PREPROCESSING OF 1D GEL ELECTROPHORESIS IMAGE

Matej Hlavatý

Bachelor Degree Programme (3), FEEC BUT

E-mail: xhlava35@stud.feec.vutbr.cz

Supervised by: Martin Vítek

E-mail: vitek@feec.vutbr.cz

Abstract: The aim of this project is an automatic analysis of 1D gel electrophoresis images. Basic principles of gel electrophoresis and types of errors and signal noise influencing the results are studied. An algorithm for automatic detection of lanes in grayscale electrophoresis images is proposed. After horizontal image segmentation, vertical summation projection profile of each segment is created. Local minima and lane borders of the profiles are detected after smoothing the profiles. Finally, points between neighbouring partition borders are interpolated and applied to the original image.

Keywords: Gel electrophoresis, lane detection, contrast enhancement, median filtering, function local extremes

1. INTRODUCTION

Gel electrophoresis is a widely used method of macromolecule fragment analysis and separation based on particle charge, weight and shape. DNA, RNA or protein fragments are spread on one end of agarose or polyacrylamide porose gel on which an electric current is applied. The fragments move through the gel pores, which resist their passage. As a result, smaller, highly charged particles travel further in the gel. The final position of the particles is visible after staining the gel as horizontal bands, organised into lanes representing each sample. A grayscale photography or scan of the gel is then acquired. [1]

To extract meaningful data from the results, lanes are detected and sorted based on their similarity. Without automated software, lab specialists are required to perform this task manually, which takes a significant amount of time for a mundane and, from a human perspective, simple process, especially in a lab processing a large amount of electrophoresis samples. The automated software, however, is prone to false detections in low quality images with irregular lane shapes. [2,3]

This project presents a method of lane detection based on image segmentation and detection of local minima in vertical summation profile. It's focused on near-full automatisation not depending on user input – the only requirement is a grayscale image with dark background and light bands.

2. METHOD

A common problem of electrophoresis images is a 'smile effect' – horizontal lane bending caused by excessive voltage or gel overheating. Detecting lane borders for the whole image would cause many defects and incorrect results. To prevent these mistakes, the image is horizontally divided into six segments, making the lane drift within one segment negligible. A contrast enhancement, summation profile and minima detection is therefore conducted for each segment separately; the resulting borders are aligned and connected afterwards, as seen in figure 1.

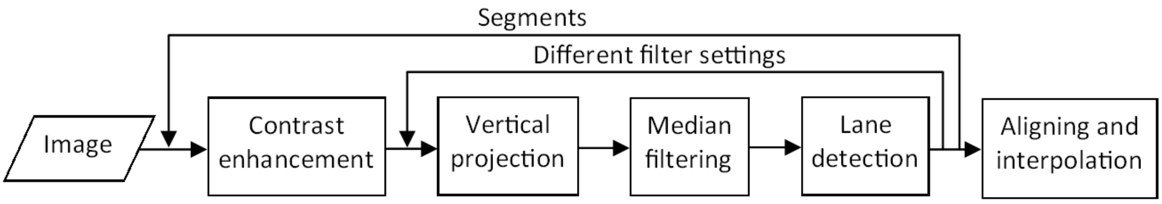

Figure 1: Block diagram

2.1. IMAGE SEGMENTATION AND LOCAL MINIMA DETECTION

Due to the quality of both image and electrophoresis results, intensity of pixels in a band is not consistent. This creates noise in the vertical profile and results in false lane border detection [4]. Therefore, contrast enhancement is used for each segment to improve the poor difference between data and background. On the pixel intensity histogram, the predominant dark grey background is visible as a highest peak. By setting all pixels darker than this peak value to black and evenly spacing the remaining values over the whole grayscale spectrum, the difference between hardly visible bands and the background is easier to detect. As the enhancement is done for each segment separately, the slight difference between bright segments - usually in the upper part of the lanes – remains, while visibility of the dark bands near the bottom of the lanes is also improved.

Median filtering of the profile with a correct window removes small irregularities in the lanes; a correct window size is difficult to determine, however. A successful filtering would result in evenly spaced borders – the brightest segment profile is therefore filtered using a variety of filter windows, between 8-30 pixels, and for each result a standard deviation of the border distances is calculated. The result with the lowest deviation, thus with the most even border spacing, determines the filter window size and average lane width estimate used for other segments.

The basis of lane detection is creation of a vertical summation profile (fig.2), showing locations of lanes as peaks on the horizontal axis [5]. In this profile, borders between lanes are detected as local minima using a profile derivative. The end of one lane and beginning of the following one are recorded separately, thus a missing lane doesn't influence the results.

