Development of semiautomatic application prototype to identify g-banded normal human karyotype by size

P Maulia¹, K Kevin², M I Tan¹* and N P Utama²

¹ School of Life Sciences and Technology, Institut Teknologi Bandung, Bandung, Indonesia
² School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, Indonesia

*marsel@sith.itb.ac.id

Abstract. Karyotype is a complete image of chromosomes arranged in accordance with the length, number, shape, and band of an individual cell's chromosomes. Generally, the karyotype is performed by taking the chromosome image at metaphase, cutting out, and identifying each chromosome image based on G-banded manually. This process consumes a lot of time, energy, and inconsistency result. Therefore, it is necessary to develop an application which helps the karyotyping process automatically or at least semi-automatically. In this experiment, we proposed a semi-automatic karyotyping process by utilizing image processing techniques of normal human karyotype. The development of semiautomatic application prototype includes three steps, they were preprocessed chromosome image, featured extraction of chromosome image, and utilized tools to facilitate the users. The methods used in preprocessed chromosome image are gamma enhancement, bilateral filtering, adaptive thresholding, and flood-fill algorithms. The featured extraction of chromosome image is used to extract the morphological features, which will measure the chromosome based on length and pair of chromosomes using skeletonize method. Functionality, tools join and cut for users to modify the results of segmentation chromosome image has been successfully created. Based on the analytical result, it can be concluded that the proposed method can provide a relatively good representation of the chromosomes to assist the further chromosome analysis process.

1. Introduction
Chromosomes are structures that contain genetic information of cells. Normal human cells consist of 22 pairs of autosome and one pair of sex chromosomes. In medical science, the detection of chromosome structure (karyotype) is one of the diagnostic stages that can be performed to detect structural chromosomal abnormalities, such as deletions, insertions, translocations, and numerical chromosomal aberrations (aneuploidy). Chromosomal abnormalities and aneuploidy have many pathological associations, such as affects human development, causing embryonic lethality or developmental defects, and chromosomal aberrations constitute an established hallmark of cancer [1].

Karyotyping is a process to identify chromosomes based on size, length, number, shape, and band of the chromosomes. Chromosomes can be distinguished based on banding patterns with the use of various staining protocols. Many banding techniques have been developed for specific purposes because chromosome banding is crucial for cytogenetics studies. These techniques can be divided into two main groups, techniques that can produce horizontal bands along the length of the entire chromosome (e.g.,
G-, Q-, and R-banding) and techniques that stain only a specific band or region of the chromosome (e.g., C-, T-, and NOR-banding). G-banding became the most widely used in clinical genetics laboratories for identifying normal and abnormal chromosomes [2]. This technique, which is nonfluorescent, is advantageous in the aspect of stability and resolution of the bands produced. Visible light dyes are more stable and capable of producing clearer bands than fluorochromes [3]. The chromosome obtained from G-banding of normal human are shown in figure 1.

![Figure 1. Normal male chromosome image obtained by G-banding technique, (a) before arranged (seen under microscope), (b) after arranged (karyotype) [4].](image)

The manual karyotyping process is usually carried out by specialist biologists or cytogeneticists. They will identify all the chromosomes then cut and place them into smaller groups based on the visual scanning. This visual inspection is a time consuming, tiring, and the costs of this incurred procedure were quite expensive. Hence, automated image chromosome analysis is still an important problem [5]. The automated karyotyping system enables many clinical advantages such as interactive and graphical environments, faster sampling, and better quality, clearer printing, better interpretation, and allows us to store the information in the form of a database for future analysis [6]. Some applications have been developed to analyse chromosome automatically, such as cytovision from Leica and GenASIs Bandview from DSS. However, the price of these applications is still expensive because it must be bought one set with the sophisticated microscope. In this paper, we present a semiautomatic karyotyping application prototype to identify G-banded normal human karyotype that is relatively easy and cost-effective. The semiautomatic karyotyping using image processing methods are divided into three steps. they were preprocessed chromosome image, featured extraction, and utilized tools to facilitate the users.

