An alternative method for electrophoretic gel image analysis in the GelMaster software.

ARTICLE *in* COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE · APRIL 2005

Impact Factor: 1.9 · DOI: 10.1016/j.cmpb.2004.09.007 · Source: PubMed

5 AUTHORS, INCLUDING:

Ivan Bajla
Slovak Academy of Sciences
46 PUBLICATIONS **245** CITATIONS
SEE PROFILE

Silvia Fluch
AIT Austrian Institute of Technology
45 PUBLICATIONS **760** CITATIONS
SEE PROFILE

Kornel Burg
AIT Austrian Institute of Technology
72 PUBLICATIONS **2,010** CITATIONS
SEE PROFILE

Computer Methods and Programs in Biomedicine (2005) **77**, 209—231

ELSEVIER

Computer Methods and Programs in Biomedicine

www.intl.elsevierhealth.com/journals/cmpb

An alternative method for electrophoretic gel image analysis in the GelMaster software

I. Bajla[a,*], I. Holländer[a], S. Fluch[b], K. Burg[b], M. Kollár[a]

[a] *Department of High Performance Image Processing, ARC Seibersdorf research GmbH, A-2444 Seibersdorf, Austria*
[b] *Biotechnology Unit, ARC Seibersdorf research GmbH, A-2444 Seibersdorf, Austria*

Received 5 May 2003; received in revised form 9 August 2004; accepted 13 September 2004

KEYWORDS
Electrophoretic gels;
Image processing;
Geometry-driven diffusion;
Edge detection

Summary A novel methodology of electrophoretic gel image analysis has been proposed that is based on two-dimensional image processing methods instead of previously used one-dimensional Gaussian deconvolution. The crucial problem of the analysis of imperfect gels, that consists in band detection, is solved using the algorithms of band boundary detection and intensity homogeneity indication. The template approach represents the core element of the developed algorithms. The GelMaster software system has been developed in which the novel algorithms are implemented. It involves two-stage interaction with the user: detection of the true bands and deleting the false band detections. The main features of the GelMaster system and the most important algorithms are described.

1. Introduction

The development of gel electrophoresis as a method of separating DNA molecules has significantly affected the progress achieved in molecular biology in the last 20 years. The electrophoresis [1—3] is a method that separates molecules on the basis of their size, electric charge, and other physical properties. The *DNA gel electrophoresis* [4,5] refers to the technique in which DNA macromolecules are forced to migrate across a span of gel, which is a colloid in solid form, motivated by an electrical current. In agarose gels stained with ethidium bromide (EtBr) or silver-stained polyacry-

Abbreviations: GAS1, gel analysis system 1 (our pilot application); EtBr, ethidium bromide; GIA, gel image analysis; GUI, graphical user interface; BB, band boundary; PAGE, polyacrylamide gel electrophoresis; DGGE, denaturated gradient gel electrophoresis; GDD, geometry-driven diffusion; GR, gray scale of an image; ROI, region of interest in an image; BFP, band finding panel in the GelMaster software; TEDEBBY, template detector of band boundary; TOODIS, tooth discrimination operator for excluding false BB detections; MAMBO, median above median below operator for excluding false BB detections; LOBBY, length of band boundary operator for excluding false BB detections; DSS, double stack strategy operator for excluding false BB detections

* Corresponding author. Tel.: +43 50550 2903.
E-mail address: ivan.bajla@arcs.ac.at (I. Bajla).

0169-2607/$ — see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2004.09.007

lamide gels, which are mostly used for DNA gel electrophoresis, the electrophoretic mobility of macromolecules is thought to be determined primarily by the volume fraction of pores within the gel that the macromolecules can enter. Since small DNA molecules can fit into more pores than large DNA molecules, small DNAs will migrate through the gel matrix faster than large DNAs. Hence, a mixture of DNA molecules of different sizes will separate into discrete blobs in the process of electrophoresis. Samples of DNA, digested by specific restriction enzymes into segments of different sizes, are loaded in the separated wells at the front edge of the gel. After turning on the power supply voltage each sample runs in its own trace. A fluorescent dye is used for marking the positions of the DNA bands in the gel which is then photographed to provide a permanent record of the electrophoresis experiment.

Besides the basic *one-dimensional* type of electrophoresis, described above, methods for *two-dimensional* electrophoresis have also been developed. In these methods, separation of molecules is organized as a two-step process in the perpendicular directions [6—8]. The methods and the software system GelMaster described in the paper are dedicated to one-dimensional electrophoresis.

In the recent years, a number of software packages, e.g., Scanalytics™, GelComparII™, Gel-Pro Analyzer™ and TotalLab™, have been developed for analysis of digitized gel images to get the final information on DNA fragment mobility, i.e., the band positions in each individual lane given in molecular weight (or DNA base pairs). To meet high throughput demands on gel image analysis (GIA) in the Biotechnology Unit, ARC Seibersdorf research GmbH, several commercial software packages have been tested. Unfortunately, the everyday practice with these packages revealed a number of problems in detection of low-contrast bands and groups of bands located closely to each other which frequently occurred in real gel images. The existing approaches are not capable to resolve small differences in DNA fragment migration and automatic mode usually leads to missing bands and falsely detected bands. In such cases, the only way out is to use time-consuming purely manual approach. Therefore, the Biotechnology Unit together with the Department of High Performance Image Processing initiated a research project ''Gel image analysis'' with the goal to improve algorithms for gel image analysis and to design a software system that could resolve or minimize the crucial problems of the GIA.

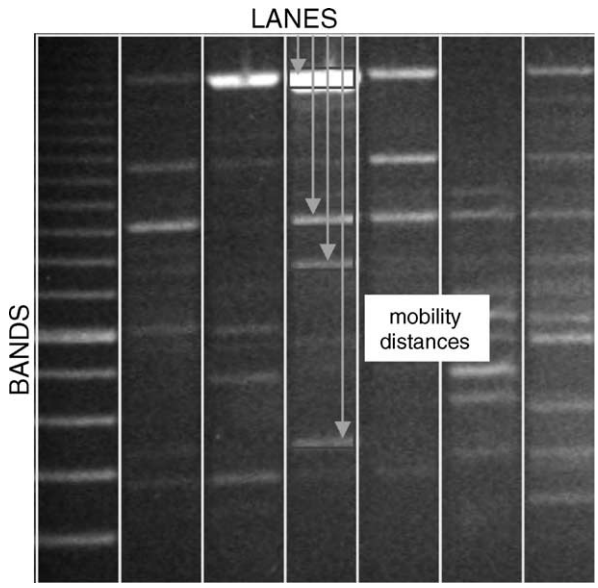

Fig. 1 Basic terminology for gel images: vertical strips are lanes, horizontally oriented blobs representing DNA deposits are bands.

2. Background

In Fig. 1, a typical gel image obtained after digitization of a film record of an electrophoretic experiment is displayed. There are three basic notions related to gel images: (i) *lanes* which denote vertical strips, each representing the information on DNA molecule separation that is uniquely related to one sample of the digested DNA, (ii) *bands* which denote horizontally oriented image blobs in each lane representing deposits of DNA molecules in the gel, and (iii) *mobility distances* which are the values representing the band positions in appropriate units (molecular weights or base pairs).

2.1. Ill-posed problem of one-dimensional signal deconvolution

In the recent decade, methods for computer-based analysis of one-dimensional electrophoretic gel images [9—12] have been proposed. The basic approach to detection of bands in individual lanes applied in these methods is based on one-dimensional profile of each lane image. Peaks of the profile, called sometimes as density profile, that is some representation of intensities in the rows (sum, mean, etc.) are being searched and then declared as band positions. This approach is successful provided the given gel image is perfect. A number of various defects, distortions and degradation can occur in gel images, most of which can be suppressed or removed by tools available in the commercial

software for GIA. However, noise and local artifacts cause essential problems of proper band detection, particularly if low contrast bands occur. The failure of the standard software tools, which work satisfactorily in perfect gel images, is related to the properties of one-dimensional deconvolution operation conventionally used in the software systems. The reasons underlying these problems can be described within the theoretical framework of one-dimensional signal deconvolution. This operation is conventionally defined as a solution of a linear first-kind Fredholm integral equation:

$$\int_0^1 K(s,t)f(t)\,dt = g(s), \quad 0 \le s \le 1, \tag{1}$$

with a convolution type kernel $K(s,t)$. In contemporary programs, the two-dimensional band image structures in a lane are represented by an averaged one-dimensional profile $g(s)$ of the intensities along the lane. The solution $f(t)$ of Eq. (1) is searched as a one-dimensional Gaussian deconvolution applied to the profile $g(s)$. According to [13], deconvolution of signals corrupted by noise and other defects encounters the following crucial problems: (i) ill-conditioning of this inverse problem, and (ii) the function $K(s,t)$ is unknown and can only be estimated from some convenient functional class. Furthermore, in the case of gel images, this function cannot serve as a model for image blurring mechanism, and the real two-dimensional problem of gel image analysis is simplified to a one-dimensional problem. For improvement of band detection operation, we propose to regularize the original gel image by a two-dimensional noise filter, and, instead of using one-dimensional deconvolution, to develop specific band boundary (edge) detectors and lane image homogeneity indicators (measures).

2.2. Need of change in interaction paradigm

Due to the high variability in quality of gels, it is not possible to resolve all practical situations in a fully automatic mode. Therefore, the existing software systems incorporate specific tools for interactive setting the numerical parameters of band detection algorithms. These tools are based on the multiparametric interaction paradigm: (i) the user is prompted to adjust several numerical parameters to control the Gaussian deconvolution operation, and (ii) a certain final result of band detection is displayed. If the user considers some detections to be false and, conversely some true band detections missing, he/she has to return and repeat the preceding step. The search for an optimum combination of the numerical parameters represents a troublesome process. The user has no possibility to control the band detection operation separately in relation to one of two crucial types of errors, i.e., missing true band boundary (BB) detections, and false BB detections. The only alternative to the *assisted complete detection* is manual marking of each individual band that is very laborious and inaccurate.

2.3. User requirements

As mentioned in Section 1, the main goal of the interdisciplinary research project ''Gel image analysis'' was to design a software system with improved functionality for everyday practice in molecular biology labs. The requirements of the common user for such a system can be divided into two groups: general requirements of easy routine gel image analysis, and specific requirements to band detection stage of the analysis.

2.3.1. General requirements

- Increased analysis throughput of complex gel images occurring particularly in research projects.
- Possibility to load images in various formats.
- User-friendly selection of geometrical region of interest (ROI) in the image.
- Simple adjusting of brightness and contrast of the image.
- Higher accuracy and precision in calculation of molecular weight and amount of bands.
- Variability of graphical tools for reporting results.
- Faster system response.

