Int. J. Electrochem. Sci., 8 **(2013)** 1609 - 1622

International Journal of
ELECTROCHEMICAL
SCIENCE
www.electrochemsci.org

Preprocessing and Classification of Electrophoresis Gel Images Using Dynamic Time Warping

Helena Skutkova[1], *Martin Vitek*[1,3], *Sona Krizkova*[2], *Rene Kizek*[2], *Ivo Provaznik*[1,3*]

[1] Department of Biomedical Engineering, Faculty of Electrical Engineering and Communication, Brno University of Technology, Kolejni 2906/4, CZ-616 00 Brno, Czech Republic
[2] Department of Chemistry and Biochemistry, Faculty of Agronomy, Mendel University in Brno, Zemedelska 1, CZ-613 00 Brno, Czech Republic
[3] International Clinical Research Center – Center of Biomedical Engineering, St. Anne's University Hospital Brno, Brno, Czech Republic
[*]E-mail: provaznik@feec.vutbr.cz

Received: 31 July 2012 / *Accepted:* 30 October 2012 / *Published:* 1 February 2013

The automation of a classification process of electrophoresis gel images is a difficult task. The result highly depends on quality of gel image digitization and on imprecisions in an electrophoretic process. The methodology proposed in the paper helps to remove most of gel image distortions and effectively overcomes the problem of non-uniform electrophoretic process.

Keywords: gel electrophoresis, cluster analysis, taxonomy, dynamic time warping, lane detection, geometric distortion, contrast adjustment, median filtering

1. INTRODUCTION

The similarity analysis with subsequent classification of electrophoretic samples is a common way to evaluate the results of one–dimensional gel electrophoresis. Taxonomy study of microbial organisms [1,2] is one of the typical applications. However, the implementation possibilities are much wider, for example fingerprint analyses of tumour tissues[3]. The basic principle of evaluation of mutual similarity by dendrogram construction can be used e.g. for comparison of DGGE (*denaturing gradient gel electrophoresis)* profiles of PCR fragments [4-8]. The resulting dendrogram can serve as an objective indicator of bacterial phylogeny [9,10], then it is called a phylogenetic tree [11]. The phylogenetic analysis of electrophoretic gel samples is usually realized manually. The reason lies

mostly in a low quality of gel images that makes precise computer processing impossible. The mutual similarity of electrophoresis lanes is often evaluated subjectively in dependence on the number of identical and different bands [6]. The automation of gel images classification requires removal of most significant image distortions such as noise, non-uniform brightness and/or contrast, and geometric disproportion. Most serious distortions are caused by imperfect procedures in the electrophoresis. Suppression or removal of gel image distortions is a known problem that can be solved by a variety of commercial software products [12-14]. Their main disadvantage is a poor adaptation to a specific electrophoretic device and their focus on a specific application.

The approach proposed in this paper combines selected tools of image and signal processing to achieve maximum possible automation in processing of digitized gel images regardless of the specific application. The basic mechanism of gel image processing usually consists of three basic steps: 1. band detection, 2. band matching, and 3. quantification and comparison [15]. These three steps are highly dependent on an electrophoretic method. Their unified automation for different methods of gel electrophoresis (protein, PCR DNA/RNA fragment, DGGE, TGGE, etc.) is difficult. Information about band positions alone does not provide sufficient description of intensity values, width of bands or vertical shift between bands. These properties have individual character for each image and make image comparison difficult. The individual character is given by non-uniformity of electrophoretic separation. The most common problems are geometric distortion (e.g. smile effect, sample hyper-focusing), sample contamination (sample background), parasitic gradient (lane flexion), poor gel image exposure (low contrast), etc. [12,16,17]. The most of these image distortions are caused by imperfect operations and wrong settings of electrophoresis [18]. However, automatic electrophoretic image analysis should take this option into account. The image distortion influences lane detection (samples do not move strictly straight), band detection (higher background level than the level of bands), evaluation of the speed of various samples (the sample from marginal well moves through the gel slowly than the middle one). We propose gel image analysis without band detection that may cause many problems and the resulting positions of bands do not provide sufficient information about the sample. Representation of each complete sample in a form of a time series (1-D signals) is potentially suitable for more detailed analysis [14]. The signal timeline corresponds to the vertical axis of the sample migration in a gel and the magnitude of the signal is determined by a gray level of the image at individual positions along the vertical axis. Thus, the band matching problem is converted to signal matching. The different positions of bands (signal peaks) can be usually adjusted by resampling. In case of nonlinear signal migration (e.g. “smile” distortion), it can be adjusted by dynamic time warping. The transformation of the gel lane to the signal representation requires precise detection of the lane shape, especially if the lane is geometrically distorted. The paper proposes a new algorithm based on innovative lane detection method in combination with signal processing methods (as dynamic time warping, peaks detection, signal filtration etc.). It allows more precise, automatic and universal gel image analysis. Moreover, it is suitable for simultaneous comparison of several gel images with different parameters.

2. METHODOLOGY

The proposed electrophoretic gel image processing for evaluation of electrophoretic samples similarity consists of five consecutive steps. The gel image is often captured by a low-quality camera. Thus, the first step of the gel image processing is the basic image quality improvement . This step is included in the proposed methodology for two reasons: 1. to ensure the complete automation of the image processing, and 2. to ensure high contrast of grayscale images. The segmentation of the gel image on sample lanes is performed as the second step. The conversion of the gel lane image to the 1-D signal is inseparably related with this step. The signal representation of the gel lane is obtained by median and it can be understood as the non-linear sample filtering. Filtering of samples enhances the useful information contained in the sample (the ratio between the bands and the noise levels of sample background). The sample filtering focuses on each sample separately and eliminates local defects. In addition to the step of median filtering, a special filtering of the sample background was designed in order to remove impurities causing unevenly gray background. Before the last step of cluster analysis used to evaluate the similarity of the samples, the innovative step based on dynamic time warping is integrated. The dynamic time warping adjusts positions of similar bands between two samples, which can be unmatched due to uneven movement of the various samples in the gel.

