IJCSC Volume 5 • Number 1 March-Sep 2014 pp.54-56 ISSN-0973-7391

A Digital Image Processing Algorithm for Automation of Human Karyotyping

Priya Chaku[1], Pooja Shah[2], Sonal Bakshi[3]
Nirma University, Ahmedabad, India
12mcec02@nirmauni.ac.in, pooja.shah@nirmauni.ac.in, sonal.bakshi@nirmauni.ac.in

Abstract. This research work aims to propose and implement a compute efficient application for augmenting the process of paired chromosome alignment called karyotyping. The methodology is a combination of various digital image processing techniques which serves as an effective solution for analyzing and processing raw microscopic images. The different phases for analysis include image acquisition, object segmentation, feature selection and extraction, and finally classification. However, certain open issues need to be tackled such as presence of marker, twisted or overlapped chromosomes. Hence, the consistency of this system relies on image quality and resolution. Manual intervention to some extent accounts for robust identification and classification.
Keywords: Microscopic Imagery, Automatic Karyotyping, Pattern Classification, Feature Extraction, Chromosomes, Digital Imaging

I. Introduction

With reference to our earlier report on automatic karyotyping [1], we further elaborate on the technical application aspects of automation in karyotyping.

This research in the area of image processing of digital photomicrographs of metaphase chromosomes is a multidisciplinary attempt employing cytogenetics and computerized digital image processing to achieve faster results, easier back-up, processing and data sharing. The karyotyping is carried out by cytogeneticists using various cell culture methods to arrest the dividing cells in metaphase and stain them following protein degrading enzyme treatment that leads to varying degree of uptake of stain leading to light and dark bands. The normal human (Homo sapiens sapiens) contains 23 pairs of chromosomes [3, 4] including 22 pairs of autosomes and one pair of sex chromosomes which differentiates between male having XY and female having XX. The pair of homologous chromosomes are identical, and made up of two sister chromatids joined to each other by a centromere. Depending on the position of the centromere a chromosome is divided in to upper and lower arm. The upper arm being smaller, is called p (petite) and the lower one is denoted 'q' as an opposite of 'p', both are equal in case of metacentric type of chromosomes. The 23 chromosome pairs are grouped in seven divisions from A to G based on size and centromere positions. The A & F being metacentric, D being acrocentric, and the rest being sub-metacentric. The X chromosome belongs to sub-metacentric and Y belongs to small sub metacentric. The positions of each pair has been defined on a typical karyotype form where the pairs are arranged in a decreasing order of size, followed by centromere location and signature light and dark banding pattern which provides milestones to the cytogeneticist to identify the chromosome pairs. The microscopes attached with a camera, developing of negative films and printing black and white photographs in the dark-room followed by cutting and pasting individual chromosome-pairs on a karyotype form used to be routine work of laboratory personnel dealing with clinical management of patients with genetic conditions including cancer, prenatal diagnosis of birth defects etc. Many of the cancer cases and prenatal cases are to be treated as medical emergency as the chromosomal analysis guide the clinical decision regarding diagnosis, therapy selection, termination of pregnancy etc. and hence was required to be as prompt as possible. The multidisciplinary approach involving state-of-the-art engineering of microscopes with high resolution, planatic and chromatic aberration free optics, cooled coupled device (CCD) camera, frame grabber cards, and computer having compatible monitor and processor has changed the face and pace of cytogenetic analysis. The field observed under the microscope following 600X to 1000X magnification is displayed real time on the computer monitor which is saved in a proprietary (or program specific) format as 'metaphase' and following various steps of analysis the final output is saved as 'karyotype' having unique id for each pair of images. Each processing step also needs to be saved as history and should be 'undoable.' As per the requirement of GLP (good laboratory practice) the unprocessed 'raw' image needs to be available as a third file for each, which is also required for medico-legal issues.