2.3.2. Specific requirements

- Tools for semi-automatic analysis which enable to achieve final proper band detection within reasonable time without need to adjust values of numerical parameters.
- The proportion between automatic and interactive analysis should follow the rule: the higher the quality of the gel image, the more analysis tasks can be solved automatically.
- Using simple and easy-to-learn graphical tools in the process of controlling the band detection process.

3. Design considerations

3.1. Novel processing philosophy

We proposed a novel GIA philosophy featured by the following characteristics [14]:

- Instead of one-dimensional data processing, algorithms for two-dimensional image data processing are used.
- Detection of bands is replaced by detection of *band boxes* which represent band boundaries.
- The process of the detection of band boundaries is divided into two stages: (i) semi-automatic detection of band boxes; in the stage I, all true detections must be detected (the false detections appearing in this stage are ignored), and (ii) coupling of the individual detections into band boxes and rejection of all false detections; stage II is also performed in semi-automatic mode.
- After finishing these two stages, a final mobility distance representation is calculated for each band box, using all intensities occurred within the band box.

To implement this strategy, novel task-specific edge detectors and band boundary indicators have been developed. An essentially modified graphical user interface has been designed which enables user-friendly interaction in both semi-automatic modes.

Based on our pilot study [15] with noisy gel images of various types, we decided to preprocess the gel images by two methods. The goal of the first method is to regularize noisy images, thereby to simplify the task of band boundary detection. The second preprocessing method is specifically aimed at correction of vertical smearing effect encountered in the class of silver-stained native polyacrylamide (PAGE) gel images.

3.2. Image preprocessing using geometry-driven diffusion filtering

In [16], we explored modifications of the nonlinear image filtering method that is based on the principle of *geometry-driven diffusion* (GDD) [17,18], given by the equation:

$$\frac{\partial I}{\partial t} = \mathrm{div}[c(|\nabla I(x,y,t)|)\nabla I(x,y,t)], \tag{2}$$

where $\nabla I(x,y,t)$ is the gradient of an image intensity function $I(x,y,t)$, the conductance $c(\cdot)$ is a function of spatial coordinates (x, y) and t is time which corresponds in discrete implementation to the index of the iteration step. This noise suppression method outperforms other linear and nonlinear image filters, as proved in papers [16—19], since the noise smoothing effect is not achieved at the expense of smoothing edges. In [16], we describe the positive effect of the GDD to the signal-to-noise ratio, which, in our case, is illustrated in Fig. 2.

3.3. Image preprocessing using Fourier notch filter

Typical PAGE images suffer from a serious degradation characterized by vertical stripes caused by smearing effect in the gels. As a consequence, one can observe considerable distortions of individual bands, largely at the location of the lane boundaries. These distortions make proper detection of band boundaries even more difficult.

It is known [20] that the periodic nature of image distortions produces bursts of concentrated

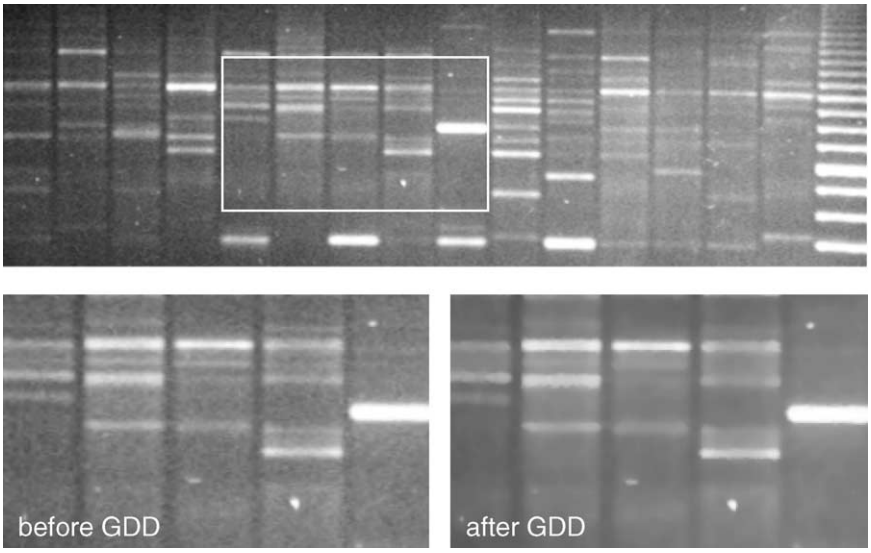

Fig. 2 Input gel image with two fragments before and after GDD filtering. Noise is suppressed while band edges are preserved.

energy in the horizontal or vertical axis of the Fourier spectrum of the image. A simple approach for reducing the effect of such kind of distortions is to use a ''notch'' filter $H(u, v)$ which attenuates the values of the Fourier transform in the coordinate axes and multiplies all other values of the transform (Fourier spectral coefficients) by 1. The corrected image $p(x, y)$ of the input corrupted image $g(x, y)$ is then obtained by the inverse Fourier transform:

$$p(x, y) = \mathcal{F}^{-1}\{H(u, v)G(u, v)\},$$

where $G(u, v)$ is the Fourier transform of the input image $g(x, y)$.

The notch filter is characterized by two parameters: (i) the *width* of the horizontal (or vertical) rectangular notches, and (ii) the size of the gap between two notches (Fig. 3, bottom right window). The original PAGE image is first converted from the colour domain (RGB) to gray scale (GR) by the transform $GR = c_R R + c_G G + c_B B$, where c_R, c_G, and c_B are the weights of individual colour components. Due to the colour characteristics of the polyacrylamide gels, we chose the values of the weights: $c_R = 0.4, c_G = 0.4$, and $c_B = 0.2$.

We have accomplished a set of computer experiments [15] with several images using various values of two parameters of the filter. The first parameter is responsible for changing the average intensity value and overall level of blurring. It also leads to deleting the spectral components corresponding to band edges (boundaries) which are therefore blurred. The filtering effect (attenuation of the vertical stripes) is significantly affected by the second parameter, the gap between two rectangular notches. The greater the gap between notches, the less Fourier coefficients in the x-axis are suppressed and more vertical stripes are preserved (not filtered out). We have found optimum values of the parameters of the notch filter: $w_1 = 0.1$ and $w_2 = 0.1$. In Fig. 4, the corrupted original input image and the result of filtering by the notch filter with the optimum parameter values is displayed.

4. System description

The GelMaster software system has been based on our pilot implementation GAS1 [21]. For research purposes, the first implementation was written in MATLAB™ (The MathWorks Inc.). The next implementation GelMaster is an application written in Visual Basic. It is suitable for practical applications in molecular biology labs.

4.1. Flow of image processing operations in GelMaster

Discussions with biologists who analyze DNA gel images in everyday practice lead to the speci-

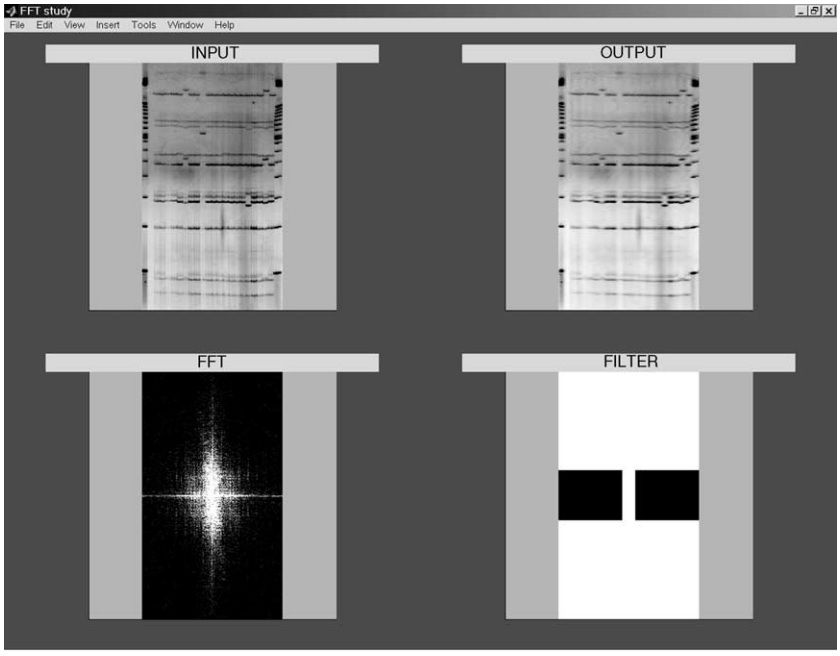

Fig. 3 The illustration of the notch filtering in Fourier domain. Top left: input gel image with vertical smears; bottom left: Fourier spectrum of the input image; bottom right: cut-off of the relevant part of the spectrum; top right: resulting image with smears suppressed.