2. Method and proposed algorithm

2.1. Chromosome isolation method with modification [7]

500 µL whole blood was inoculated in sterile plastic 15 ml tubes containing 8 ml of prewarmed (37°C) RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS), 200 µL phytohaemaglutinin (PHA), and 1% Antibiotic-Antimycotic (100X). The tubes were gently mixed and placed in a rack in an incubation chamber for 48 hours at 37°C ±0.1°C in a 5% CO2 atmosphere. Lymphocytes were arrested at different stages during metaphase adding 200 µL of 10 µg/mL colcemid solution (Gibo Karyomax colcemid, lot No. 15 212-012) to the medium and incubated at 37°C ±0.1°C for 12 min. The culture was centrifuged during 10 minutes at 1300 rpm. The supernatant discarded and the cell pellet was first resuspended with 1.5 ml prewarmed (37°C) KCl 0.075 M and then incubated at 37°C 13
min. The centrifugation step was repeated, the supernatant discarded and the pellet were fixed adding 8 ml of 3:1 methanol:acetic acid solution. The centrifugation step was repeated with methanol:acetic acid solution two times more until the pellet and supernatant looked clear. Finally, the pellet was resuspended in 2 ml fixative. Three drops of resuspended pellet were added per each new, clean slide (wet slide, there are water on the slide’s surface), which were then air dried and checked under phase contrast microscope to assess an adequate cell density. Aging slide for a week in box that contain silica gel. G-banding was performed by dipping slides in 0.125% trypsin solution at 25°C for certain second (according to the incubation, start from 8 second), washed with sterile saline solution two times and then 6% Giemsa stain (Gibco Karyomax Giemsa stain improved R6 solution) was applied for 5 minutes. The slides were washed with aquades, air dried, mounting slide using entellan, and ready to be observed under microscope on maximal magnification.

2.2. Proposed algorithm on image processing

![Figure 2. Proposed algorithm on image processing.](attachment:image.png)

2.2.1. Pre-processing data. The programming language that is used in this program is python and OpenCV. The pre-processing stage is required to get image with a better representation. Pre-processing improves the chromosome image for better quality by doing image enhancement, thresholding, filling, and followed by image segmentation. The result of this stage will help to separate the images from the background. The segmentation process helps in extracting the region of interest which is to be processed. Pre-processing methods are adapted to the literature that has been found associated with chromosome image.

- Image enhancement. The chromosome image quality may be degraded by many reasons such as staining, defective sample, and imaging conditions. Thus, an image enhancement processing algorithm is needed before segmentation of chromosomes image. In this paper, we propose power law transformation or gamma correction algorithm for chromosome contrast enhancement [8].
- Thresholding. After chromosome images have better quality, thresholding is done by setting a threshold for the chromosome area. Thresholding will facilitate the next stage of segmentation. This step using adaptive thresholding algorithm. The result of this algorithm will convert the original image into binary ones. One of the advantages of this method compared to global thresholding is that it reduces the possibility of breaking one chromosome into two parts [9].
- Filling. The results of enlarging and enhancing the quality of chromosome images are improved by filling areas that are less clear and require additional areas inside of chromosome objects that do not appear full. The filling process utilizes the flood fill algorithm [10].
- Segmentation (Contour Extraction). Active contour (contour extraction) is a type of segmentation technique which can be defined as segregation of the pixels of interest from the image for further processing and analysis. Active contour described as active model for the
process of segmentation. This process is used to specifically segregate chromosome area and non-chromosome area [11]. The obtained result is images of chromosomes that have been successfully extracted.

2.2.2. Feature extraction. This stage is done to obtain features that are used to numbering and arrange chromosome images. The feature extraction involves of contouring, centrelines, chromosome length, centromeres determination, and banding patterns of each segmented chromosome images [12].

- Centreline extraction. The centreline of chromosome image is extracted to obtain the length of the chromosome and determine the centromere position. Algorithm that used in this method is skeletonization or medial axis transformation (MAT). This algorithm performs morphological transformation by reducing the outermost region of the chromosome object. This algorithm continues to thin while maintaining the connectedness of the two end points belonging to the line segment. The results of the previous process give a diameter that both ends have not yet reached the boundary area of the chromosome object. Therefore, we need a method for extending the centreline. The line lengthening method is done by the line extrapolation method. Elongation is done on two end points on the midline until it reaches the boundary area of the chromosomes.

- Length chromosome extraction. The length of chromosome image is calculated based on the centreline extraction after the centreline is obtained. The centreline of the chromosome images was computed based on the Euclidian distance transform of the binary image, calculate the distance between 2 points.

- Centromere position determination and banding pattern extraction. The centromere position of the chromosome is generally found in the constricted chromosome region (smaller in body width). The banding pattern of chromosomes is extracted based on chromosome image input and the grey level value of the point traversed by the chromosome centreline.

3. Results and discussion

Normal human chromosomes have been successfully isolated after optimizing the available methods. The results of chromosome image can be seen at figure 3. Based on that results, it is known that female chromosomes provide a clearer banding pattern compared to male chromosomes. The number of lysis cells of female are much better than those of male because female blood cells respond better to PHA than male blood cells. Therefore, the next step on image processing will be used female chromosome image as input.

![Figure 3. Chromosome image obtained by chromosome isolation method and observed under microscope, (a) male chromosome and (b) female chromosome.](image-url)
The first step in image processing is to input the chromosome image to the program. The program has been successfully recognizing the input image as chromosome image. After that, the pre-processing is implemented to the chromosome image and the output as shown in figure 4. The pre-processing method consists of four processes, namely image enhancement, thresholding, filling, and segmentation. The function of image enhancement is to make the image smoother, thresholding method is to separate image from background, the filling process is to cover the gap that appears inside the thresholding chromosome object, and segmentation process is to segregate chromosome area and non-chromosome area specifically. Furthermore, the original image is converted into binary ones.