2.1. Image quality improvement

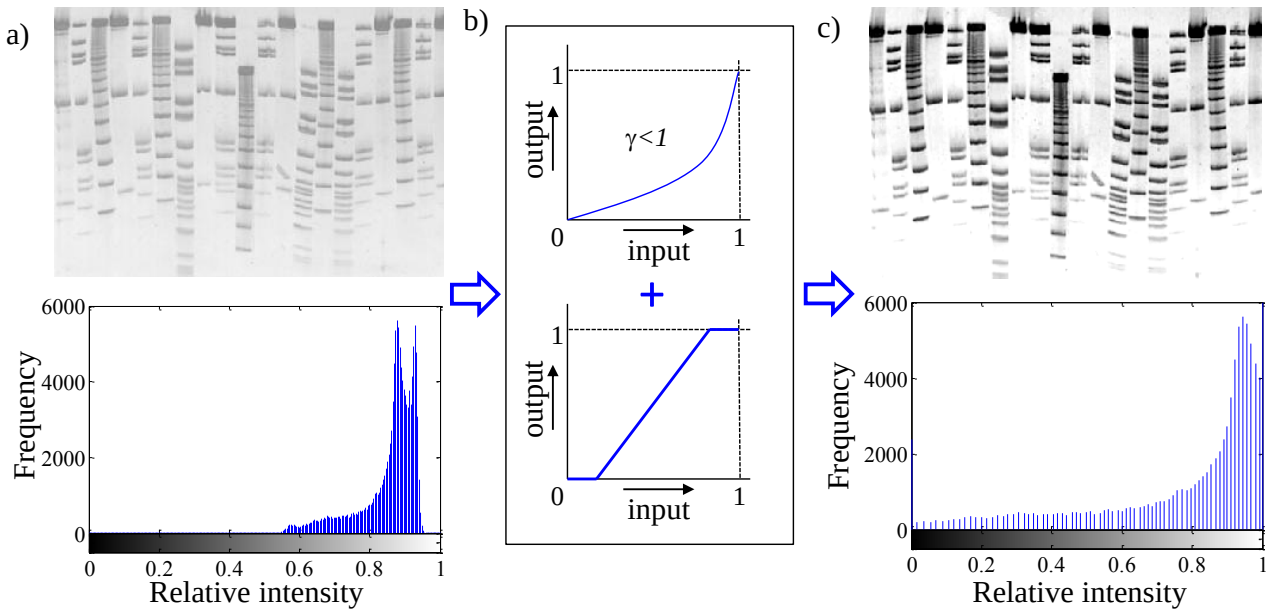

Figure 1. Enhancement of image contrast using a combination of two transformation functions. a) The original electrophoretic gel image (upper) and the grayscale histogram of the original image (lower). b) The transformation function of gamma correction γ=0.65 (upper) for modification of intensity values and the piece-wise linear function to emphasize marginal values (lower). c) The electrophoretic gel image with enhanced contrast and its grayscale histogram. Image from www.shimadzu-biotech.net.

All the other processing steps of gel image analysis require the grayscale image with the white background and the grey level of bands reaching black in the limit. The gray scale can be normalized in the range from 0 to 1 (0 value corresponds to black, 1 corresponds to white) for subsequent computational processing. The grayscale image histogram expresses frequency of the pixels with particular intensity level in relative grayscale range from 0 to 1. The grayscale histogram of the gel image should contain pixels on all levels of the relative grayscale range. A significant increase of the number of pixels is in range from 0.55 to 1, i.e. increased frequency of pixels of white color representing white background. An overexposed gel image has the histogram characteristic tend to 1 (white level), underexposed to 0. The contrast enhancement is realized by transformation function applied to grayscale range of the gel image histogram.

Another useful image enhancement method incorporates gamma correction. The gamma parameter is chosen in dependance on exposition of the image. The overexposed images should be corrected with $\gamma < 1$, underexposed image with $\gamma > 1$. The range of contrast enhancement for piece-wise linear contrast adjustment is set according to a level of noise in background. Combination of contrast transformations with gamma correction and piece-wise linear contrast adjustment allows the extension of histogram spectrum to entire grayscale range, highlights bands and suppresses background impurities. Figure 1 shows the example of contrast adjustment of an overexposed gel image. The contrast enhancement of the image in panel a) was realized by transformation functions in the panel b). Gamma parameter was experimentally chosen as $\gamma=0.65$ and linearly increasing trend of piece-wise linear function was in range between 0.15 and 0.85 of grayscale [19,20].

The choice of parameters of the transformation functions should be manually corrected, because this step is particularly useful to improve the appearance of the result gel image, which is a subjective criterion. The contrast enhancement is not necessary for the similarity analysis of the electrophoretic gel image; the result is changed only slightly and the taxonomy classification only in exceptional cases for very similar samples. This step can be completely skipped or the parameters of the transformation function can be configured based on a density of the histogram for fully automated analysis.

2.2 Automatic detection of lanes

The next step of the electrophoresis image processing is the automatic detection of lanes. Until now, there has been presented a number of approaches to handle the lane detection problem, such as [14,21-23]. These methods are based on a principle called “lane tracking” which lies in tracking of the centre of each lane. The disadvantage of such an approach lies in ignoring of the information from the area outside the central line. In this work, we present a new alternative method for automatic lane detection based on a tracking of lane borders. The result of such an approach is the full segmentation of the lanes. Thus, we can use the information from all pixels of each lane.

The proposed method can be divided into two main parts: 1) detection of the first pixel of each lane border, and 2) tracking of the border through other pixels. The algorithm requires single input value that is the number of lanes of the analysed electrophoretic image. Depending on this value are set

the other parameters of the algorithm. The rest of the segmentation process is fully automatic and requires no further human intervention.