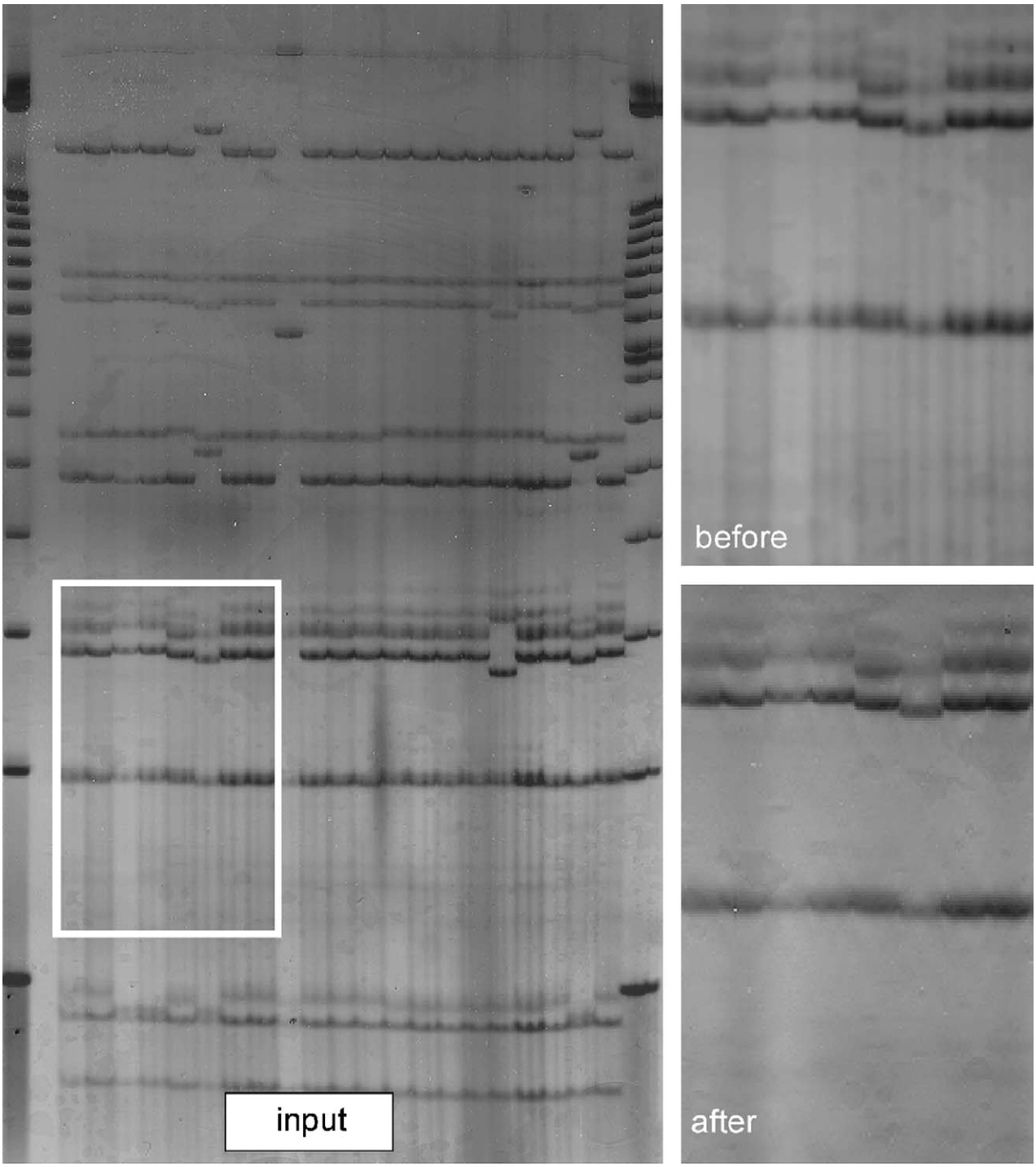

Fig. 4 Original input image with two fragments before and after filtering by the notch filter in the Fourier domain, with the optimal parameters $w_1 = 0.1$ and $w_2 = 0.1$. Disturbing vertical smears are removed.

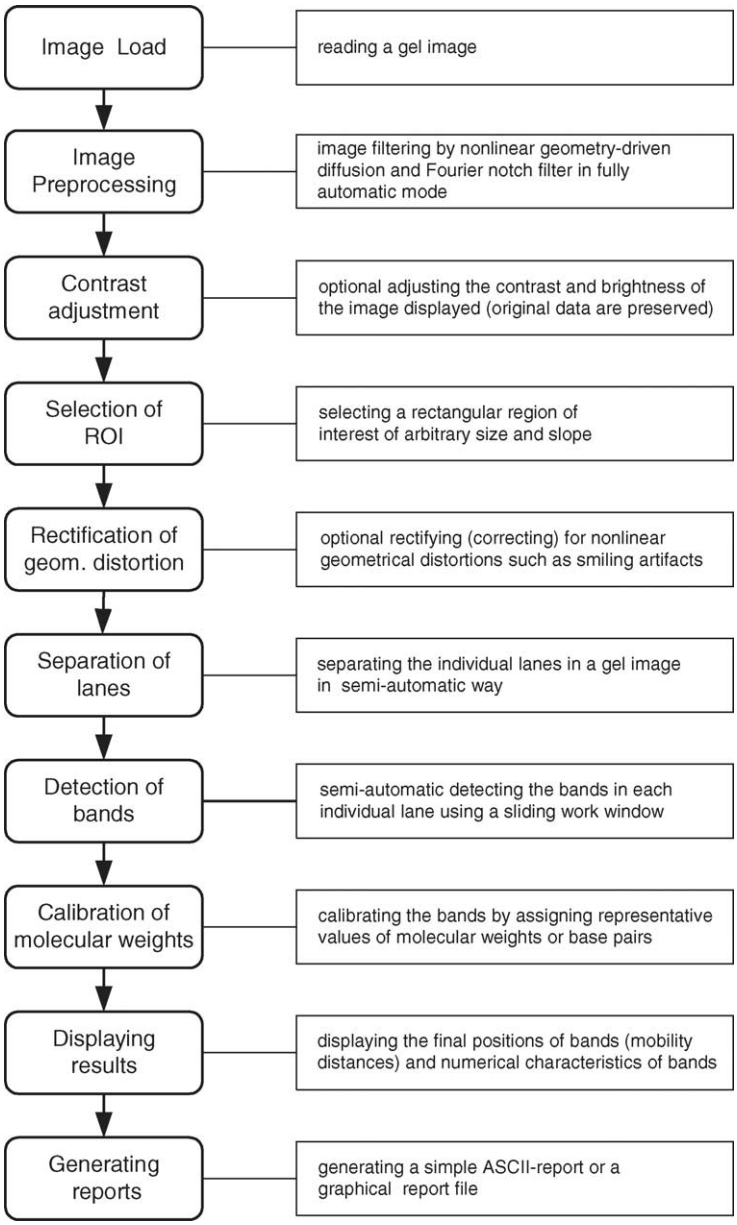

Fig. 5 Main flow of operations in the GelMaster.

fication of the appropriate sequence of individual operations which constitutes a standard gel image analysis session. The sequence of operations, which served as a basis for the design of the GelMaster main structure, is depicted in Fig. 5.

The operations (except image preprocessing) are semi-automatically oriented. They are implemented with the emphasis on simplicity of user interaction, namely, graphical tools are used instead of setting numerical parameters by the user. The operations of contrast adjustment and rectification of geometrical distortions in gel images are optional. The detection of bands represents the core operation within the whole sequence. The novelties of the methods proposed and algorithms implemented into the program system GelMaster are concentrated exactly in this operation.

4.2. The layout of the basic windows in GelMaster

The basic layout of the GelMaster is displayed in Fig. 6. The main window is divided into three interaction areas:

Button panel (left-hand side of the window, marked by 1): This panel contains buttons which control the processing stages, i.e., principal steps

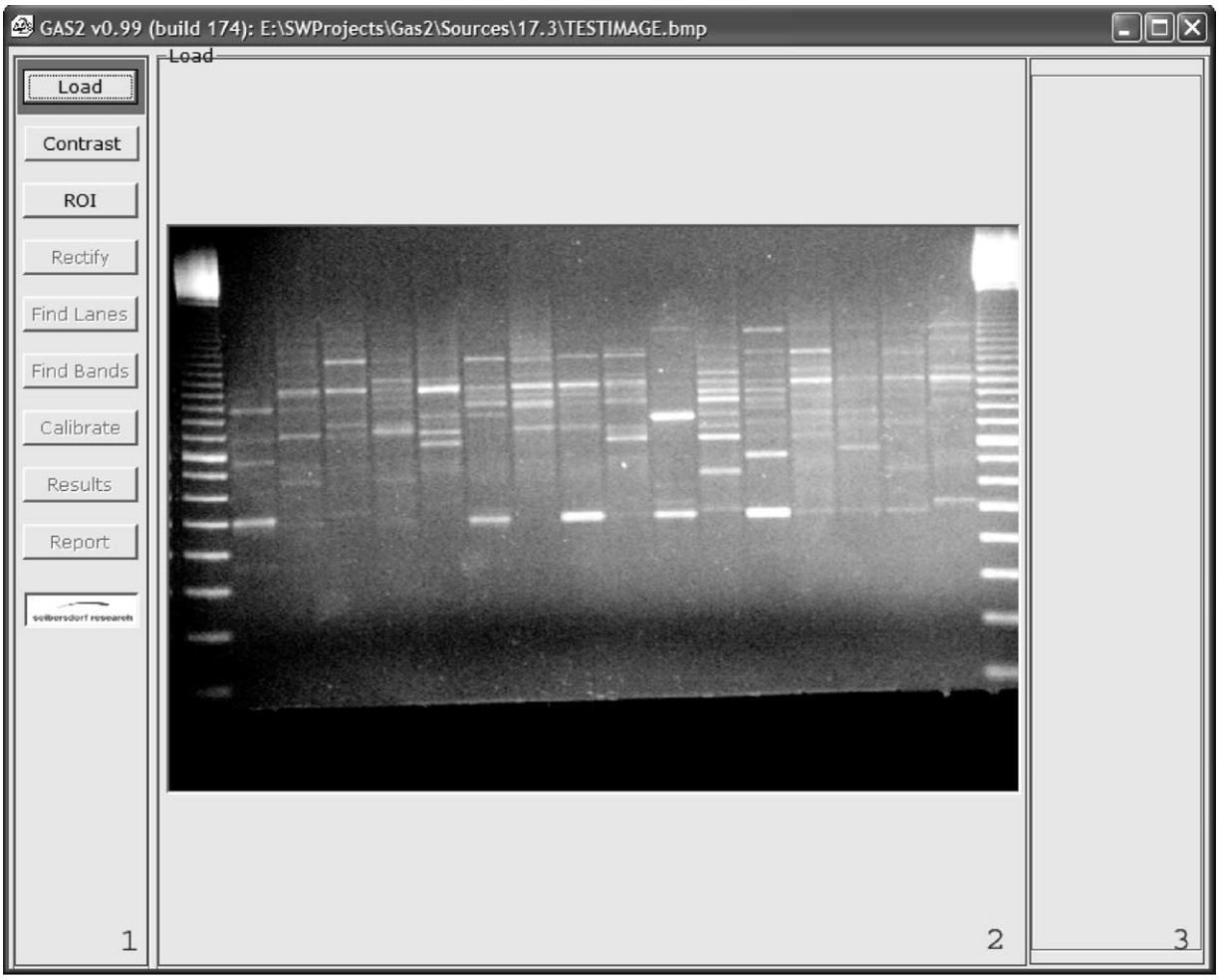

Fig. 6 Basic layout of GelMaster.

of the gel image analysis process. The buttons are arranged from top to bottom in the normal processing order. The individual buttons are only enabled when it makes sense to use them (e.g., it is nonsense to search for bands if the lanes have not been found yet). Nevertheless, the user can return to any previous stage to check or adjust the settings just by pressing one of the buttons in virtually any situation. If the setting made would affect the following stages, the corresponding buttons are disabled. The button representing the currently active stage is highlighted.

Main image display (center of the window, marked by 2): The main image display shows the image under processing in a suitable form such as full image, selected region of interest, separated lanes or similar. Virtually all image interaction takes place in this panel.

Panel for stage-specific controls (right-hand side of the window, marked by 3): This panel contains buttons, sliders, edit windows, etc. which are stage-specific. A brief explanation of the controls in this panel will be given in the corresponding description parts which follow.

4.3. Image load

The GelMaster can load gel images in several standard file formats, like BMP, GIF, JPEG, etc. The GelMaster copes with one gel image at a time.

4.4. Contrast adjustment

Interactive working with images requires setting of contrast and brightness of the display in a way most convenient for the user. The implementation of the operation ''Contrast'' is based on the following simple paradigm. The brightness is controlled (from maximum to minimum value) by moving the directional cursor up and down (mouse button pressed). Moving the cursor left and right changes

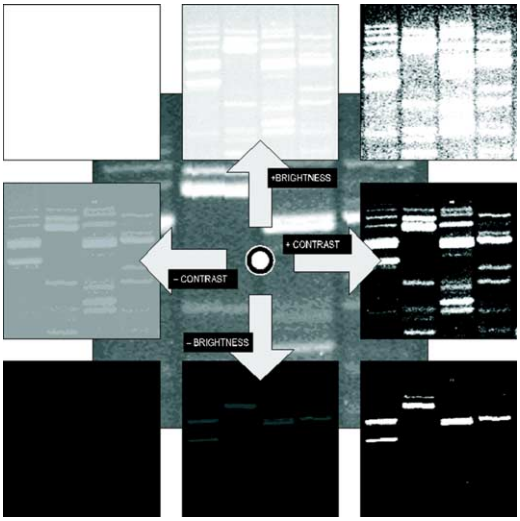

Fig. 7 The scheme of interactive control of the image contrast. Moving the cursor up and down controls image brightness, while moving it left and right changes image contrast.

the contrast (from minimum to maximum value). The cursor position in the middle corresponds to the standard brightness and contrast settings. The setting is fixed just by depressing the mouse button in the desired position. In Fig. 7, the ranges of variability of the contrast/brightness setting are displayed. A picture with minimum contrast (left) is coloured uniformly—gray, completely white, or completely black, dependent on the brightness setting. On the other hand, maximum contrast means only two images shades have been left: white and black. This effectively means an image processing operation called thresholding. The threshold value can be controlled by moving the cursor up and down. Image data processing is independent on the contrast/brightness setting. Therefore, this setting can be made anytime in the session.

4.5. Selection of ROI

A typical gel image usually covers a larger area than the region with important lanes and bands (it may show excessive margins which are not only irrelevant but they can contain a number of misleading artifacts). Therefore, it is necessary to restrict the GIA only to a region of interest. This should cover all lanes and all bands but not more. Such a region is typically a rectangle which may be slightly rotated with respect to the rectangular frame of the input image. In the GelMaster, the selection of a rectangular ROI with arbitrary rotation can

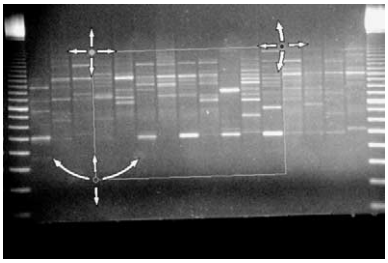

Fig. 8 Interactive tool for ROI selection. Possible types of movements and settings are symbolized by arrows.

be performed in the following interactive way: after pressing the ''ROI'' button in the left panel, an ROI definition marker is overlaid on the image display (see Fig. 8). It consists of a rectangle and three control points (green and blue in original layout). The interior of this rectangle represents the ROI of the given gel image. The position, size, and slope of the ROI can be adjusted using three control points. The top left (basic) control point controls the translation of the rectangle (the slope and size are preserved). The top right and bottom left control points simultaneously control size and rotation of the rectangle. Dragging the latter two points away from the top left control point (which is static) modifies the length of the respective side of the rectangle. Dragging the two above-mentioned control points around the basic point rotates the rectangle. While a modifier is pressed, dragging these control points modifies only the length of the respective side of the rectangle and the slope is preserved.

4.6. Rectification of geometrical distortions

In most contemporary GIA software systems, geometrical distortions (nonlinear in general) are corrected on the basis of the following paradigm: estimated correction curves are superimposed on the underlying gel image to enable the calculations of the proper quantities of mobility distances; the image itself is not transformed and its distorted appearance is preserved. We chose another approach, namely, the rectification (warping) of the entire image. It means that all subsequent operations, starting with lane separations, are applied to an already corrected image. The rectification algorithm we implemented is based on superposition of a regular rectangular grid on the input image. The grid consists of interconnected control nodes which represent target positions. As default (Fig. 9), control nodes are arranged in six rows by nine columns (additional rows

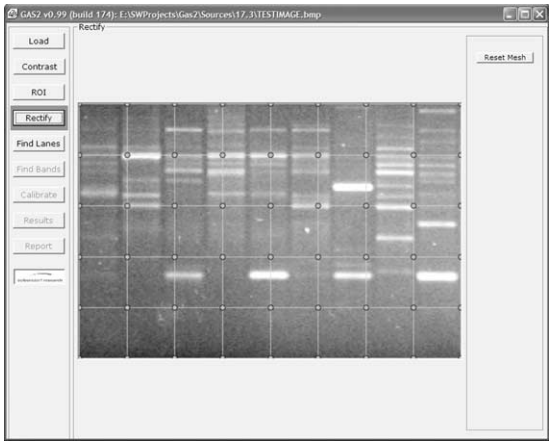

Fig. 9 The image rectification window of GelMaster with the initial regular grid of nodes. The user selects some of them for moving them to follow the shape of geometrical distortions.

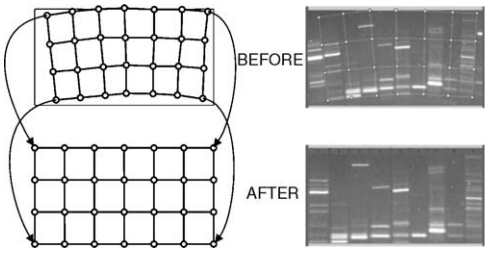

Fig. 10 The image rectification principle used for geometrical corrections. Top right: input distorted image with the grid deformed interactively by the user; bottom right: the result of rectification image transformation.

and columns of nodes can be added interactively).

Let us assume that the image being processed is geometrically distorted. Most frequent distortions (smiles) are best visible on bended margins of the lanes. The user moves selected grid nodes to simulate the distortion of the underlying gel image. Other nodes remain unchanged. The resulting distorted grid and the initial regular grid (that remains transparent for the user) define a geometrical transformation for the needed correction of the input gel image. The parameters of this transformation, calculated for grid nodes, are then applied to the entire image raster and the final corrected image is calculated (Fig. 10).

4.7. Separation of lanes

After the ROI selection and optional rectification, the image should consist of parallel vertical lanes.

Fig. 11 The GelMaster window for lane separation depicted by graphical overlay of the vertical lane margins.

There should be no lateral margins in the image before starting the operation lane separation. After pressing the ''Find Lanes'' button in the left panel, the main image display is overlaid with a system of markers (vertical lines) (Fig. 11). The user has to set the number of lanes using a slider in the upper part of the right panel. The system of lane markers is updated continuously with the setting the slider to proper position. When the pointer moves over the lane, the lane is highlighted. By pressing the right mouse button over the highlighted lane, the user can invoke the context-sensitive menu for that particular lane. For each value of the number of lanes, the algorithm calculates the positions of the lane boundaries in the case of gaps among them. In some cases (extremely noisy gel images or specific artifacts overlapping the gaps), it is better to use a trivial algorithm which divides the entire width of the ROI into required number of equally spaced lanes.

Although the lane separation (detection of their boundaries) was not in the focus of the GIA research project, we explored various approaches to this non-trivial problem, including an algorithm based on a compound criterion of accepting the candidates of lane boundaries generated by a functional metric, an algorithm based on the theory of testing the outliers in statistical data applied to lane candidates or a method based on cross-correlation of intensity column vectors. Finally, we proposed the algorithm based on one-dimensional cumulative indicator of lane edges coupled with shifted regular initial grid calculated from the a priori information on lane number. We briefly describe this method.

Let $[G(i,j)]$ be the $M \times N$ intensity matrix representing an image of the gel, and let $[g(i,j)]$ be

an $m \times n$ lane image matrix for which we set $m = M, n = N/\text{NoLanes}$, where NoLanes is the number of lanes in the given gel image. Further, we denote by sum_j the sum of the intensities in the j th column: $\text{sum}_j = \sum_{i=1}^{M} G(i, j)$. Due to prevailingly oblong shape and horizontal orientation of the bands, these column sums differ for neighboring columns within one lane less than for a pair of the columns in which one column represents a lane boundary, while the other one comes from a lane interior or a gap between neighboring lanes. The presence of local variations of band intensities or artifacts within the lane background affects the cumulative quantity of sum_j only weakly. To further minimize these defects the vector $\mathbf{Sum} = [\text{sum}_1, \text{sum}_2, \ldots, \text{sum}_N]$ is smoothed by a low-pass mean filter. Then, the resulting vector **FiltSum** serves as a basis for lane boundary indication. Namely, we can calculate the absolute value of the first differences of the **FiltSum** to get a robust lane boundary indicator $\mathbf{DifFiltSum}(j) = |\mathbf{FiltSum}(j+1) - \mathbf{FiltSum}(j)|$.

Our computer experiments showed that for real gel images, it is difficult to find an appropriate global threshold of the quantity $\mathbf{DifFiltSum}(j)$ that could discriminate individual lanes or lanes and gaps. This is basically due to local variations of intensities overlapping the gaps between lanes. Therefore, we developed another strategy for lane separation that is based on the assumption of regularity of the lane widths which is given by the regularity of the combs used in electrophoretic experiments. Calculating and setting the *initial grid* of vertical lines given by the overall number of columns N in the gel image and the number of lanes NoLanes provided by the user, we obtain a grid step $\text{Gstep} = N/\text{NoLanes}$ (in Fig. 12, the initial grid is displayed as a group of white regular vertical bars). The positions of the initial grid are given by a Gstep/2 shifts from both sides of the input gel image and the re-calculated grid step $\text{Gstep}' = N/(\text{NoLanes} - 1)$ between the initially set marginal positions. This enables to identify the first and the last lane boundaries independently on the defect of cutting them in the input image. Within each strip given by the Gstep' lines, there is a local maximum of the indicator $\mathbf{DifFiltSum}(j)$ (see Fig. 12 in which the values of this function are plotted as a curve at the bottom). We declare this maximum to be a *supporting* maximum for the given strip. If the input gel image does not comprise gaps this maximum determines the lane boundary location. If there are gaps, we construct penalty functions PenaltyLeft and PenaltyRight aimed at searching for the next maximum that should correspond to the supporting maximum (as gap boundaries).

