Volume : 2 | Issue : 12 | December 2013 • ISSN No 2277 - 8179

Research Paper

Automatic Karyotyping of Human Chromosomes Using Band Patterns

Engineering

KEYWORDS : automatic karyotyping, chromosomes, microscopic image processing, pattern classification

Priya Chaku Institute of Technology, Nirma University, Ahmedabad, India

Pooja Shah Institute of Technology, Nirma University, Ahmedabad, India

ABSTRACT

The target of this research is to develop a simple, improved and efficient system for automated karyotyping. It involves the analysis and processing of a native image of captured chromosomes so as to produce a paired and classified arrangement of the same, also called as a karyotype. The procedure follows enhancement, extraction and classification. The limiting factors for efficiency could be image resolution, overlapped or over-segmented chromosomes, and presence of artifacts. Thus, the reliability and accuracy of the system developed largely depends on solving these issues. This research work offers implementations to overcome these challenges.

INTRODUCTION

Microscopic image processing involves usage of digital image processing to acquire, analyze and process images from a microscope. Digital image processing uses spatial and frequency-altering techniques for processing digital images so as to improve their quality; which is now applicable in various fields like medical imaging, biological research, metallurgy etc. Acquisition of images is incorporated using a high resolution CCD camera attached to a microscope. Image processing allows large scale in-depth evaluation of images which are not visible to the naked eye; and used in the areas of cytology and histology. Microscopic imagery targets images captured at microscopic levels and helps to derive useful information from them. The cells behave in a unique manner at a micro level, and hence need to be processed to extract required information; which is the objective of this research.

For chromosome analysis, a culture of rapidly dividing cells needs to be achieved using by extracting a sample of blood and centrifuging it to single out the white blood cells. These are further cultured in a medium which accelerates mitosis. Colchicine is added after 72 hours of incubation, so as to stop mitosis after metaphase is reached. A hypotonic solution is also supplemented to enlarge the cells and separate out the chromatids. Finally, these cells are mounted on glass-slides and are subjected to Giesma stains. It suffices for dark and light colored bands on the chromosomes which help in uniquely identifying individual pairs of chromosomes. They are then trimmed out from the micro-photograph and aligned in pairs. This procedure is termed as karyotyping. Chromosomes are supercoiled DNA associated with its chromatin. They consist of two sister chromatids conjoined at the centromere. It partitions the chromosome into two pairs of short and long, namely the p and q arms. A global representation has been accepted to distinguish the different band patterns developed by Giemsa staining.

There are 23 pairs of human chromosomes in a single cell. Each autosome is designated a number, 1 to 22, and the pair of sex chromosomes are attributed as X and Y. The arms are denoted with a p or q.

MOTIVATION

Efforts have been made in this area to address the limitations encountered by automatic karyotyping systems and also improve its efficiency. The objective of this work is to address the issues and propose solutions as compared to the previous works for similar kinds of problems. This is due to the motivation behind the work to make the proposed techniques innovative, simple and less computational.

CLASSIFICATION TECHNIQUES

- **Wavelet-based Enhancement**

Cubic-spline basis functions show high similarity to the shape of chromosome bands. This wavelet-based transform reproduces the chromosome as a weighted-sum of basis functions, which allows segregating the chromosomes into band-like units in the transform-domain. Thus we can efficiently augment the band-patterns by carefully analyzing the wavelet-transform coefficient image at desired neighborhood, orientation and scale using space-frequency localization and multi-resolution representation.

- **Appearance-based Object Recognition**

Chromosome images are directly projected into a subspace for reduction in the dimension. LDA (linear discriminant analysis) is a subspace optimized over discrimination between classes and PCA (principal component analysis) is a subspace optimized over representing data with least principal components.

- **Automatic Landmark Detection**

The length and centromeric index of a chromosome play an important role in locating the important landmarks for classification purposes. These include the centromere, the medial axes; and the end- and branch-points in the medial axis.