The principle of detection of the first pixel of the lane border is shown in Figure 2a). In the first step, the intensity mean value is calculated for each column of pixels from the upper third of the image. The reason for using only upper third of the image is suppression of the lane shifts in x-axis direction which is greater in the middle and the lower third of the image. The peaks in the obtained 1D signal representation mark the position of the first pixel of each border. The used peak detection algorithm is based on finding local maxima in the signal. It returns only peaks with indices separated by more than the value of variable *mpd* (minimum peak distance). The variable *mpd* is calculated as the mean width of the lane multiplied by the constant 0.7. The necessary mean width of the lanes can be easily estimated as the length of the signal divided by the provided number of the lanes in the image.

The next step lies in tracking of the other pixels of each border line. The tracking algorithm starts from the already known positions of the first pixels. These pixels correspond with the uppermost points of blue lines in Figure 2b). In each tracking step, the algorithm compares the intensity values of the three pixels below the current pixel and selects the pixel with the highest intensity (the whitest one). In order to prioritize the most direct path as possible, the value of the middle pixel is a slightly increased by multiplicative constant 1.04. The tracking algorithm stops when reaching the bottom of the gel. The final result of the lane segmentation process is shown in Figure 2b) where the blue lines represent the found lanes boundaries. (The algorithm constants 0.7 and 1.04 were set to achieve best results on the training set of 50 gel images, see example in Figure 7.)

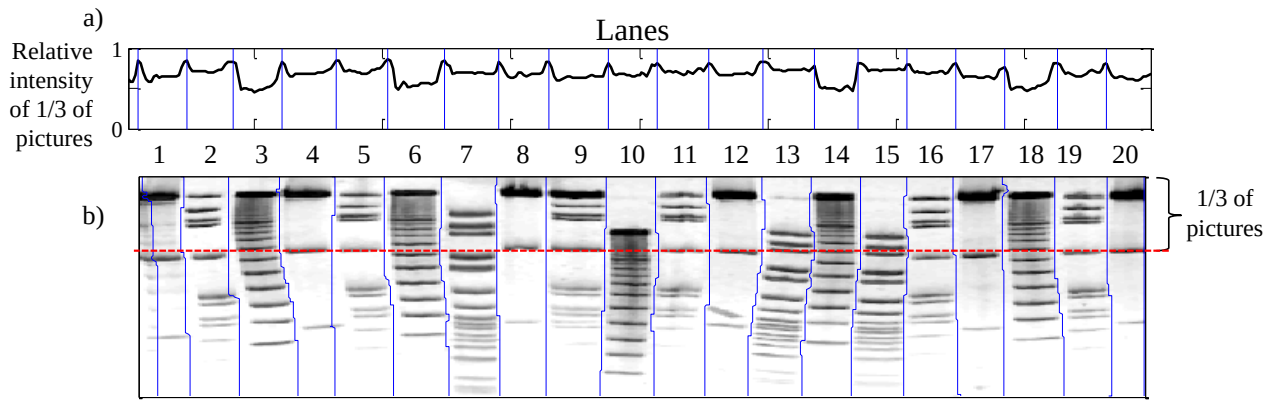

Figure 2. Principle of gel image segmentation using tracking of lane boundaries. a) The detected positions of intensity peaks (blue lines) correspond to the positions of the first pixel of each boarder. b) The final segmented image with marked boundaries of each lane (blue lines).

2.3 Conversion of 2D grayscale image to 1D signal representation

The digitization of 2D grayscale profiles of electrophoretic gel samples to 1D signal representation is the next necessary step of computational analysis of electrophoretic gel used in most applications [8,12,17,22,24]. We can effectively emphasize the important features of the electrophoresis image and suppress some types of distortions with the known boundaries of each lane

in the image. The important features necessary for subsequent classification are position, thickness and intensity of the bands in each lane. Shifts in the direction of the x-axis and the background noise are the types of image distortion which can be suppressed. In the case of the absence of the shift, the detected boundary lines would be straight. In the case of the presence of the shift, the proposed boundary detection algorithm can trace it. Both emphasize of the features and disorders suppression lies in conversion of 2D lanes to 1D signal representations. The pixels with the same y-axis coordinates within the range of each lane boundaries are replaced by their median value. The possible shift in x-axis direction is eliminated and median also effectively suppresses the image noise. The result of such a conversion can be seen in Figure 3. The figure shows original lane before conversion, the final 1D signal representation and 2D median lane constructed by simple repeating of the 1D signal.

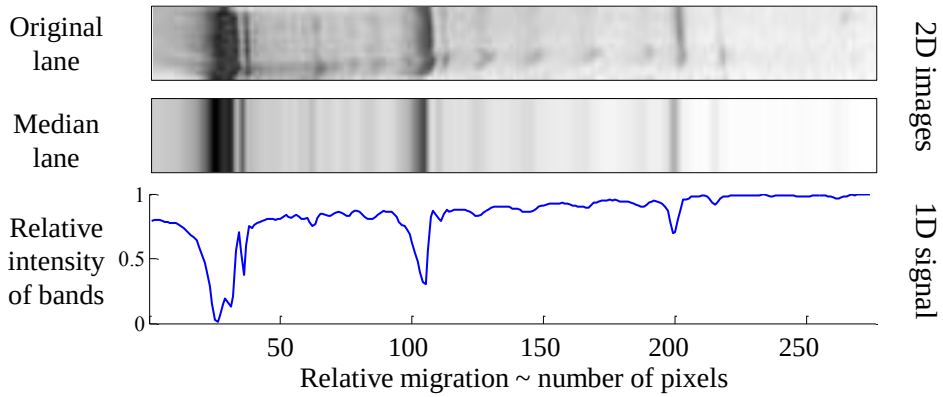

Figure 3. Conversion of a lane image to signal representation using median. a) The original 2D lane after segmentation. b) The 2D lane obtained by median filtering of the original 2D lane. c) Final 1D signal representation of the lane obtained from median values of the original 2D lane.