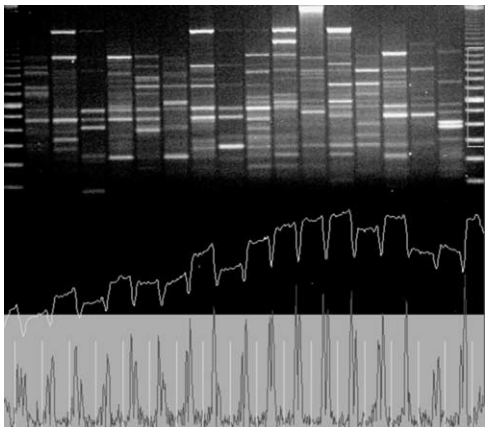

Fig. 12 The illustration of the method for lane separation given the number of lanes which are generally separated by gaps. Top: input image; bottom: regular initial grid at places of lane axes together with the plot of the **DifFiltSum** function; middle (dark) part: plot of the **FiltSum** function.

First, the next left maximum LeftMax and the next right maximum RightMax to the supporting maximum are found. Their indices are denoted as SupInd, LeftInd and RightInd. The distances between the supporting maximum and both of these maxima are given as $\text{LeftDist} = \text{SupInd} - \text{LeftInd}$ and $\text{RightDist} = \text{RightInd} - \text{SupInd}$. Introducing appropriate weights w_1 (the value 0.7 was used) for intensities and w_2 (equal to 0.3 in experiments) for distances, we define the penalty functions PenaltyLeft and PenaltyRight by the following formulae:

$$\text{PenaltyLeft} = w_1 \left(1 - \frac{\text{LeftMax}}{\text{SupMax}}\right) + w_2 \,\text{LeftDist}, \tag{3}$$

$$\text{PenaltyRight} = w_1 \left(1 - \frac{\text{RightMax}}{\text{SupMax}}\right) + w_2 \,\text{RightDist}. \tag{4}$$

For the given supporting maximum, we choose as its pair the maximum for which the penalty value is smaller. It means that the closer the next maximum is to SupMax and the smaller its distance to the index SupInd, the lower value of the penalty function. Thus, the lane boundaries are set indirectly as the boundaries of the corresponding gaps. In Fig. 12, the continuous curve in the dark area illustrates behavior of the function $\mathbf{FiltSum}(j)$

4.8. Detection of bands

The band detection operation represents the crucial point of the GIA. The implementation of this

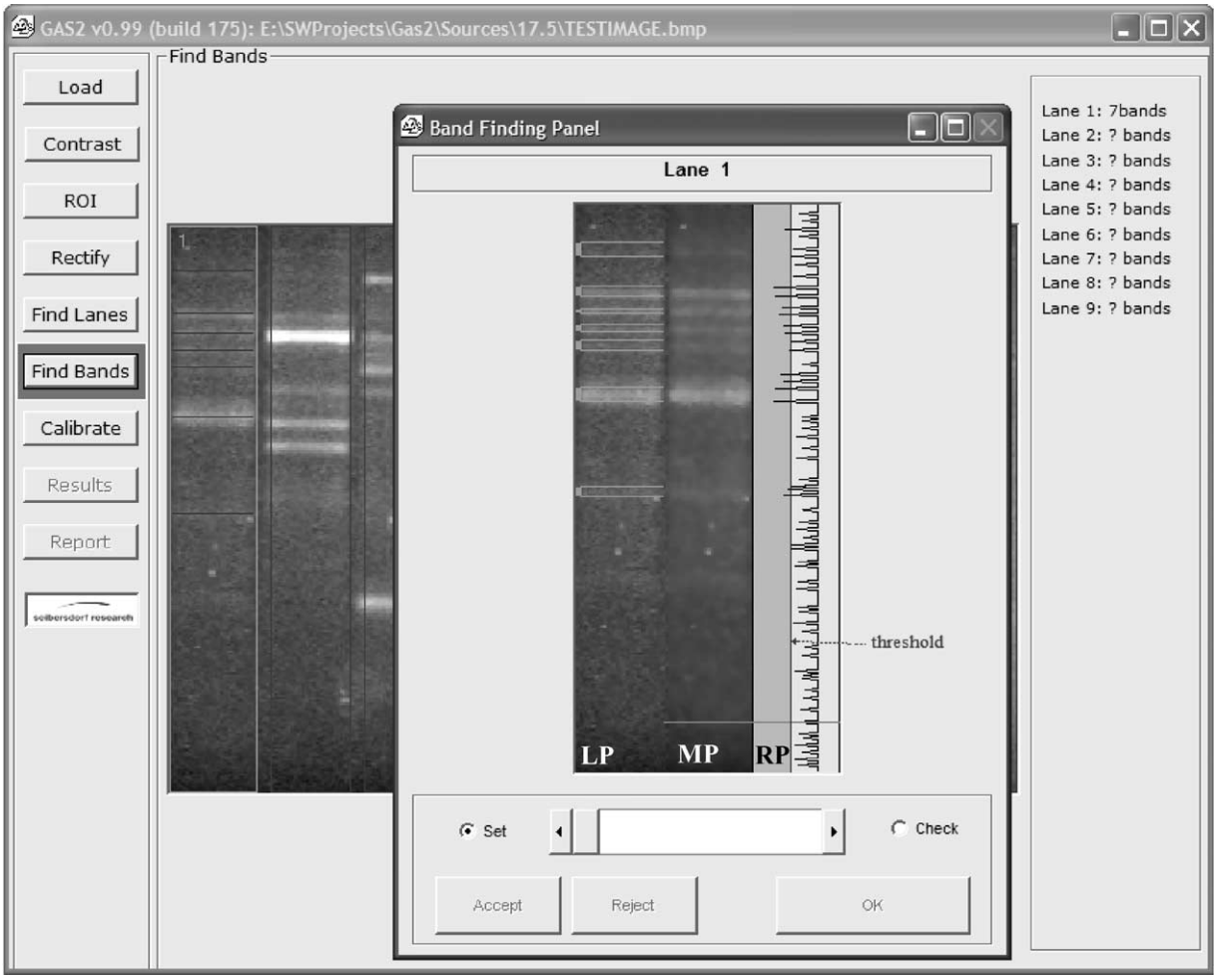

Fig. 13 The GelMaster window for band detection. LP: original lane with band boundary candidates depicted; MP: filtered replica of the lane enabling visual analysis; RP: plot of the maximum values of the band boundary detector with vertical threshold line.

operation in the GelMaster follows the two-stage philosophy we proposed in Section 3. As default result, the band boundary detections in all lanes are calculated using a constant threshold value of detector maxima. The representative positions of bands are marked in all lanes of the gel image. For simplicity, in Fig. 13, band marks in the first lane are displayed only. After pressing the ''Find Bands'' button, a specific window is opened (band finding panel (BFP)). This panel is dedicated for individual analysis of every lane, if necessary. The higher quality of the gel image, the less interaction is needed in subsequent steps. In the BFP, the stage I is invoked by the ''Set'' radio button and the stage II by the ''Check'' radio button (Fig. 13). The processing is organized lane by lane. Moving to the arbitrary lane is accomplished by the slider located in the bottom part of the BFP. The right panel of the main image window contains a check list of the number of bands identified in each individual lane.

For semi-interactive work of the user in both band detection stages, we have designed a triple of the adjacent image (graphical) panes which occupy the main part of the BFP panel. The left pane (LP) contains the original (unfiltered) image of the selected lane. In the ''Set'' mode, the image is overlaid with markers of the positions of the potential band boundaries. The boundaries are shown only if they are paired (which is marked by a rectangle connecting the two boundary lines). Manual insertion of band boundaries is also possible. The middle pane (MP) contains the lane image smoothed by the GDD filter and its primary purpose is to display the lane image without any overlays so the user can examine it in detail with no obstruction. The line passing through this lane is a marker of the actual local maximum of the

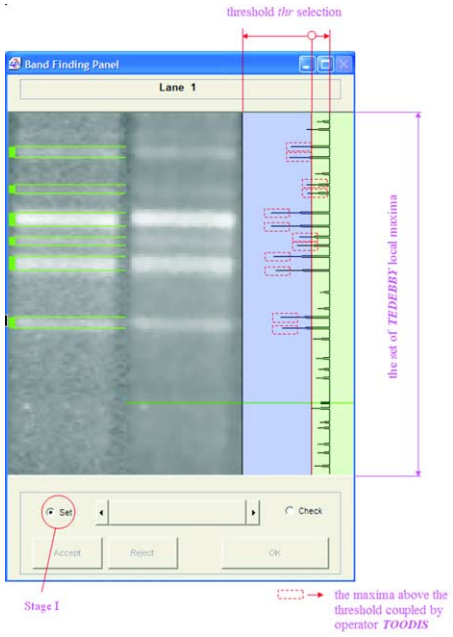

Fig. 14 The GelMaster window in stage I.

boundary detector responses for the given lane. Finally, the third, rightmost pane (RP) of the BFP, contains a plot of the band boundary detector values calculated for the given lane. Only local maxima are displayed (their peaks point to the left). These maxima represent the potential candidates of band boundaries. The user is provided with a graphical tool (vertical line separating the left and right area) to control the threshold of the band boundary detector. By moving with this line (always by one local maxima downwards or upwards) the individual band boundaries detected are marked and dynamically paired to constitute boundaries of the potential bands. This principle and the corresponding visualization of detected band boundaries differ considerably from the ''one-dimensional lane profile'' processing used in common GIA software systems that is based on searching for representative position of each individual band (as relevant peaks of the one-dimensional lane profile). The aim of the stage I, implemented in the BFP under the ''Set'' mode, is to detect all true band boundaries at the expense of some false detections. Our approach is semi-automatic, however, it uses no numerical parameters. The final solution of the stage I is reached for the most cases by fast adjusting the proper position of the threshold line, that is further simplified by two adjacent lane images and graphical information on the relevant peaks in the rightmost pane. Very rare cases of missing true bands can be resolved by manual adding the missing band boundaries. In Fig. 14, a situation for the BFP in the stage I is illustrated. The function of the BFP in ''Check'' mode will be described in the next subsection after explaining the TEDEBBY (template detector of band boundary).

4.8.1. TEDEBBY operator

We explored several approaches to detection of band edges using standard two-dimensional edge detectors and other operators which measure the level of intensity nonhomogeneity in rectangular neighborhoods (e.g., neighborhood nonhomogeneity operator $h(\cdot)$ explored within GDD research [16] or a specific template matching operator). However, the results achieved by these operators in the preliminary computer experiments were not satisfactory. Since the band boundaries manifest prevailing horizontal orientation, an operator accumulating vertical intensity differences has been tested. It proved to work quite satisfactorily in a number of situations except the gel images containing low-contrast bands and the bands located closely to each other. The further research led to a template approach.