![Figure 4.](image-url) Chromosome image (a) original input before pre-processing method (b) binary image after pre-processing method.

The binary image is proceeded further to the next step, namely feature extraction. This process consists of three steps. First step is contour extraction. In this step the object of chromosome was separated one to another so that the number of chromosomes can be counted. Second, extract the centreline of the contour using the skeletonize method. Third, measure the length of chromosome using extrapolated centreline. The extrapolation process is performed because the two ends of the midline have not reached the contours of the chromosome object. The result of this feature extraction can be seen at figure 5.

![Figure 5.](image-url) Chromosome image (a) binary image before performed by feature extraction method, (b) the result after processing by feature extraction method.

The graphic interface of the application consists of three areas, which are original image area, working area, and result area. The functionality tools that have been successfully developed are cut and join tools. The cut tool helps the user to edit two or more images of adjacent chromosomes that are considered as one object. The process of cut tools shows at figure 6. Initially the number of chromosomes identified
was 40 and after it performed by cut tool, the chromosome count become 41. It shows that the cut tool is successfully edited the object.

![Figure 6. Functionality of cut tool](image)

The join tool helps the user to join separate or mis segmentation an of chromosome image. The example of process that performed by join tool shows at figure 7. The black arrow shows the area that is applied by join tool. The segmentation area of the image become wider after join tool is applied.

![Figure 7. Functionality of join tool](image)

The forth interface is an area for the final result after the chromosome image has been edited entirely. This application classifies the chromosome image based on their size but not based on the banding features of the chromosome (Figure 8). Process of chromosome classification based on the banding pattern is still on progress due to the complexity of its pattern.
Figure 8. Karyotype chromosome image classified by size, (a) Display of image classified by size in our application we developed, (b) Overall display, (c) Future display of image classified by size.
4. Conclusions
Normal human chromosomes (male and female) have been isolated and the images are successfully captured with clear banding. Image processing algorithms succeed to be implemented to the chromosome image and chromosome can be classified by size. The functionality tools that have been successfully developed are cut and join tools.

Acknowledgment
This research was funded by the Indonesian Ministry of Research, Technology and Higher Education in 2019, through the Community Service Program by the Institute of Research and Community Service (LPPM) Bandung Institute of Technology (ITB), and it was developed in two faculties namely the School of Electrical and Informatics Engineering (STEI) and the School of Life Sciences (SITH).

References
[1] Potapova T A, Unruh J R, Box A C, Bradford W D, Seidel C W, Slaughter B D and Li R 2015 Karyotyping human and mouse cells using probes from single-sorted chromosomes and open source software Biotechniques 59(6) 335-346
[2] Yang X, Wen D, Wu X, Zhao Z, Lacny J and Tseng C 2009 A comprehensive cytogenetics tutorial program, encompassing changeable G-band resolutions Comput. Methods Programs Biomed. 99(1) 66–74
[3] Estandarte A K C 2012 A review of the different staining techniques for human metaphase chromosomes (London: Department of Chemistry, University College London, University of London)
[4] Department of Pathology Cytogenetics Gallery [Online] retrieved from http://www.pathology.washington.edu/galleries/Cytogallery/main.php?file=human karyotypes
[5] Karvelis P, Likas A and Fotiadis D I 2010 Identifying touching and overlapping chromosomes using the watershed transform and gradient paths Pattern Recognit. Lett. 31(16) 2474–2488
[6] Joshi P, Munot M, Kulkarni P and Joshi M 2013 Efficient karyotyping of metaphase chromosomes using incremental learning IET Sci. Meas. Technol. 1–9
[7] Bang C D and Donlon T A 2005 Metaphase chromosome preparation from cultured peripheral blood cells Current Protocols in Human Genetics 45(1) 4-1
[8] Yelmanov S, Hranovska O and Romanyshyn Y 2019 Image Enhancement Technique based on Power-Law Transformation of Cumulative Distribution Function In 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT) 120-124
[9] Somasundaram D and Kumar V V 2014 Separation of overlapped chromosomes and pairing of similar chromosomes for karyotyping analysis Measurement 48 274-281
[10] Wayalun P, Chomphuwiset P, Laopracha N and Wanchanthuek P 2012 Images Enhancement of G-band Chromosome Using histogram equalization, OTSU thresholding, morphological dilation and flood fill techniques In 2012 8th International Conference on Computing and Networking Technology (INC, ICCIS and ICMIC) 163-168
[11] Yang X, Wu W and Tseng C C 2012 Algorithms for modeling structural changes in human chromosomes Comput. Methods Programs Biomed. 110(2) 171–182
[12] Abbase S, Khtanlou H, Afrasiabi M and Asgari A 2015 Automatic Identification of Chromosomal abnormalities in Metaphase Karyotype Using Paired Images in Human Chromosomes in 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI)