DNA) along with protein. The DNA is primarily responsible for genetic inheritance and behavioral patterns of a human being. The genetic makeup and the familiar physical resemblance of a human chromosome is achieved due to the genetic condensation during cell division (mitosis). Therefore by studying the chromosome structure during mitosis, cytogeneticists can identify genetic disorders caused by genetic translocation, deletion, trisomy, monosomy and radiation exposure etc.

2. Proposed Algorithm.

The algorithm requires the user to manually pick a point within (or close to)each chromosome in order to proceed with the rest of the process autonomously. The algorithm assumes that the marked chromosome does not either touch or overlap with other chromosomes in the cell image. This assumption is reasonable due tothe use of a content based ranking algorithm proposed by Kobayashi et al. in thisapproach.

The output of this algorithm was a ranked set of metaphase images where chromosome images that were spread well with minimal overlaps and were complete (contain all 46 chromosomes) were ranked higher. Typically from a given set of cell images, only the highest ranked 5% were selected for further processing. This is a critical step required to improve the accuracy of the proposed algorithm.

3. Preprocessing and segmentation

Chromosome metaphase images are often subjected to uneven illumination and could contain nuclei which appear as bright blobs under the microscope. Since these artifactscan adversely affect the segmentation process, each chromosome was processed individually. A user was required to mark a pixel within or close by the chromosome,which in turn was used to extract a fixed window containing the chromosome as the 'Region Of Interest' (ROI) for further processing in figure 3.

This window had to completely include the chromosome of interest while also including some portion of the background as well. The dimensions of this ROI were set to 201x201 empirically. This value was observed to be sufficient to include all chromosomes in the given data set(collected using the standard 100X magnification). However, if needed, the value maybe changed to accommodate more elongated chromosomes in the future.

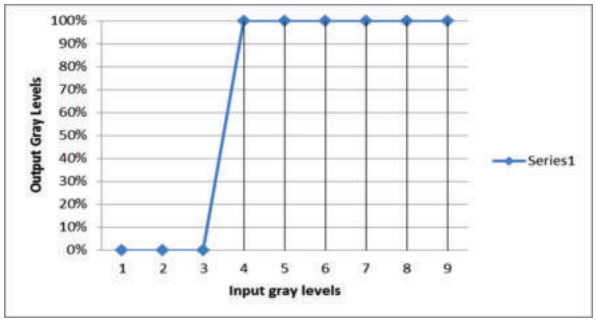

Figure 3.1: The flow diagram of the preprocessing and segmentation stage of the Proposed method.

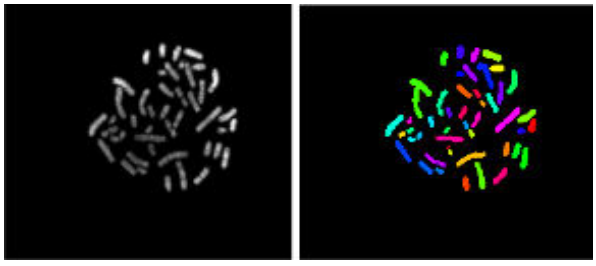

Figure 3.2:The window-center intensity mapping scheme which was used to map a certain intensity range (defined by the window and the center) to the full range(intensity levels 0 - 255).

DAPI images are composed of chromosomes with brighter (higher) intensitieswith a dark (low) intensity background as opposed to Giemsa images which havedarker intensity chromosomes in the bright background. In order to process imageswith both staining methods, intensity values of the DAPI cell images were invertedto obtain an appearance consistent with Giemsa stained images.

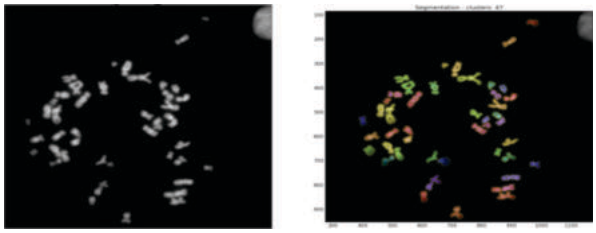

Figure 3.3: effect of intensity normalization on a DAPI

stained image window. Figure 5(a) & (b) depicts the original image window and the corresponding histogram. Similarly the figures 5(c) & (d) depicts the intensity normalized image window and the corresponding histogram.