We recall the denotation $[g(i,j)]$ for an $m \times n$ pixel matrix representing the image of the lane. We will call (i,j) *the index pair*. The index pair represents *the pixel position*. If we say 'pixel (i,j)', we mean the pixel with position (i,j). *The value of the pixel* in position (i,j) is $g(i,j)$. Let $\{(i,j)\}$ be the set of all possible index pairs in the matrix $[g(i,j)]$. We will call *image region* any subset of $\{(i,j)\}$. *The* (*digital*) *curve* in the image is also regarded to be an image region.

Example 1. The lane image matrix (for the lane image displayed in Fig. 16a) is:

$$[g(i,j)] = \begin{bmatrix} 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 5 & 0 & 0 \\ 0 & 6 & 3 & 6 & 0 \\ 4 & 5 & 6 & 7 & 5 \\ 3 & 5 & 6 & 4 & 3 \\ 2 & 4 & 5 & 2 & 0 \\ 0 & 0 & 2 & 0 & 6 \\ 0 & 2 & 6 & 0 & 3 \\ 0 & 3 & 0 & 4 & 0 \\ 3 & 4 & 0 & 0 & 0 \\ 2 & 5 & 6 & 5 & 1 \\ 0 & 3 & 4 & 4 & 7 \\ 0 & 0 & 4 & 5 & 3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \end{bmatrix}.$$

Here, $m = 15$ and $n = 5$.

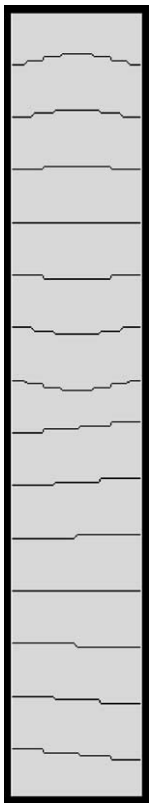

Fig. 15 All 13 templates of band boundary shapes for $S = 3$ depicted as stepwise lines.

The possible shapes of BBs are given by *templates*. *The template of type s* for lane width n is a vector $\boldsymbol{t}^s = (t_1^s, t_2^s, \ldots, t_n^s)$, where t_j^s are small integers. The number t_j^s represents the relative row position of the pixel belonging to the band boundary curve in the column j, with respect to the reference position. It is reasonable to require $t_1^s = 0$ for any s, so that the position of the band boundary curve in the first column is identical to the reference position of the band boundary. In particular, we have two basic types of templates: (i) *slanted line*, and (ii) *arc*. These template types can be further specified by an integer parameter defining their *skewness*. If the skewness is limited by number S, we have in total $4S + 1$ different templates (for both basic types, *skewness* can be $\pm 1, \pm 2, \ldots, \pm S$, and we have one trivial template type for skewness $= 0$). The templates are generated by regular repeating of components whose absolute value does not exceed skewness. In Fig. 15, the set of all templates for $S = 3$ is graphically represented by stepwise lines.

Example 2. For $S = 2$, $n = 5$, we have the following templates:

$$\boldsymbol{t}^1 = \boldsymbol{t}_{\text{arc}}^{\text{skew}=-2} = (0, -1, -2, -1, 0),$$

$$\boldsymbol{t}^2 = \boldsymbol{t}_{\text{arc}}^{\text{skew}=-1} = (0, -1, -1, -1, 0),$$

$$\boldsymbol{t}^3 = \boldsymbol{t}_{\text{arc}}^{\text{skew}=1} = (0, 1, 1, 1, 0),$$

$$\boldsymbol{t}^4 = \boldsymbol{t}_{\text{arc}}^{\text{skew}=2} = (0, 1, 2, 1, 0),$$

$$\boldsymbol{t}^5 = \boldsymbol{t}_{\text{slanted}}^{\text{skew}=-2} = (0, 0, -1, -1, -2),$$

$$\boldsymbol{t}^6 = \boldsymbol{t}_{\text{slanted}}^{\text{skew}=-1} = (0, 0, 0, -1, -1),$$

$$\boldsymbol{t}^7 = \boldsymbol{t}_{\text{slanted}}^{\text{skew}=1} = (0, 0, 0, 1, 1),$$

$$\boldsymbol{t}^8 = \boldsymbol{t}_{\text{slanted}}^{\text{skew}=2} = (0, 0, 1, 1, 2),$$

$$\boldsymbol{t}^9 = \boldsymbol{t}^{\text{skew}=0} = (0, 0, 0, 0, 0).$$

Note. In the text, we combine the notation for the basic type (*arc* or *slanted*) and for the parameter *skewness* in one symbol. The superscript s in t^s (and in similar notations) denotes any of the $4S + 1$ combinations of the basic type and the parameter *skewness*. The range of s is $\{1, 2, 3, \ldots, 4S + 1\}$.

For each template, we define *the cumulative vertical difference* $d^s(i)$:

$$d^s(i) = \sum_{j=1}^{n} |g(i + t_j^s, j) - g(i + t_j^s - 1, j)|. \tag{5}$$

This quantity represents the strength of a potential band boundary of the given shape with the reference position in the row i.

Note. In the algorithm, we only compute $d^s(i)$ for the rows (i) where all quantities on the right-hand side are defined. For other rows, we simply define $d^s(i) = 0$.

For each selected (designed) template $\boldsymbol{t}^s$, we can characterize the lane $[g(i, j)]$ by means of the column vector $\boldsymbol{d}^s = (d^s(1), d^s(2), \ldots, d^s(m))'$. We search for all local maxima of the components of the vector $\boldsymbol{d}^s$, i.e., all such $d^s(i)$, for some $i \in \{1, 2, \ldots, m\}$ for which $d^s(i-1) < d^s(i) > d^s(i+1)$, since these maxima may indicate band boundaries. Then, we arrange these maxima in a matrix $[M_{i,s}]$:

$$M_{i,s} = \begin{cases} d^s(i), & \text{if } d^s(i-1) < d^s(i) > d^s(i+1), \\ 0, & \text{otherwise.} \end{cases}$$

Example 3. For the lane image given in Example 1 and the templates given in Example 2, we have the following column vectors of cumulative

differences:

$$\boldsymbol{d}^1 = (0, 0, 0, 26, 7, 10, 11, 10, 8, 12, 13, 14, 13, 0, 0)',$$
$$\boldsymbol{d}^2 = (0, 0, 0, 23, 8, 7, 12, 12, 9, 14, 7, 20, 9, 0, 0)',$$
$$\boldsymbol{d}^3 = (0, 0, 0, 12, 7, 13, 14, 14, 8, 15, 7, 12, 12, 0, 0)',$$
$$\vdots$$
$$\boldsymbol{d}^9 = (0, 0, 0, 14, 6, 8, 17, 9, 14, 8, 14, 13, 8, 0, 0)'.$$

The matrix $[M_{i,s}]$ is:

$$[M_{i,s}] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 26 & 23 & 12 & 13 & 13 & 14 & 13 & 11 & 14 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 11 & 0 & 0 & 16 & 11 & 0 & 0 & 13 & 17 \\ 0 & 0 & 0 & 0 & 0 & 0 & 13 & 0 & 0 \\ 0 & 0 & 0 & 14 & 0 & 0 & 0 & 0 & 14 \\ 0 & 14 & 15 & 0 & 17 & 0 & 0 & 21 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 15 & 0 & 14 \\ 14 & 20 & 0 & 15 & 15 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In each row i of the matrix $[M_{i,s}]$, we are interested in the template yielding maximum response, as well as in the value of the response itself. Thus, we define the m-dimensional column vectors $\boldsymbol{m}_{\text{init}} = (m_{\text{init}}(1), m_{\text{init}}(2), \ldots, m_{\text{init}}(m))'$ and $(\boldsymbol{a}_{\text{init}} = a_{\text{init}}(1), a_{\text{init}}(2), \ldots, a_{\text{init}}(m))'$ such that:

$$m_{\text{init}}(i) = \max_s(M_{i,s}),$$

$$a_{\text{init}}(i) = arg \ \max_s(M_{i,s}).$$

If there are more equal maxima in the row i, we simply take a randomly selected s of one of them for $a_{\text{init}}(i)$. If $m_{\text{init}}(i) = 0$, we define $a_{\text{init}}(i) = 0$.

A nonzero value $a_{\text{init}}(i)$ represents the most probable shape of the band boundary (simply: *the BB candidate*) positioned in the row i. If $a_{\text{init}}(i) = 0$, we do not expect a band boundary in the row i. The corresponding value $m_{\text{init}}(i)$ represents the 'strength' (likelihood) of the potential band boundary.

Example 4. From the matrix in Example 3, we have:

$$\boldsymbol{m}_{\text{init}} = (0, 0, 0, 26, 0, 0, 17, 13, 14, 21, 15, 20, 11, 0, 0)',$$
$$\boldsymbol{a}_{\text{init}} = (0, 0, 0, 1, 0, 0, 9, 7, 4, 8, 7, 2, 8, 0, 0)'.$$

In Fig. 16b, the obtained BB candidates, represented by the vectors $m_{\text{init}}(i)$, $a_{\text{init}}(i)$, are illustrated by white lines superimposed on the input image. The digital curve representing the band boundary for the given vectors $\boldsymbol{m}_{\text{init}}$ and $\boldsymbol{a}_{\text{init}}$ is denoted by $B_{\text{init},k}$. If $1 \le k_1 < k_2 < \cdots < k_p \le m$ are indices of all nonzero components in the vector $\boldsymbol{m}_{\text{init}}$, then the BB candidates are:

$$B_{\text{init},1} = \{(b_1^1, 1), (b_2^1, 2), \ldots, (b_n^1, n)\},$$
$$B_{\text{init},2} = \{(b_1^2, 1), (b_2^2, 2), \ldots, (b_n^2, n)\},$$
$$\vdots$$
$$B_{\text{init},p} = \{(b_1^p, 1), (b_2^p, 2), \ldots, (b_n^p, n)\},$$

where

$$b_1^1 = k_1 + t_1^{a_{\text{init}}(k_1)},$$
$$b_2^1 = k_1 + t_2^{a_{\text{init}}(k_1)},$$
$$\vdots$$
$$b_n^1 = k_1 + t_n^{a_{\text{init}}(k_1)},$$
$$\vdots$$
$$b_j^k = k + t_j^{a_{\text{init}}(k)},$$
$$\vdots$$
$$b_n^p = k_p + t_n^{a_{\text{init}}(k_p)}.$$

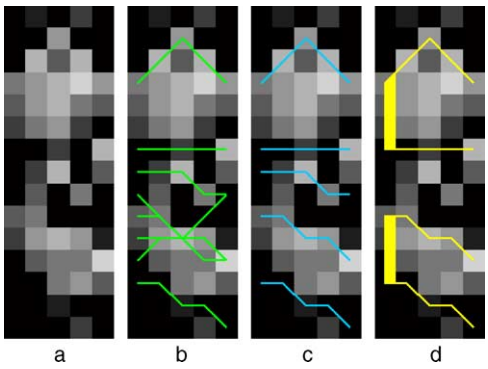

Fig. 16 The illustration of the TEDEBBY operator on examples: (a) the input model image; (b) the BB candidates detected in the first step are displayed as white lines; (c) the BB candidates after compatibility check; (d) the remaining BB candidates coupled by the TOODIS operator application.