The cluster of 46 chromosomes in normal condition as well as few more or less in various cancer and other disease conditions are expected. In addition, due to the nature of the method, many artifacts, varying degrees of staining intensities etc. require skilled human intervention and prevent total walk-away automation. However, the current methods of computer image analysis has proven to be valuable for automation and requirement of skilled human intervention has never been the final goal as it will be unrealistic and technically not sound.

II. Approach

We have prepared the algorithm to maximize automation in an error-free manner. A bigger dataset as well as variation in quality of native images has brought about a considerable change in the proposed technique, so as to enhance the accuracy of the output and reducing the computation required.

To develop an automated system for chromosome orientation and pairing, the images in gray scale format are processed and enhanced. A cluster of chromosomes along with slide artifacts need to be separated using intelligent algorithm. The target is to obtain an output image with well-defined vertically aligned chromosomes in the order of length, centromere position and band patterns. The first phase starts with segmentation [2] of objects for future extraction and analysis. Global threshold value for the image accounts for segmentation of objects from the unwanted background pixels. Thereafter, extraction is performed to administer individual object properties. Each chromosome is extracted the way it is present in the native image. Hence, for karyotyping proper rotation is necessary. Edge linking techniques like local processing and global processing can be used for detecting edge lines. However, the former implementation largely depends on pixel similarity measures like magnitude and angle of gradient vector. Thus the space and time complexity increases for processing each and every pixel. A global processing technique such as Hough Transform on the other hand is a way of finding edge points in an image that lie along a straight line, curves and other structures if their parametric equation is known. It is resistant to noise and can produce multiple instances of a model in a single pass, which proves to be a favorable choice. While vertical rotation takes place, it has to be taken care that the chromosome is properly placed. This implies that the short arm (p-arm for petite) should be positioned above the longer arm (q-arm for queue). This can be achieved by computing the distance of the centromere from both the ends using Medial Axis Transformation [5, 6, 7]. Following phase deals with merging the individual chromosomes into a karyotype image in the order of decreasing length so as to conform to the ideogram rules.

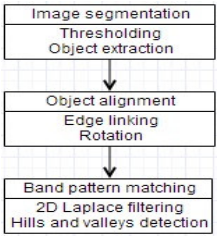

Fig. 1. Proposal overview.

For band pattern matching, which is the final stage in automatic karyotyping, a density profile signal [5] has to be computed which produces a 1D vector, the values of which contain the median values of the pixels orthogonal to the skeleton or medial axis. 2D Laplace filtering [8] is then used for detecting hills and valleys in gray values which is analogous to the band patterns in chromosomes.

III. Implementation

MATLAB 7.12.0.635 (2011a) along with the Image Processing Toolbox has been used for executing the defined proposal. The system takes a native image of any file format as input, and if required converts it to gray scale for processing. A global threshold value is computed using Otsu's method [2], which chooses the threshold to minimize the interclass variance of the thresholded black and white pixels. Also, minute impurities which may affect processing are removed using morphological operator. After labeling the connected components, their region properties need to be measured so as to accurately define the objects and extract them to individual image files. These extracted chromosome files may also contain certain images with large unwanted artifacts or conjoined chromosomes. These may be either deleted or re-segmented using manual intervention, depending on user choice. Next, boundary thresholding is performed which takes into consideration the use of morphological operation, by setting a pixel to zero if its 4-connected neighbors are all ones, thus leaving only boundary pixels. Skeletonization is yet another morphological operation which removes pixels on the boundaries of objects without allowing objects to break apart. These steps are required in order to facilitate proper chromosome rotation. As mentioned before, Hough Transform computes the edge linkages of the images. The angle of the edge determines the amount of rotation required for vertical chromosome alignment. As a result of bi-cubic interpolation due to rotation, pixel values are produced which lie outside the original range. Removal of such pixels is obligatory which may otherwise interfere with further image processing techniques. Lengthwise arrangement of chromosomes is also done in the meanwhile, which may ease the process of ordered karyotyping.