2.4 Shading correction

Impurities in the gel sample often cause unevenly distributed gray background. The level of gray in the background can be very close to the gray value of some bands. Moreover, the background of impurities is specific for each sample but not for sample characteristics, and then the consecutive similarity analysis differentiates two identical samples only on the basis of different impurities. This problem can be solved for example by localization of particular bands using peak detection technique [15,17,23,25]. However, the proposed method compares whole courses of each gel sample, not only bands positions of samples. A simple filtering technique consists of three steps: the estimation of the background envelope in signal representation (Figure 4 b, upper part), subtraction of background trend from sample course and normalization of signal values in range <0, 1> (Figure 4 b, lower part).

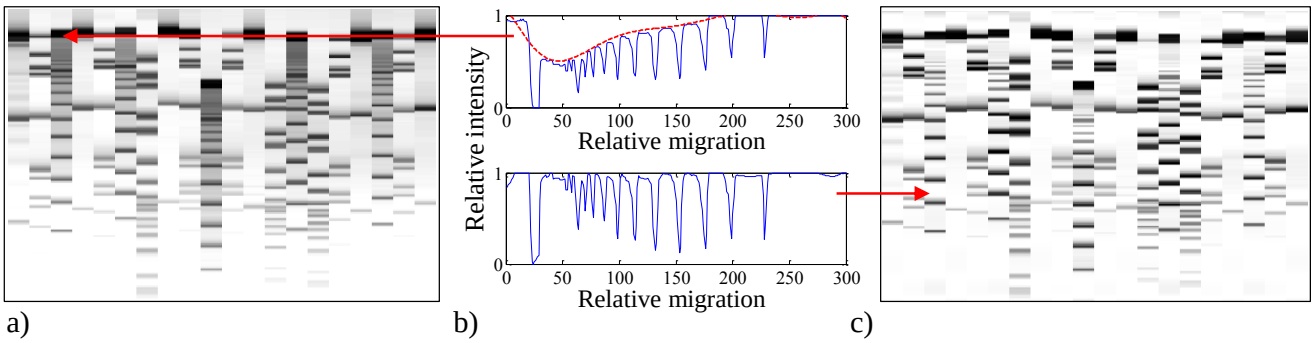

Figure 4. Correction of shading in electrophoretic lanes. a) The gel image after lane detection and median filtering. b) Upper: The signal of the lane 3 (blue line) with estimation of the signal envelope (red line). Lower: The signal of the lane 3 with removed envelope and normalized values. c) The gel image restored from corrected lane signals.

The estimation of the trend of gray background is realized as finding the signal maxima in a floating window. The window size L_W depends on the image resolution and width of the largest band, that the envelope of the background does not follow the course of a signal. The proper window size is twice the size of the largest band and the steps of the floating window with overlap L_o equal one quarter of the window size L_W. The window overlap requires the extension of each signal by one length of the window on both sides. The values of extension are the first and last L_W values of signal repeated at the beginning and the end. The background envelope estimation is performed separately for each signal (sample). The window parameters were set $L_W = 20$ px and $L_o = 5$ px for the testing image shown in Figure 4 with the size of 550 × 300 px.

2.5 Adjustment of mutual positions of bands using DTW

The uneven speed of movement of the samples on the gel is caused by many factors. These are usually too high electrophoretic voltage, uneven density of the gel or sample application error [18]. The increased speed of samples movement in the middle wells compared to marginal wells (smile effect) is a type of distortion which can be modelled and thus compensated [12,13,15-17,26]. The combination of more factors causing geometric distortions makes the modelling difficult. In this case, the sample speed does not change equally, i.e. the middle sample lane is not only shifted to marginal, but also unevenly stretched. Compensation of such a distortion requires individual adjustment of each sample. Methods based on detection of bands positions [12,15,23,25] correct positions to correspond to the ladder scale. However, determination of the exact position of each band is complicated. The use of the signal course instead of band positions maintains more information about sample. The correction of uneven sample movement can be realized by local adjustment of the signal length. If we consider the signal representation of the sample lane as the time course, the dynamic time warping can be used for this local adjustment. The length correction is performed pairwise (all samples with each other), so one incorrect adjustment has less influence on the final result of the similarity analysis. The one pairwise adjustment affects the mutual similarity of these two samples, but a sufficient number of other samples can compensate this inaccuracy.

The technique called dynamic time warping is was originally used in speech analysis [27]. The same spoken word in the speech of different people has the same meaning (signals are of the nearly same shape), but its timing and offset is specific for each person. The method of dynamic time warping can adapt the timing and offset of such signals [28]. This attribute can be advantageously used for adjustment of bands positions between two gel lanes.

The principle of the DTW is well known in signal processing community [28]. The principle of signals alignment using DTW is shown in the Figure 5. The technique aligns samples values using minimization of the distance between the pairs of samples. Stretching of one or both signals is realized by repeating the selected samples. The criterion for alignment and repetition of samples is determined by the table of accumulated distances (Figure 5 b). The value of accumulated distance is calculated from pairwise distance for each pair of samples in accordance with (1).