After binarization, the horizontal and vertical projection vectors are measured by integrating the pixel values along the projection lines. After plotting the values along the horizontal projection lines, a central minimum is obtained on the graph, which is considered as the centromere position. Thus, the centromere location on the medial axis is also approximated using the same projection plot. Starting from the centromere, if we plot right angles across medial axis, only one pixel of the medial axis would intersect with the orthogonal lines.

- **Local Band Patterns**

Every chromosome is considered to be an array of sub-regions, and searching of the sub-regions is performed for recognizing individual chromosomes. Templates for each chromosome are prepared which are divided into sub regions called local band patterns. The chromosome is to be searched for the sub-region similar to the local band-pattern. If found; the adjacent sub-regions are iteratively examined for the analogous local band-patterns. Thus the entire chromosome region is detected in the image.

- **Artificial Neural Networks**

After obtaining the medial axis of the binary image of the chromosome smear, the central curve determines the longitudinal dimension of the chromosome. It is further extrapolated at the branch points of the medial axis with the chromosome boundary and the median formed by the triangle formed by the two branches.

The density profile is computed by the intensity variations across the chromosome axis, which represents the characteristic bands. A typical three-layered feed-forward perceptron neural-network is used for classification, which is to be trained using the back-propagation learning rule. The dataset is partitioned into training, validation and testing sets. The training subset is responsible for modifying the ANN arguments; the testing subset is responsible for evaluation, and the validation

Volume : 2 | Issue : 12 | December 2013 • ISSN No 2277 - 8179

set is responsible for monitoring the classification error caused during the training process so as to prevent over-fitting.

COMPARISON

- **Wavelet-based Enhancement**

Wavelet-based techniques allow interactive detection and enhancement of chromosome images at desired scale, orientation, neighborhood and degree, thus enabling flexible and efficient automated analysis. Also, the band patterns can be detected meticulously and selectively enhanced using basis functions for denoting the bands. However, the results depend highly on shift, gyration and scale of the provided image, because of down sampling of DWT.

- **Appearance-based Object Recognition**

PCA is believed to be better for small training dataset while LDA supports classification with nearest neighbor classifier.

- **Automatic Landmark Detection**

The computational positions of the landmarks tend to show high accuracy. The procedure is simple and is based on the fact that chromosomes are geometric objects. However, highly bent or twisted chromosomes are not suitable for such practices.

- **Local Band Patterns**

Proper initial adjustments of dimensions and intensity of the given image with those of the provided templates improves the efficiency of local band pattern recognition. Also, the technique cannot differentiate the chromosome regions from the background.

- **Artificial Neural Networks**

The neural networks that have been trained with six-dimensional feature vectors gives positive results for classification rates. More attention needs to be paid to the feature-extraction process so as to ascertain complete separation of the classes by the new feature set. The accuracy of these values is limited by the illumination conditions and the feature extraction process.

CONCLUSION

The development of an accurate and automated karyotyping system is very important for cytogeneticists to diagnose diseases and abnormalities. Till now, no such fully automated system developed by the numerous manufacturers has been 100% reliable. Therefore the approach for this research aimed at the limitations on efficiency and to project simple effective solutions for further enhancement.

PROPOSAL

For developing an automated system for analysis of chromosomes and their classification into pairs, certain uniquely defined features need to be taken into consideration. The system is developed for high resolution image formats like .bmp, .tiff and .jpg. Also, the system will support images having non-overlapping straight chromosomes.