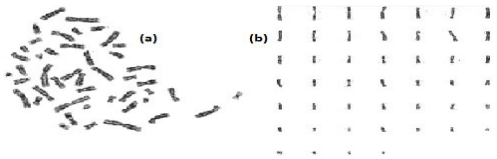

Fig. 2. (a) Chromosome smear and (b) Arranged chromosomes.

There may be disputes where the chromosome after being vertically aligned needs to be flipped upside down. This facility is provided to the users for pointing out the particular chromosome. After complete pre-processing of individual chromosomes, it is essential to merge them together in a single karyotype image. Image cleaning is performed to eliminate background pixels due to presence of residual artifacts or interpolated pixels. For attaining a clear output, the image is traced with green boundary pixels around the chromosome outline. It has been identified that presence of green pixel gives a truth value, 255, in the g-component of the image already decomposed into r-, g- and b-components. Hence, pixels lying outside green boundaries of the chromosome outline can be excluded. Band pattern matching can be accomplished using Laplace filtering by matching the band components of chromosomes which are represented in the form of hills and valleys in gray values. The user interface also offers a facility to manually mark the region of interest in the native image using mouse pointer positions.

IV. Conclusion

Although the system works effectively on most real-time images, more calibration is still essential to solve the disputable concerns so as to generalize and cater to the varying needs of academicians, bioinformaticians and cytogeneticists. Presence of marker chromosomes which resemble artifacts might be accidentally removed. Conjoined, twisted, overlapped or over-segmented chromosomes [3, 7] also pose a challenge in image capture, extraction and analysis.

Fig. 3. (a) Overlapped, (b) Twisted, (c) Marker, (d) Over-segmented and (e) Conjoined chromosomes.

The native images can be acquired at various intensity levels; hence thresholding ought to be performed accordingly. Genetic disorders, where the human chromosomes vary in number (less than or greater than 46), can affect the ordering and pairing of karyotype. There may be circumstances where vertical alignment may result in positioning the chromosomes upside down. Thus, overturning the specified chromosomes is mandatory for precise arrangement. At times, manual input becomes vital for masking the region of interest in the microscopic images. Lastly, exact matching or coupling might also require human involvement with features of drag and drop or control buttons.

Acknowledgements. This research was supported by the Institute of Technology and Institute of Science, Nirma University, Ahmedabad, India. The authors are grateful to the eminent and knowledgeable scientists of Unipath Specialty Laboratory Pvt. Ltd., Ahmedabad, India for their fruitful comments and guidance. The authors would also like to thank the notable researchers working in Supratech Micropath Laboratory & Research Institute, Ahmedabad, India for their help in providing images and useful comments.

References

[1] Chaku, P., Shah, P. Automatic Karyotyping of Human Chromosomes using Band Patterns. International Journal of Scientific Research, 2013.
[2] Wayalun, P., Laopracha N., Songrum P. et al. Quality Evaluation for Edge Detection of Chromosome G-band Images for Segmentation. Applied Medical Informatics, Romania, 2013.
[3] Holland, D. Chromosome Analysis: Banding Patterns and Structural Aberrations, 1980.
[4] Xiong, Z, Wu Q., Castleman K. Enhancement, Classification and Compression of Chromosome Images.
[5] Moradi, M. and Setarehdan, S. New features for automatic classification of human chromosomes: A feasibility study. Pattern Recognition Letters, Elsevier, 2006.
[6] Morardi,M., Setarehdan S., Ghaffari S. Automatic Landmark Detection on Chromosomes' Images for Feature Extraction Purposes. Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, 2003.
[7] Abe,T., Hamada C., Kinoshita T. Chromosome Region Recognition Based on Local Band Patterns.
[8] Wu, Q., Castleman, K. Wavelet-based Enhancement of Human Chromosome Images. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1998.
[9] Shaffer, L., Sloak, M., Campbell, L. ISCN 2009: An International System for Human Cytogenetic Nomenclature. Basal: S Karger, 2009.