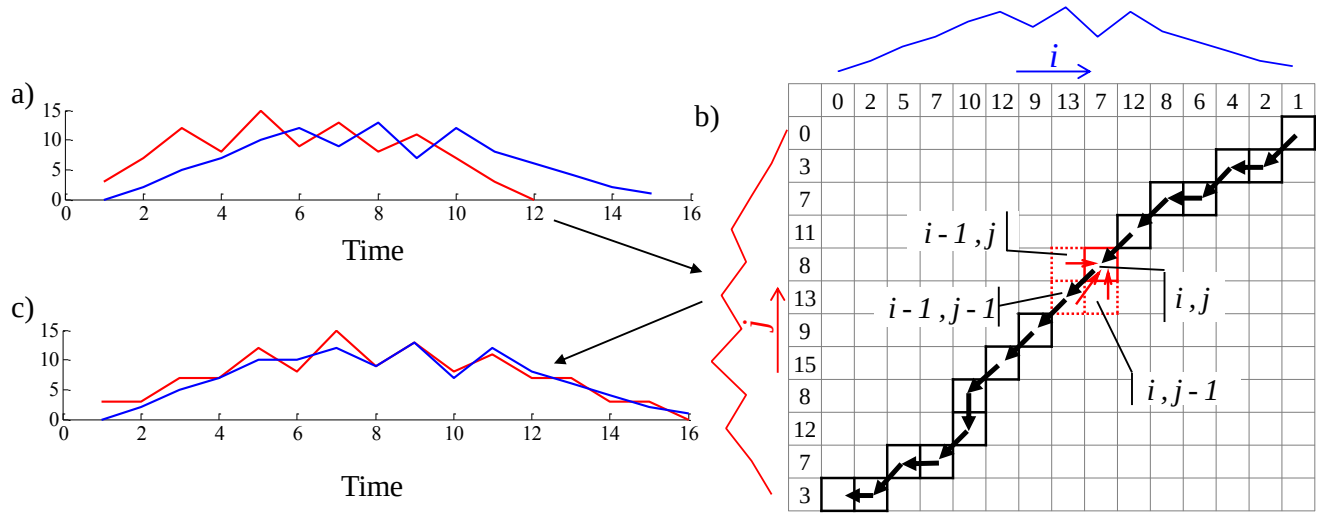

Figure 5. General principle of DTW. a) Two original signals of different lengths. b) Determination of the matrix of accumulated distances and optimal path backtracking. Numbers in the first column represent values of the first signal and numbers in the first row represent values of the second signal. The black arrows compose optimal path used for the signal adaptation. The red arrows show the possible directions in each step. c) The resulting signals after DTW (same lengths).

$$D(i, j) = \min[D(i-1, j-1), D(i-1, j), D(i, j-1)] + d(i, j) \quad (1)$$

where D is an accumulated distance and d is a value of a pairwise distance. The value of the accumulated distance $D(i,j)$ is determined by the pairwise distance $d(i,j)$ and minimum from the previous values of accumulated distances as shown in Figure 5b). This set of accumulated distances for each pair of samples forms a matrix. The result sequence warping is derived on the basis of minimization of the backward way from the right upper corner to the left lower corner of the matrix. There are three allowed directions for the each path step, see Figure 5b) (red arrows). The original signals from Figure 5a) are aligned to resulting signals shown in Figure 5c).

The DTW adapts positions of the same bands but the specificity of each lane is maintained. The principle of adjustment using DTW is shown in the Figure 6. The Figure 6a) shows two similar lanes with different migration speed in image and signal representation. Traditional similarity analysis evaluates only a small amount of mutual similarity in this case. The application of DTW on a signal form of these two signals in Figure 6b) modifies lengths of signals for positional adjustment. The adapted signals were converted again to the image form.

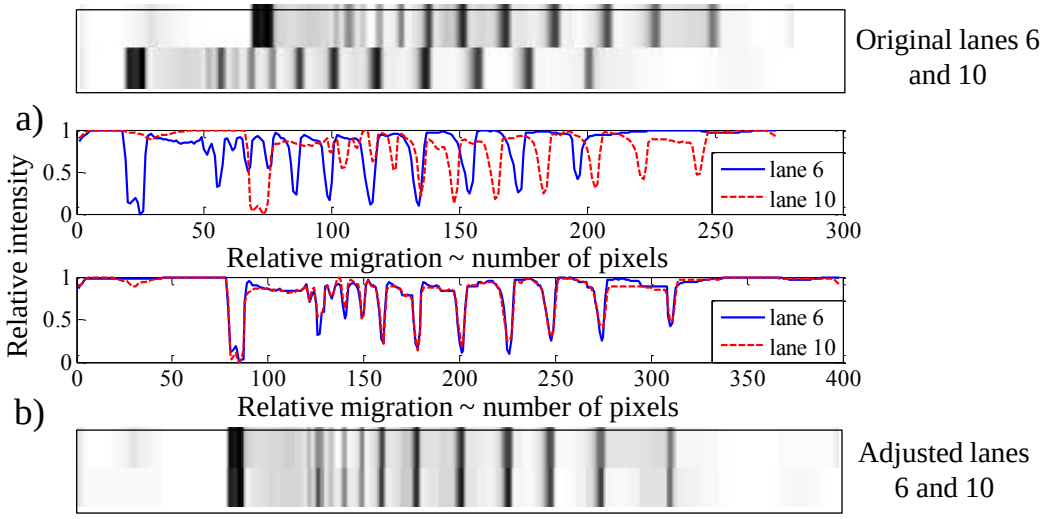

Figure 6. Adjustment of band positions in a pair of lanes using DTW. a) Lanes containing similar bands which are not aligned (upper) and signal representation of the lanes (lower). b) Positional adjustment of signal peaks of the lanes (upper) and corresponding image representation of adjusted lanes (lower).

2.6 Similarity analysis of gel lanes

The result of the similarity analysis can be represented by a dendrogram. For gel electrophoresis samples, dendrogram is called a phylogenetic tree [9-11]. The obtained information does not describe only similarity, but also affinity. UPGMA method was used for cluster analysis. Although the method is simple, it is still the most widely used [4-6]. Moreover, testing of cluster analysis techniques is not the subject of this article. The main goal is the extension of common methods by implementing the DTW, new line detection algorithm and the gel image improvement. The dendrogram is reconstructed from pairwise distances determined by (2) as a simple Euclidean distance d_{xy} of two signals X and Y with length k.

$$d_{xy} = \sqrt{\sum_{i=1}^{k} (X_i - Y_i)^2} \qquad (2)$$

The distance matrix was reconstructed from pairwise distances of each pair of signals (sample lanes). Several methods use the Pearson correlation coefficient for distance measure [4,12]. The correlation coefficient can compensate the sample offset caused by unequal moving sample on the gel. The correlation can be used only for linear positional adjustment; the unequal stretching of the samples

is not solved. The application of DTW in previous processing step solves both the problems together; therefore the use of correlation coefficient is unnecessary.