Note that since t_1^s is always zero, b_1^k can be interpreted as *the reference position* of the BB candidate $B_{\text{init},k}$.

Example 5. In our example, we have $p = 8$ band boundary candidates:

$$B_{\text{init},1} = \{(4,1),(3,2),(2,3),(3,4),(4,5)\},$$
$$B_{\text{init},2} = \{(7,1),(7,2),(7,3),(7,4),(7,5)\},$$
$$B_{\text{init},3} = \{(8,1),(8,2),(8,3),(9,4),(9,5)\},$$
$$B_{\text{init},4} = \{(9,1),(10,2),(11,3),(10,4),(9,5)\},$$
$$B_{\text{init},5} = \{(10,1),(10,2),(11,3),(11,4),(12,5)\},$$
$$B_{\text{init},6} = \{(11,1),(11,2),(11,3),(12,4),(12,5)\},$$
$$B_{\text{init},7} = \{(12,1),(11,2),(11,3),(11,4),(12,5)\},$$
$$B_{\text{init},8} = \{(13,1),(13,2),(14,3),(14,4),(15,5)\}.$$

The reference positions of these BB candidates are 4, 7, 8, 9, 10, 11, 12, and 13, respectively. Note that these are the indices of the nonzero elements in the vectors $\boldsymbol{m}_{\text{init}}$ and $\boldsymbol{a}_{\text{init}}$ in Example 4. The applied templates are $\boldsymbol{t}^1, \boldsymbol{t}^9, \boldsymbol{t}^7, \boldsymbol{t}^4, \boldsymbol{t}^8, \boldsymbol{t}^7, \boldsymbol{t}^2, \boldsymbol{t}^8$.

Check for compatibility of the band boundary approximations by templates

The adjacent BB candidates $B_{\text{init},k}$ and $B_{\text{init},k+1}$ must not intersect or touch each other. Two adjacent BB candidates satisfying this condition are called *mutually compatible*. We want to check mutual compatibility of all BB candidates. The goal is to build new vectors $\boldsymbol{m}_{\text{comp}}$ and $\boldsymbol{a}_{\text{comp}}$ (and thereby new BB candidates) which are guaranteed to be mutually compatible. The new vectors are built from the initial vectors $\boldsymbol{m}_{\text{init}}$ and $\boldsymbol{a}_{\text{init}}$ by setting some of the nonzero components to zero (this effectively means deleting some BB candidates).

Due to the limited extent of the templates, it is sufficient to check only the compatibility of the BB candidates whose reference positions do not differ by more that $\pm 2S$. If two BB candidates turn out to be mutually incompatible, one of them (preferably the one which is less likely) has to be deleted. The algorithms works like follows:

- The components of the vector $\boldsymbol{m}_{\text{init}}$ are sorted in decreasing order, this permutation yields the vector $\boldsymbol{m}^*_{\text{init}}$; the components of the vector $\boldsymbol{a}_{\text{init}}$ are reordered in accordance with the vector $\boldsymbol{m}^*_{\text{init}}$; the resulting vector is denoted by $\boldsymbol{a}^*_{\text{init}}$.
- The compatibility check starts from the first components of the $\boldsymbol{m}^*_{\text{init}}$ and $\boldsymbol{a}^*_{\text{init}}$; which correspond to the boundary approximation of the most intensive band; this approximation is checked against its neighbors, and components of $\boldsymbol{m}^*_{\text{init}}$ and $\boldsymbol{a}^*_{\text{init}}$, corresponding to the incompatible neighbors, are replaced by zeros, the obtained vectors are denoted by $\boldsymbol{m}^{**}_{\text{init}}$ and by $\boldsymbol{a}^{**}_{\text{init}}$.
- The respective components in $\boldsymbol{m}_{\text{init}}$ and $\boldsymbol{a}_{\text{init}}$ are also replaced by zero.
- We proceed to the next nonzero components of $\boldsymbol{m}^*_{\text{init}}$ and $\boldsymbol{a}^*_{\text{init}}$.
- The check and deletion continue up to the last nonzero component of $\boldsymbol{m}^*_{\text{init}}$.

In this way, the greater components of $\boldsymbol{m}_{\text{init}}$ are preferred. The modified vector $\boldsymbol{m}_{\text{init}}$ obtained after this process is denoted by $\boldsymbol{m}_{\text{comp}}$. The corresponding BB candidates are denoted by B_{comp}.

Example 6. By applying this algorithm to the vectors from Examples 4 and 5, we obtain progressively the following reordered and modified vectors:

$$\boldsymbol{m}^*_{\text{init}} = (26,21,20,17,15,14,13,11,0,0,0,0,0,0,0,0)',$$
$$\boldsymbol{m}^{**}_{\text{init}} = (26,21,0,17,0,0,13,11,0,0,0,0,0,0,0,0)',$$
$$\boldsymbol{m}_{\text{comp}} = (0,0,0,26,0,0,17,13,0,21,0,0,11,0,0)',$$
$$\boldsymbol{a}^*_{\text{init}} = (1,8,2,9,7,4,7,8,0,0,0,0,0,0,0,0)'$$
$$\boldsymbol{a}^{**}_{\text{init}} = (1,8,0,9,0,0,7,8,0,0,0,0,0,0,0,0)'$$
$$\boldsymbol{a}_{\text{comp}} = (0,0,0,1,0,0,9,7,0,8,0,0,8,0,0)'.$$

The reduced set of BB candidates obtained after compatibility check, which is represented by the vector $\boldsymbol{a}_{\text{comp}}$, is illustrated in Fig. 16c.

4.8.2. TOODIS operator

In the algorithms of the stage I described up to now, we have maximized the number of detections of true band boundaries at the expense of false detections. Therefore, in the set of BB detections corresponding to the vector $\boldsymbol{m}_{\text{comp}}$ false BB detections (candidates) may be comprised. A greater part of false BB candidates can be deleted automatically in the stage I already. This can simplify the necessary user interaction in the stage II. The proposed algorithm TOODIS (tooth discrimination) enables such a deletion of false BB candidates in two steps.

The first step

The BB candidates $B_{\text{comp},k}$ resulting from the above-mentioned procedure and represented by the vector $\boldsymbol{m}_{\text{comp}}$ divide the lane image into disjoint pixel regions which can constitute subregions of the background, or subregions of a band. Let us denote by R_{k-1} the region which is bounded by the boundary candidates $B_{\text{comp},k-1}$ and $B_{\text{comp},k}$. The homogeneity property of the background intensities in gel images (measured by statistical characteristic *sample median*) can be used for further automatic deletion of false boundary candidates. To avoid influence of neighboring structures (namely, high-contrast bands with curvilinear boundaries) on medians, the inter-boundary regions are splitted into

subregions by a digital curve C defined in the following way. For each column index $j = 1, 2, \ldots, n$ and two subsequent BB candidates $(k-1), k$, we calculate the index $l_j = [(b_j^{k-1} + b_j^k)/2]$. Then, the *splitting curve* C is represented by the set of index pairs $\{(l_1, 1), (l_2, 2), \ldots, (l_n, n)\}$. This curve divides the region R_{k-1} between the BB candidates $B_{\mathrm{comp},k-1}$ and $B_{\mathrm{comp},k}$ into two approximately symmetric subregions $S_{k-1,u}$ and $S_{k-1,l}$, where u denotes the upper subregion, l denotes the lower subregion and $R_{k-1} = S_{k-1,u} \cup S_{k-1,l} \cup C$, where $S_{k-1,u}, S_{k-1,l}$, and C are mutually disjunctive. Let us consider now two neighboring regions R_{i-1} and R_i, bounded by the BB candidates $B_{\mathrm{comp},k-1}, B_{\mathrm{comp},k}$, and $B_{\mathrm{comp},k+1}$. Thus, we have four subregions for which the medians $\mathrm{med}(S_{k-1,u}), \mathrm{med}(S_{k-1,l}), \mathrm{med}(S_{k,u})$, and $\mathrm{med}(S_{k,l})$ are calculated. (*Note*: for a statistic ''stat'' and a region S the symbol $\mathrm{stat}(S)$ means: $\mathrm{stat}(S) = \mathrm{stat}(\{g(i,j)\})$, where $(i,j) \in S$.) If for neighboring subregions $S_{k-1,l}$ and $S_{k,u}$ from two different regions R_{k-1} and R_k, $\mathrm{med}(S_{k-1,l}) = \mathrm{med}(S_{k,u})$, it is assumed that these subregions are subsets of one homogeneous region (background or band) and the i th BB candidate is deleted. This procedure is repeated for all BB candidates in $\boldsymbol{m}_{\mathrm{comp}}$. The resulting candidate set is represented by the vector $\boldsymbol{m}_{\mathrm{med}}$.