Using thresholding on an extracted window as opposed to the cell image, reduced the adverse effects of uneven illumination in cell images. However since thresholding being a point processing method, this segmentation method often yields noisy results. This can show up as both individual pixels being marked as objects or as noisy object boundaries due to intensity fading in the vicinity of the chromosome boundary In order for the final stage of the segmentation algorithm to perform well, noise in a digital image needs to be removed or attenuated. Metaphase cells often have background pixels with highly variable intensity values as artifacts created during the light microscopy imaging process. A median filter with element dimensions of5x5 was utilized in order to remove these artifacts from the ROI as a pre-processing step. Unlike Gaussian filtering, median filtering is a non linear filtering process which effectively removes noise from images without blurring object boundaries. Although the amount of noise removal is directly proportional to the size of the filtering element,it also dislocates the object boundary. Therefore a relatively smaller element (5x5)was utilized in this research. Once the image window was filtered for noise, theimage window was then subjected to the next stage of the segmentation algorithmwhere the object boundary from global threshold was utilized as the starting pointsfor a Gradient Vector Flow (GVF) active contour model, which is a variation of the standard parametric active contour model.

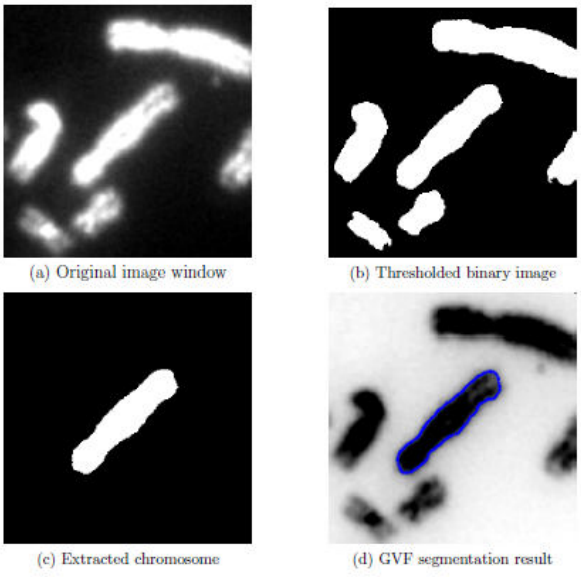

Figure 4.6: image window at different stages of the segmentation algorithm where figure 6 (a) gives the original window containing the chromosome prior to segmentation. Figures 6 (b) and (c) contain the threshold output and the extracted binary object. The GVF outcome is given in figure 6 (d)

The GVF active contour model has the ability to both contract or expand depending on the static vector field created using the object boundary. GVFalso has the capability to converge into boundary concavities which is an important property when analyzing chromosomes with high shape variability. The main internal parameters of the GVF were set at _ = 0:05 (elasticity factor), _ = 0 (rigidityfactor), _ = 0:2 (GVF regularization factor) and _ = 2 (external force weight). This set of values were obtained empirically and yielded satisfactory segmentation results with good convergence into boundary concavities across the entire data set. External energy models for increasing the capture range, along with the advantages of using GVF snakes are discussed in detail in section A.1. The 'Canny edge detection operator' which uses a multi-stage algorithm to accurately detect image boundary edges was utilized for generating the edge map.

REFERENCES

[1] C. Xu and J. L. Prince, "http://iacl.ece.jhu.edu/projects/gvf." A web site for the GVF snake demos and examples.
[2] L. J. Latecki and R. Lak¨amper, "Convexity rule for shape decomposition based on discrete contour evolution," Computer Vision and Image Understanding, vol. 73, pp. 441 – 454, March 1999.
[3] D. Pinkel and J. Landegent, "Fluorescence in situ hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4," Proc. Nati. Acad. Sci. USA, vol. 85, pp. 9138–9142, December 1988.
[4] R. King et al, A dictionary of genetics. Oxford university press, fifth ed., 1968.
[5] M. Moradi and S. K. Saterahdan, "New features for automatic classification of human chromosomes : A feasibility study," Pattern Recognition Letters, no. 27, pp. 19–28, 2006.
[6] H. Choi, A. C. Bovik, and K. R. Castleman, "Normalization via expectationmaximization and unsupervised nonparametric classification form-fish chromosome images," Submitted to, IEEE Transaction on MedicalImaging, 2006.
[7] H. Choi, K. R. Castleman, and A. C. Bovik, "Color compensation ofmulti-color fish imags," Submitted to, IEEE transactions on medical imaging, 2006.
[8] "Segmentation and fuzzy-logic classification of m-fish chromosomeimages," International conference on image processing, October 2006.
[9] H. Choi, A. C. Bovik, and K. R. Castleman, "Maximum-likelihood decompositionof overlapping and touching m-fish chromosomes using shape,size and color information," IEEE International Conference of the Engineering in Medicine and Biology Society, September 2006.
[10] "Automatic karyotyping via maximum-likelihood decomposition ofoverlapping and touching m-fish chromosomes using geometry, size, andpixel classification," In preparation, 2006.
[11] A. J. Vander, J. H. Sherman, and D. S. Luciano, Human Physiology, 8th ed. New York: McGraw-Hill, 2001.