If the similarity of samples depends on the mutual offset between two sample lanes, the length of backward way in DTW can be used for weighting of Euclidean distances.

3. RESULTS AND DISCUSSION

The proposed methodology is composed of several innovative steps: the new method of lane detection, conversion of a sample lane to 1D signal and its processing without band detection and application of DTW for signal adjustment. The awareness of their impact to the result is important for their appropriate utilization.

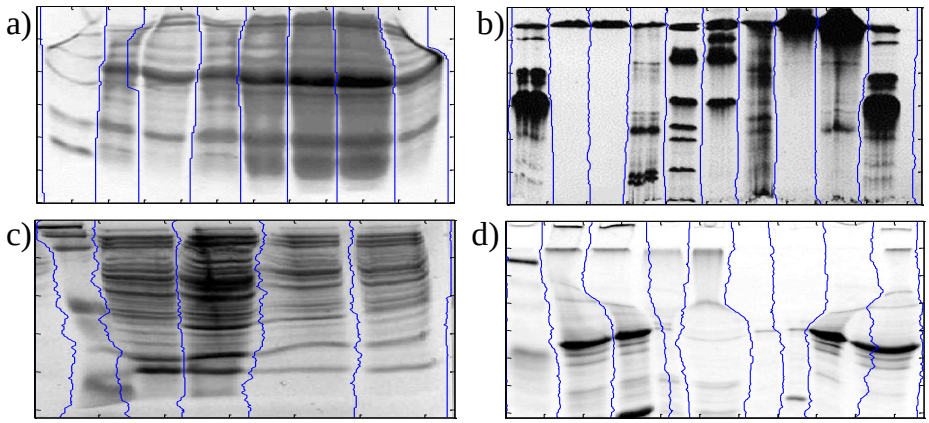

Figure 7. Examples of correct lane detections in four gel images with problematic lane boundaries because lanes overlap each other. a) distortion of protein pattern due to the delay between loading the samples and running the gel, b) bands smearing caused by proteins precipitation during the electrophoresis and uneven proteins concentration in the wells, c) distortion of protein pattern due to delay between samples loading and gel running and high salt concentration in the samples, d) distortion of protein pattern caused by high salts content in the samples, image from http://nu-distance.unl.edu/homer/class/4/mastery/text/geltips.html.

The first innovative step is the segmentation of the gel image on individual sample lanes. The incorrect identification of the sample lanes trend leads to total degradation of the similarity analysis result. Moreover, this step is problematic and greatly influenced by any image distortion. The conventional detection approach is more susceptible to errors because it detects only middle line of a sample lane. Our approach detects the boundaries of each lane and estimates true values as median values between the lane boundaries. The algorithm was tested on 50 images obtained from student laboratories and from website which discusses the troubleshooting in gel electrophoresis (http://nu-distance.unl.edu/homer/class/4/mastery/text/geltips.html). The Figure 7 shows four examples of problematic lane detections. These four gel images are visibly distorted, lanes overlap each other and such a complicated type of distortion can be hardly modeled. The proposed method of lanes detection reliably detects borders of each lane. The median filtering compensates small overlapping of border

lines with sample lanes and corrects some types of image noise and distortions. The detection works even if the sample lane does not contain any bands (Figure 7d: 6th and 7th lane) and traditional algorithms of sample middle line detection has nothing to detect.

The testing of influence of image and signal pre-processing in combination with DTW was performed on a gel image shown in Figure 1a). The gel image contains four groups of samples which differ in the number of bands and their positions. The division of samples into these groups is color coded in Figure 8. The similarity within groups of samples is relatively clear; the difficult samples to determine are especially the 7th and 10th sample. These sample lanes were degraded by offset (10th sample) and offset with stretching (7th sample). The first group (No. I) is significantly different from the others, because these samples consist of only 3 bands, while the other groups have samples with much more bands. The separation of the first group should be clear from others. The diversification of samples from the remaining three groups is complicated and every small of band positions can cause an error. The criteria for the correct result of similarity analysis are the division into four clusters, the separation of the first group, the classification of distorted samples No. 7 and No. 10 and the threshold of clusters separation. The last criterion takes into account the classification sensitivity; the higher threshold value means worse classification of very similar groups. We set three parameters for an objective evaluation of classification accuracy: *th* – threshold for clusters separation, *root* – distance between root and terminal branches, *sens* – classification sensitivity as a ratio *root* : *th*.

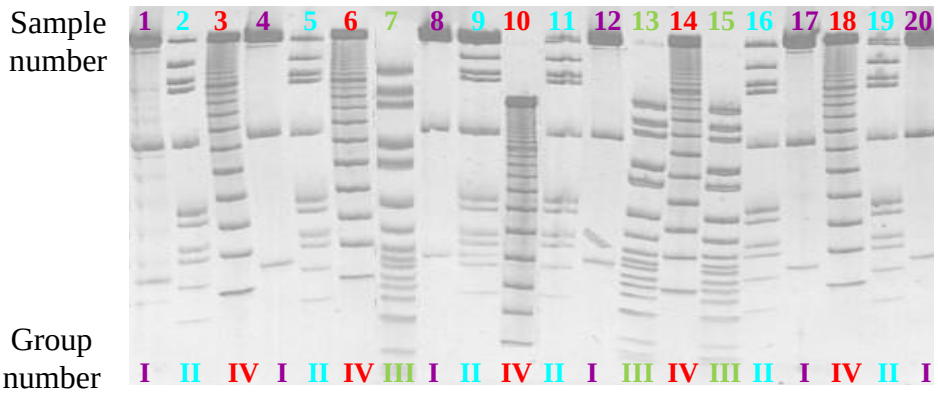

Figure 8. The original gel image with colour coded numbers of samples and groups by mutual similarity.

