ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 9, No. 4, 2014, *pp*. 243-251

Bands detection and Lanes segmentation in DNA Fingerprint images

I. Ismail, Gh. S. Eltaweel, H. Nassar
Department of Computer Science
Faculty of Computers & Informatics
Suez Canal University, Ismailia 51422, Egypt
Is-raa@hotmail.com, Ghada_eltawel@ci.suez.edu.eg,
(*Received February 10, 2012, accepted July 23, 2014*)

Abstract. Gel electrophoresis (GE) is a widely used technique to separate DNA sequence according to their size and weight, GE results are presented using images. In this paper, we present a scheme that aims to detect and segment the lanes in DNA gel images without any human interventions. We have successfully implemented an image processing scheme to automatically detect and segment the lanes in DNA gel images, applying this scheme to several DNA fingerprint gel images ,all the lanes are successfully segmented. Also we have obtained up to 99.5% accuracy for the segmentation of lanes in good quality images. The proposed scheme is compared with other techniques and the comparison shows that it has a minimum error rate.

Keywords: Gel electrophoresis (GE), matched filter, watershed segmentation algorithm**.**

1. Introduction

Gel Electrophoresis (GE) is a method used in clinical chemistry to separate proteins by charge and/or size also in \ble tool in many applications such as forensic studies, paternity analysis, protein profile comparison, gene identification, isolation, purification and population genetic analysis.

GE technique generates images that called DNA fingerprinting images which are an efficient and highly accurate means of identities and relationships. DNA fingerprinting images consist of several vertical *lanes*, each lane corresponding to one sample. Each lane contains a number of horizontal *bands*. Each band represents a part of the sample. The positions of the horizontal bands in the lane represent the molecular weights of that part of the sample. Two samples are considered to be the same if their lanes have the same pattern, Fig. 1 shows an example of a DNA fingerprint gel electrophoresis image.

Previous work regarding this problem can be found in [2]–[7], the semi-automatic lane detection method of Elder and Southern (ES) [2] is based on equispaced lanes with constant width, where the center of the first and the last lanes in the gel image are manually specified and the number of lanes between the first and the last lanes is given by the user.

Kaabouch, N. et al [3] proposed an algorithm that consists of main steps: automatic thresholding, shifting, filtering and data processing. They use the automatic thresholding to equalize the grey values of the gel electrophoresis image background.

Akbari A. et al [4] presented an effective noise filtering technique for DNA gel images. Lin, C.Y et al [5] designed a computer method to compare the lanes and identify the identical ones. This method segments the lanes and bands in the GE images. In order to describe the position of the bands, they introduce a position vector normalization technique. Then the compared lanes become equivalent to the position vectors. As a result, this method could accurately identify identical lanes. Cheng, W.Z. et al [6] presents a method that lanes in a GE image are first segmented and converted into a chain code representation. The lane comparison is performed by calculating the longest common subsequence (LCS) in two chain codes. Akbari, A.et al [7] present the (ES) semi-automatic lane detection method, iterative moving average filter (IMA) and continuous wavelet transform (CWT) followed by two new methods for lane separation.

Many factors affect the image quality and the patterns in the lanes, such as the applied voltage, field strength, pulse time, reorientation angle, agarose type, concentration, and buffer chamber temperature [8]. In electrophoresis, DNA or other charged molecules are forced to move through the maze formed by the polymers. The mobility is guided by two factors, the mass and the shape of the molecules. The smaller the

mass, the faster it moves. As the samples move farther away from the original starting point, the effects of the shape on the molecules start to appear and the bands become blurry.
In this paper, we present a scheme that detect and segment DNA fingerprint gel images using image processing techniques to automate routine analysis process of GE DNA images to help identifying humans.

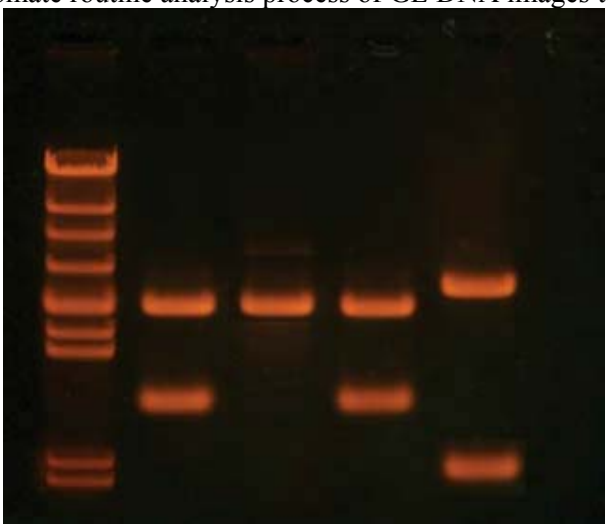

Fig.1 DNA fingerprint gel electrophoresis image.

Automatic tools speed up the routine of the biological processes. Much repetitive work in molecular biology can be in a computerized analysis form that is reproducible and avoids various forms of human error. Automatic techniques with an interactive check on the results speed up the analysis and reduce the error, and that's the motivation of this scheme.

2. Materials and Methods

The proposed scheme consists of two stages, the pre-processing stage starts by converting the image into a gray scale image, then image enhancement and background removal, The main purpose of the pre-processing stage is to enhance the image and make it sharper, The lanes detection and segmentation stage, starts by using the intensity profile to detect lanes and matched filter to enhance the bands' shape, then the watershed segmentation algorithm is applied which actually segments the lanes. Fig. 2 shows the block diagram for the proposed scheme.

2.1 Preprocessing

The pre-processing stage starts with converting the original RGB image into a grey scale image, as shown in fig.3 (a), and then image enhancement process is applied to improve the visual appearance of an image and converts the image to a form better suited for analysis by a human or a machine. Using un-sharp / local contrast stretching [9], because most of the un-sharp masking methods are not effective for low-contrast images. Therefore, after converting the image into gray scale, the contrasts of these images are stretched.

The main objective of using local contras stretching is to highlight faint bands that can be washed out because of their low gray levels. Local contrast stretching (LCS) is an enhancement method performed on an image for locally adjusting each picture element value to improve the visualization of structures in both darkest and lightest portions of the image at the same time. LCS is performed by sliding windows (called the kernel) across the image and adjusting the center element using the formula

$$D(x,y) = f((x,y) - min\,)/(max - min\,)) * N \qquad (1)$$

Where N is the number of intensity levels, "min" and "max" are the minimum intensity value and the maximum intensity value in the input image. For example, normally in the gray-level standard, the lowest possible intensity is 0, and the highest intensity value is 255. Thus N is equal to 255. After enhancing the DNA gel image by using LCS, apply the Un-sharp masking that yields to increase either sharpness or local contrast because these are both forms of increasing differences between values and increasing slope