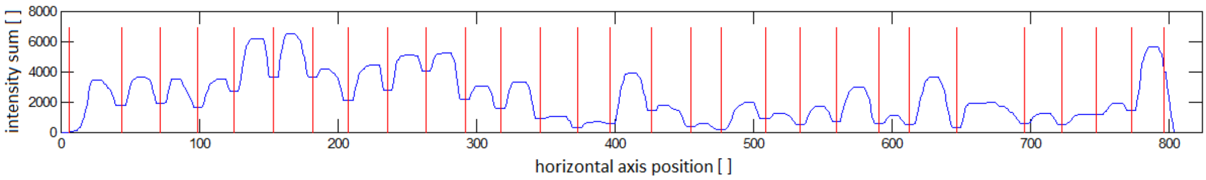

Figure 2: Vertical summation profile with detected lane borders

2.2. BORDER ALIGNMENT AND INTERPOLATION

Even with the best filter, local minima detection is not always perfect. Empty lanes, variable lane width, or interlacing adjacent bands may result in unevenly spaced border profiles. In order to correct these mistakes, each segment is searched for irregularities: in case the distance between borders is below 0,7*lane width a border is deleted, in case it is between 1,6*lane width and 2,2* lane width a new border is added by interpolating between the closest borders.

Before connecting the segments together, the borders of neighbouring segments are aligned, filling the possible differences. Values between the segments are then interpolated, creating a set of borders for each row of pixels of the image (fig. 3, left).

As the last step, lanes from the original image are extracted. For each row of the lane a median value is calculated and used as a new value for the whole row. The lane is displayed with a constant width, thus creating a clear representation of the lane, ready for comparison or analysis, without sacrificing any important data contained in the original image (fig. 3, right).

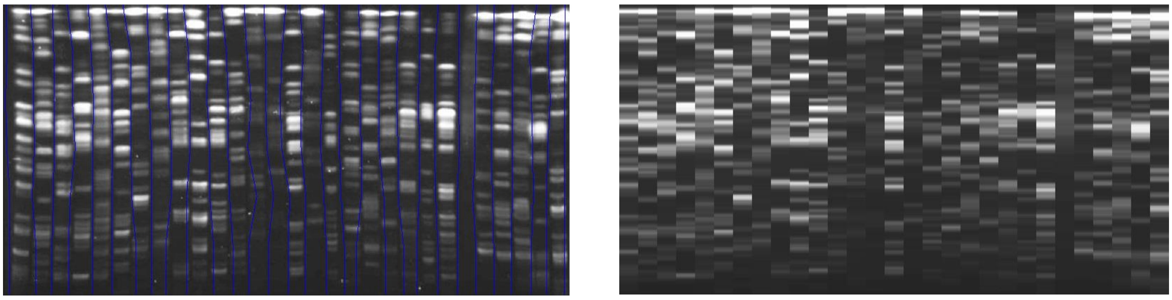

Figure 3: Original image with detected borders (left) and extracted lanes (right)

3. CONCLUSION

In this paper a method of lane detection in 1D gel electrophoresis images is presented. Automated pre-processing of electrophoresis results is crucial to effectively process a large amount of samples. Various contributions [3,6,7] have been presented before with a similar approach.

Common obstacles of lane detection include low image quality and lane distortion. The method approaches them using contrast enhancement to decrease background influence, image segmentation to adapt to lane drift, and median filtering to prevent false lane border detection. The resulting borders are applied to the original image and median value of each row is extracted.

The method shows 98% success rate in images with regularly spaced lanes. However, as median filtration and detection of irregularities depend heavily on lane width estimate, error in distorted images with changing line width increases rapidly. Another important factor is image quality and influence of background. Better methods of suppressing the background without smoothing the useful data together could improve the result.

Future work may also include lane comparison based on the final extracted lanes, starting with detection of empty lanes and false data. Lanes could then be sorted into groups, based on their mutual similarities.

REFERENCES

[1] J. Pizzonia, "Electrophoresis Gel Image Processing and Analysis Using the KODAK 1D Software," BioTechniques, vol. 30, pp. 1316-1320, 2001.

[2] A. Akbari, et.al. "Automatic lane detection and separation in one dimensional gel images using continuous wavelet transform," Anal. Methods, vol. 2, pp. 1360-1371, 2010.

[3] R. Wong, S. Flibotte, R. Corbett, P. Saeedi, Steven J. M. Jones, Marco A. Marra, Jacqueline E. Schein, and I. Birol, "LaneRuler: Automated Lane Tracking for DNA Electrophoresis Gel Images," IEEE Trans. Autom. Sci. Eng., vol. 7, No. 3, pp. 706-708, 2010.

[4] V. Sousa, et. al. "Automatic Lane and Band Detection in Images of Thin Layer Chromatography," A. Campilho, M. Kamel (Eds.): ICIAR 2004, LNCS 3212, pp. 158–165, 2004. Springer-Verlag Berlin Heidelberg, 2004.

[5] Park, et. al. "Lanes Detection in PCR Gel Electrophoresis Images," 11th IEEE International Conference on Computer and Information Technology, 2011.

[6] Park, et. al. "Lane detection and tracking in PCR gel electrophoresis images," Computers and Electronics in Agriculture, vol. 83, pp. 85–91, 2012.

[7] Skutkova, et.al. "Preprocessing and Classification of Electrophoresis Gel Images Using Dynamic Time Warping," International Journal of Electrochemical Science, vol. 8, pp. 1609-1622, 2013.