The cluster analysis was applied on four different modifications of the electrophoretic gel image from Figure 8. The first step using lane detection with median filtering and 2D to 1D transformation is identical for all variants. Further, the ability of classification was tested in dependence on the use of DTW and combination of proposed image adjustment techniques. The first result in **Error! Reference source not found.**a) shows classification of gel samples with the use of filtering and DTW utilization. The division of samples to four clusters was performed correctly. The group No. I forms a separate cluster. The second main cluster is composed from three subclusters represented by groups II, III and IV. The distortion of samples No. 7 and No. 10 has slightly affected length of branches inside clusters, but their classification remains correct. Almost the same clustering

result was achieved without any adjustment technique in Figure 9c) with DTW utilization. The question is whether to use the image adjustment in combination with DTW or not. The answer lies in mentioned parameters: the clustering threshold and sensitivity. The sensitivity of classification technique with filtration and DTW is *sens = 8.04 : 1.52 ≈ 5 : 1* and without filtration *sens = 3.02 : 0.68 ≈ 4 : 1*. Although the sensitivity of classification without the image filtering is still high, the occurrence of errors will be more possible in the case of more similar sample groups. The proposed filtering technique does not change the result; it only increases the clustering sensitivity.

The remaining dendrograms with reorganized sample lanes in Figure 9b) and 9d) demonstrate the result of classification without using DTW. The cluster analysis alone is not sufficient to classify the damaged samples 7 and 10, even with the help of the image filtering. The internal clusters division is inconsistent, the significantly different group No. I is incorrectly linked with others. Moreover, the sensitivity of clusters separation is significantly below ratio 2:1. The DTW utilization appears more important than the image filtration.

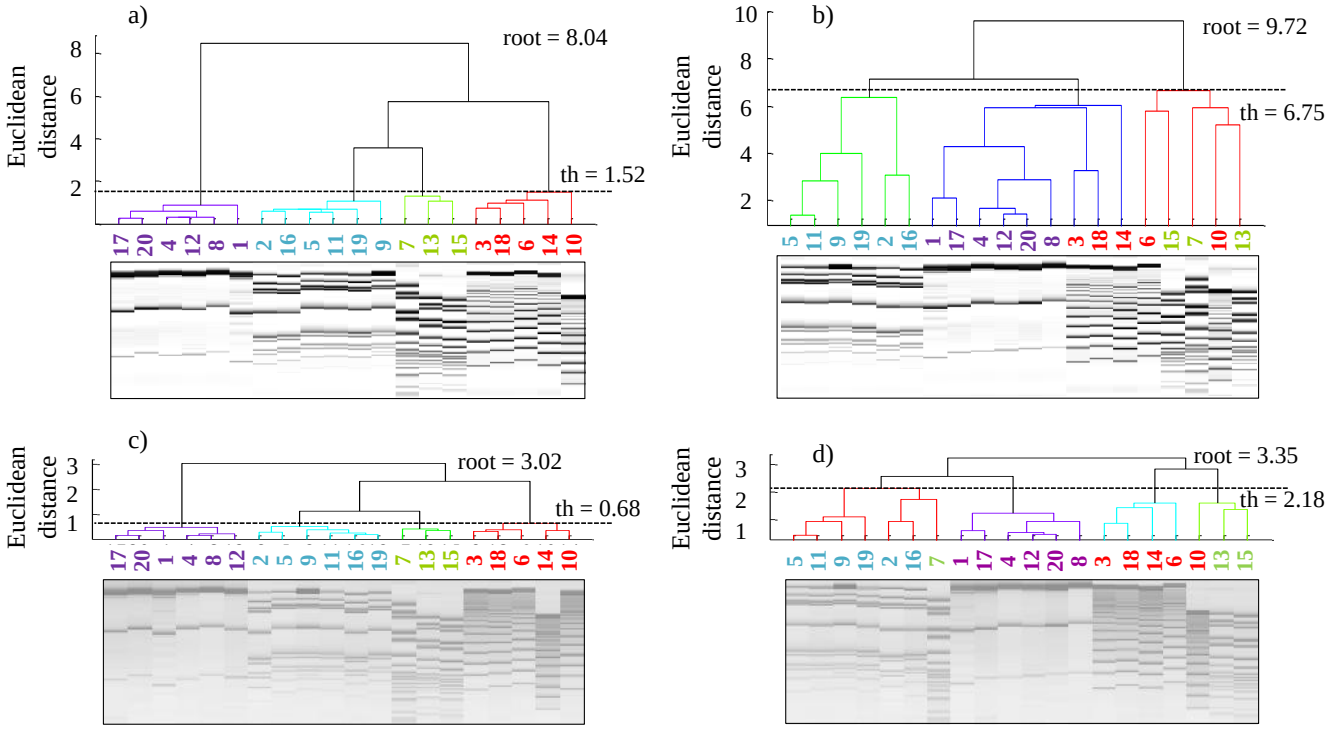

Figure 9. Influence of image and signal pre-processing and DTW utilization on result of sample classification. Classification of gel samples a) with pre-processing and DTW, b) with pre-processing and without DTW, c) without pre-processing and with DTW, d) without pre-processing and DTW. Distance values: th – threshold for clusters separation, root – the length of the tree.