The objective of the proposed system is to develop a cost-effective alternative for commercially available softwares. The initial preprocessing step involves enhancing the image for clear representation of individual chromosomes. The analysis begins with discriminating the chromosomes with respect to their length, centromere position, and finally band patterns. For enhancement, a cubic-spline wavelet-based technique is to be applied which shows stark similarity with the shape of the chromosomes as well as decomposes the bands into weighted sums of its basis functions. Edge detection techniques also require histogram equalization to be applied for better results. Robert's edge detection algorithm is found to provide best results for detecting edges in the chromosomes due to contrasting bands, and eventually segmentation of chromosome objects using oriented bounded boxes. The next stage involves computation of medial axis for thinning of binarized chromosome objects. Hence, a density profile signal is computed for finding out the banding patterns. This signal is a 1D vector which contains the average values of the pixels perpendicular to the medial axis. For centromere position, the CI (centromere index) is to be found out which is the ratio of the short arm of the chromosome to the long arm. The projection vectors in the horizontal and vertical directions are measured by integrating the pixel values along the projection lines. After plotting the values along the horizontal projection lines, a central minimum is obtained on the graph, which is considered as the centromere position. The local-band descriptor technique is based on two-dimensional Laplace filtering of the image, which results in the detection of hills and valleys in grey value images. This facilitates pattern matching for chromosome bands.

EXPERIMENTAL RESULTS

MATLAB 7.6.0.324 (R2008a) has been used for performing the initial thresholding, segmentation, extraction and rotation steps using user-defined as well as built-in functions provided by the Image Processing Toolbox. The dataset used as input image is in JPEG format. But other picture formats like TIFF and BMP can also be used.

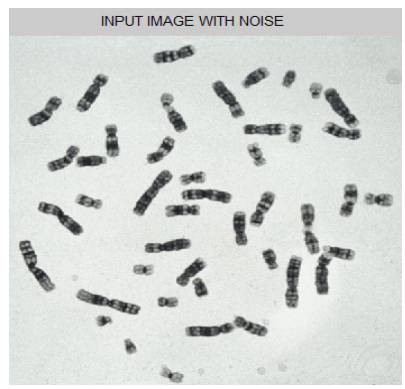

Figure 1: Original image
Sources: www.wellcomeimages.org/ B0000278

Four different Matlab script files are used to obtain the rotated chromosome images. For image segmentation and extraction, the first Matlab function is used to display the original image and perform basic functions for further steps. After converting the image to grey-scale, built-in function is used to generate the global threshold value for future segmentation of individual objects. Also, the different connected components are labelled and are tagged using 'boundingbox' property of 'regionprops' function. For boundary thresholding, a second matlab function is used for every connected component labelled by the boundingbox; the original image is extracted according to their respective coordinate positions.

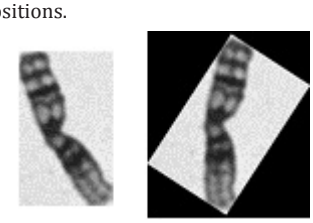

Figure 2: Extracted chromosome object and rotation for uniform vertical alignment
Source: Matlab implementation output

For skeletonization of segmented chromosome objects, this step is necessary to find out the boundary pixels of the objects, which can be used in the future to determine the centromere position. Also, the built-in morphological function 'bwmorph' is used to compute boundary pixels. For rotation of individual chromosomes to uniform vertical alignment, built-in functions for Hough transformation are used to determine the angle of alignment of the medial axes of the chromosomes and are rotated accordingly.

REFERENCE

[1] Wayalun, P., Laopracha, N., Songrum, P. and Wanchanthuek, P. (2013), "Quality Evaluation for Edge Detection of Chromosome G-band Images for Segmentation." Applied Medical Informatics, Romania. | [2] Holland, D. (1980), "Chromosome Analysis: Banding Patterns and Structural Aberrations." | [3] Xiong, Z., Wu, Q. and Castleman, K. "Enhancement, Classification and Compression of Chromosome Images." | [4] Moradi, M. and Setarehdan, S. (2006), "New features for automatic classification of human chromosomes: A feasibility study." Pattern Recognition Letters, Elsevier. | [5] Morardi,M., Setarehdan, S. and Ghaffari, S. (2003), "Automatic Landmark Detection on Chromosomes' Images for Feature Extraction Purposes." Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis. | [6] Abe,T., Hamada, C. and Kinoshita, T., "Chromosome Region Recognition Based on Local Band Patterns." | [7] Wu, Q. and Castleman, K. (1998), "Wavelet-Based Enhancement of Human Chromosome Images." Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.