4. CONCLUSIONS

The computer processing of one dimensional gel electrophoresis may appear outdated compared to later methods such as the capillary, chip or 2D electrophoresis. However, it is still the

most affordable technology for most (bio)chemical laboratories because of its low purchase and operational costs. However, this fact often corresponds to low quality output which complicates the computer processing of the result gel image. The most common types of outcome degradation are geometric distortion of sample lanes on electrophoretic gel, which occurs already during an electrophoretic process, and the low contrast of the gel image caused by contamination of samples and camera quality. Because the origin of these problems is different and varies, it is appropriate to solve them separately.

Many authors solve a geometric distortion of gel images by modelling their characteristics with subsequent compensation, but a nontrivial distortion is often unpredictable. The presented solution lies in two steps, which separately compensate the distortion in direction of x and y axes regardless of the type and shape of the distortion. The special designed algorithm of the sample lanes detection searches line trend of samples borders and depends only on the changes of gray levels of pixels in the gel image. The trend of line does not have any important characteristic. The correct lane segmentation compensates the geometric distortion in horizontal direction (the lane flexion). The compensation of various relative speeds of different samples on gel in vertical direction is realized by DTW. This step does not solve typical distortions only (as smile affect) but also the individual distortions of each sample. Again, it does not model distortion, because this method adjusts samples to each other, so more precise result is achieved by multiple samples. This fact is related with requirements of similarity analysis, which also needs more samples (10 samples are sufficient).

The elimination efficiency of geometric distortion influence depends on a good visibility of samples in a gel image. The effectiveness of both previous steps increases in dependence on the image quality. The special combination of techniques for image adjustment and signal filtering was proposed for this reason. The contrast adjustment by piecewise linear transform function in combination with the gamma correction improves the lane detection. Filtering of unequal background of each sample improves the position adjustment by DTW and we obtain more accurate evaluation of mutual similarity independent on the background impurities.

The most important fact is that all these steps are adaptive and independent on the used electrophoretic method. It offers universal use and full automaticity of electrophoretic gel image similarity analysis.

ACKNOWLEDGEMENTS

Supported by European Regional Development Fund - Project FNUSA-ICRC (No. CZ.1.05/1.1.00/02.0123) and by the grant project GACR P102/11/1068 NanoBioTECell.

References

1. M. Feizabadi, I. Robertson and D. Cousins, *Microbiology*, 143 (Pt 4 (1997) 1461.
2. J. a. Morris, *Journal of general microbiology*, 76 (1973) 231.
3. S. Krizkova, M. Ryvolova, J. Gumulec, M. Masarik, V. Adam, P. Majzlik, J. Hubalek, I. Provaznik and R. Kizek, *Electrophoresis*, 32 (2011) 1952.
4. C. Malin and P. Illmer, *Microbiological research*, 163 (2008) 503.

5. a. Gelsomino, a. C. Keijzer-Wolters, G. Cacco and J. D. van Elsas, *Journal of microbiological methods*, 38 (1999) 1.
6. F. Amp and E. Miambi, *International journal of food microbiology*, 60 (2000) 91.
7. M. H. Nicolaisen and N. B. Ramsing, *Journal of microbiological methods*, 50 (2002) 189.
8. T. Zhang, *Biotechnology letters* (2000) 399.
9. T. M. LaPara, C. H. Nakatsu, L. Pantea and J. E. Alleman, *Applied and environmental microbiology*, 66 (2000) 3951.
10. A. E. Murray, J. T. Hollibaugh, C. Orrego, A. E. Murray and J. T. Hollibaugh, *Applied and Enviromental microbiology*, 62 (1996) 2676.
11. C. R. Woese, *Proceedings of the National Academy of Sciences of the United States of America*, 97 (2000) 8392.
12. I. Bajla, I. Holländer, S. Fluch, K. Burg and M. Kollár, *Computer methods and programs in biomedicine*, 77 (2005) 209.
13. J. Pizzonia, *BioTechniques*, 30 (2001) 1316.
14. R. T. F. Wong, S. Flibotte, R. Corbett, P. Saeedi, S. J. M. Jones, M. a. Marra, J. E. Schein and I. n. Birol, *IEEE Transactions on Automation Science and Engineering*, 7 (2010) 706.
15. N. Kaabouch, R. R. Schultz and B. B. Singh, *2007 IEEE International Conference on Electro/Information Technology* (2007) 577.
16. P. Salas and P. Alvarado, in Conference on Technologies for Sustainable Development TSD2011, 2011, p. 53.
17. X. Ye, C. Suen and M. Cheriet, *Vision Iterface* (1999) 19.
18. D. C. Rio, M. Ares, G. J. Hannon and T. W. Nilsen, *Cold Spring Harbor protocols*, 2010 (2010) pdb.prot5444.
19. J. Jan, Medical Image Processing, Reconstruction And Restoration: Concepts And Methods, Taylor & Francis, 2006.
20. C. Dah-Chung and W. Wen-Rong, *Medical Imaging, IEEE Transactions on*, 17 (1998) 518.
21. A. Machado and M. Campos, in Computer Graphics and Image Processing, 1997. Proceedings., X Brazilian Symposium on, 1997, p. 140.
22. S. C. Park, I. S. Na, T. H. Han, S. H. Kim and G. S. Lee, *Computers and Electronics in Agriculture*, 83 (2012) 85.
23. A. Sousa, R. Aguiar, A. Mendonca and A. Campilho, *Image Analysis and Recognition*, 3212 (2004) 158.
24. S. E. Shadle, D. F. Allen, H. Guo, W. K. Pogozelski, J. S. Bashkin and T. D. Tullius, *Nucleic acids research*, 25 (1997) 850.
25. Y.-K. Chan, S.-W. Guo, H.-M. Cheng and P.-H. You, *2011 International Conference on Electronics, Communications and Control (ICECC)* (2011) 3586.
26. I. Bajla and I. Hollander, *Measurement science review*, 1 (2001) 5.
27. H. Sakoe and S. Chiba, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26 (1978) 43.
28. T. Giorgino, *Journal of Statistical Software*, 31 (2009) 1.