The second step

After automatic deletion of the essentially unacceptable BB candidates (detections) in the first step, the user sets in the graphical mode the flexible threshold ''*thr*'' to some local maximum of the values of the TEDEBBY operator, thereby excluding all such components of the vector $\boldsymbol{m}_{\mathrm{med}}$ which are smaller or equal to the threshold ''*thr*''. The layout of this graphical mode for the user interaction is shown in Fig. 14. The resulting vector $\boldsymbol{m}_{\mathrm{med}}^{\mathrm{thr}}$ is obtained. The ultimate goal of the TOODIS operator is to couple the appropriate BB candidates. It operates on the vectors $\boldsymbol{m}_{\mathrm{med}}^{\mathrm{thr}}$ generated by the particular values of the threshold ''*thr*''. For each set $\{B_{\mathrm{comp},k}\}_{\mathrm{thr}}$, of the BB candidates, represented by the vector $\boldsymbol{m}_{\mathrm{med}}^{\mathrm{thr}}$, that corresponds to the threshold ''*thr*'', a set of image regions R_k, each bounded by two adjacent BB candidates $B_{\mathrm{comp},k}$ and $B_{\mathrm{comp},k+1}$, $k \in \{1, 2, \ldots, m-1\}$ can be considered. These regions can again be characterized by appropriate statistical characteristic $\boldsymbol{f}$ of region intensities $f(R_k)$ and the entire lane image by the vector $\boldsymbol{f}_{\mathrm{lane}} = (f(R_1), f(R_2), \ldots, f(R_p))$, for $p \leq m-1$. The coupling of neighboring BB candidates is based on the following assumption. Let us denote the three adjacent components (triple) of the vector $\boldsymbol{f}_{\mathrm{lane}}$: $f(R_{k-1}), f(R_k), f(R_{k+1})$. Provided band image structures are brighter than background (this is the standard situation in gel image analysis) and the relation $f(R_{k-1}) < f(R_k) > f(R_{k+1})$ is true, the central region is assumed to be brighter than two neighboring regions. So, we interpret the region R_k as a band surrounded by background and call the component $f(R_k)$, a *positive tooth*. Again, to avoid influence of neighboring image structures we do not apply this scheme to the regions R_k directly. Instead, we decompose them, similarly as in the previous step, into two approximately symmetrical subregions $S_{k,u}$ and $S_{k,l}$ and select the characteristic of one of these subregions as representative for the whole region R_k. The medians $\mathrm{med}(S_{k,u})$ and $\mathrm{med}(S_{k,l})$ for these subregions are calculated. We can use minimum or maximum in each median pair as a representative value for the whole region R_k. Since we maximize the detection of true band boundaries (at the expense of additional false detections, which will be ultimately deleted in the stage II), the following scheme for representative median selection has been proposed. In the following, we abbreviate $m_{ku} = \mathrm{med}(S_k, u)$ and $m_{kl} = \mathrm{med}(S_k, l)$ and we denote $f_{k-1} = \min\{m_{k-1,u}, m_{k-1,l}\}$, $f_k = \max\{m_{ku}, m_{kl}\}$, and $f_{k+1} = \min\{m_{k+1,u}, m_{k+1,l}\}$. Then, for all relevant k, we:

- generate a triple of values (f_{k-1}, f_k, f_{k+1}) for the central position k;
- check, whether $f_{k-1} < f_k > f_{k+1}$;
- if so, the BB candidate pair $B_{\mathrm{comp},k}$, $B_{\mathrm{comp},k+1}$, corresponds to a positive tooth and the next triple is generated for the central position $k+2$;
- if not, the BB candidate pair $B_{\mathrm{comp},k}$, $B_{\mathrm{comp},k+1}$ does not correspond to a positive tooth and the next triple is generated for the central position $k+1$.

4.8.3. Operators MAMBO, LOBBY and DSS

The main part of the false detection rejection is done implicitly by the TOODIS operator before starting the stage II. However, due to fluctuations of background intensity homogeneity in lanes (often occurring in imperfect gel images), we have to admit some false detections also in pairs of detections which have been accepted for the stage II. So, the basic purpose of the BB indicators in this stage is to characterize a degree (probability) of the given curve to be a boundary of true band. If all BB candidates are increasingly ordered with respect to this degree, a process of checking and removing false BB candidates can be organized within the stage II efficiently. Again, it is a semi-automatic procedure in which the BB candidates are subsequently displayed to the user. The user can decide, if it is necessary to continue the procedure or he (she) can accept all remaining detections at once. The stage II is im-

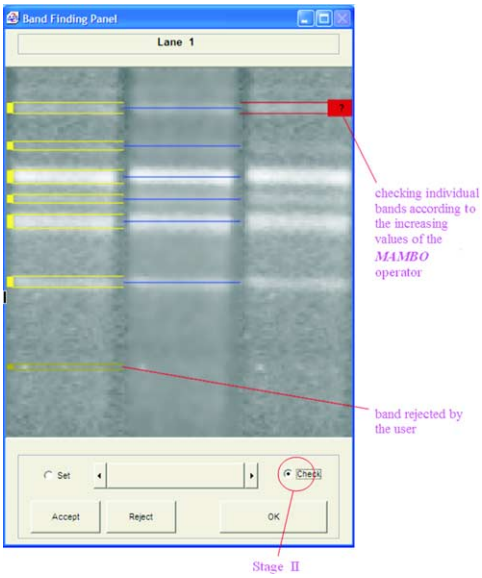

Fig. 17 The GelMaster window in stage II. Left pane: original lane with finally accepted band boundaries; middle pane: graphical information on representative positions of the detected bands; right pane: suspected band boundaries displayed to the user for final checking.

plemented in the BFP panel which is switched over to the ''Check'' mode (see Fig. 17). In this mode, the rightmost graphical pane is replaced automatically with the replica of the original lane image. A marker with a question mark overlaid in this pane points to the actually checked band position. The user can reject the detection under question by corresponding button or he (she) can accept such a BB detection. The goal of this interactive phase is to finally reject all false BB detections. In most cases, there is no need to reject BB detections, since only true band boundaries have been detected. In such cases, immediate decision of the user to abandon the ''Check'' mode is executable by pressing the ''OK'' button. Usually in case of some false BB detections present, there is a very short way how to reject them, because all detections are ordered according to the specifically designed operators yielding a probability of false detections. We will give a short explanation of the construction of these operators which are actually second order indicators of the band boundaries.

We have proposed three improved BB indicators which are based on different measures of the intensity homogeneity applied to two horizontally oriented adjacent regions defined by an estimate of their common curvilinear boundary. For each individual band boundary from a set of all paired detections, a rectangular pixel neighborhood is defined. The size of the neighborhood is limited by the type of templates used (maximum slant or curvature). The array of the magnitude of intensity gradients (shortly gradients) is calculated for this neighborhood. We assume that any band boundary can be represented as a simple curve that intersects each column of the rectangular neighborhood just in one point. Therefore, for each column of the neighborhood, the position of the maximum gradient value represents a reasonable estimate of a band boundary point. If the given detection represents a false boundary position, the gradient maxima in each column are distributed randomly according to the random fluctuations of the background. A set of these points constitutes an improved estimate of the whole band boundary which separates the neighborhood into two disjunctive regions. Statistical and other characteristics of these regions (or digital boundary curves) may serve for measuring the difference between the background and band intensity distributions. In case of a false detection, this difference will obviously be small because both regions belong to the background (or to the band).

The first indicator MAMBO (median above median below) based on the improved curvilinear estimate of the band boundary is defined as follows. We propose to compute the intensity median for all the neighborhood pixels above the boundary and intensity median below it, and to define the indicator B simply as the absolute difference of the medians.

The second indicator LOBBY (length of band boundary) based on the identical boundary estimate is defined as a measure of curve continuity by means of digital length of the boundary. A horizontal straight line is the shortest boundary which is used for normalization. Randomly distributed points of gradient maxima in columns yield larger values of the LOBBY indicator.

Stage II requires ordering of the obtained values of the indicators MAMBO and LOBBY with preserved information on detection indices. Smaller values of these indicators indicate smaller difference between the two regions (above and below the boundary), i.e., higher probability of the background or band interior that in turn means higher probability of a false detection. The user can check all detections starting from the least indicator value up to the last potentional false detection in the list. The total number of searched (checked) values (positions) up to the last false detection (N^{I}_{searched} for the indicator I) is a measure of potential effort of the user.

The MAMBO and LOBBY indicators differ in the ability to concentrate their values belonging to false detections in lower part of the range. An ideal discrimination is reached when the maximum indicator value calculated for false detections (from

the stage I) is smaller than all values obtained for true detections. It is not known a priori which of the two indicators would produce smaller number $N^{I}_{searched}$ of searched values in the ordered list. However, based on computer experiments we can summarize that in approximately 50% of cases one of the indicators gives faster results, i.e., the list of searched values necessary for rejecting all false detections was shorter than for the second indicator. In 50% cases, the situation is reverse. If we decide to use just one indicator, it means that in 50% of cases this choice will be worse than the other. We can represent the ordered lists of indicator values calculated for all detections as two number stacks. Let us consider an algorithm that uses the values from both stacks in a combined search for which the number of ''averaged'' searched positions is denoted as $N^{*}_{searched}$. If, in average, $N^{*}_{searched} < N^{I}_{searched}$, and $N^{*}_{searched} < N^{II}_{searched}$, then this algorithm outperforms both indicators.

The double stack strategy (DSS) is based on this idea. It has been implemented and successfully tested on real gel image data. Its construction was based on an assumption that the ordered indicator values corresponding to false and true detections constitute ties. It means that if a false detection is encountered in one stack, it is reasonable to continue the search in the same stack. On the contrary, if a true detection is encountered, it is more advantageous to switch to the other stack and then to continue searching. The main features of the DSS algorithm can be described as rules according which the search procedure is switched from one stack to the another:

1. The searching procedure can start from any of the two stacks.
2. The element chosen in one stack is automatically excluded also from the other stack.
3. If the given element in one stack proved to be ''true'' searching is switched to another stack.
4. If the given element is declared as ''false'' we proceed in the same stack.

We have tested these three BB indicators on real image data with two basic goals:

- To quantitatively characterize the efficiency of the individual operators and to find out whether it exists statistically significant difference between them.
- To collect the ground truth data (accepted and rejected generic detections, and additional true detections) from an expert.

The testing itself was performed using an interactive software tool which we developed specifically for this purpose. Without using any numerical parameters of the image analysis, the expert (biologist) set a subjectively optimal response of the band boundary operators in each lane separately. Then, based on visual analysis, he marked all false detections and added all missing detections (called *extra* detections). All numerical parameters necessary for the subsequent statistical evaluation have been recorded automatically by the software. We denote N_{false} as the number of false detections found by the expert in the set D of all detections generated by TEDEBBY.

It is reasonable to characterize the theoretical limits of the improved BB indicators which could be reached by an individual operator without preceding application of the TOODIS operator. In other words, we consider the values of the operators for all false detections generated by the TEDEBBY detector. After having ordered all operator values in increasing order, the values corresponding to false detections are intermixed with those for true detections. In such a case, to get rid of all false detections in the stage II, we need to search for all values starting from the minimum up to the last (greatest value) belonging to a false detection. Then, the quality of the operator can be characterized by an effort (a number of steps) needed for achieving this state. This can be measured by a coefficient $coef = (N_{searched} - N_{false})/N_{false}$, where $N_{searched}$ is the number of all values of the given indicator which have to be checked for false detection presence. In our test study [14], we have calculated this characteristic for each of the operators in test. The operator is considered to be better, if the mean value (mean(coef)) is closer to zero (zero is obviously the ideal case).

To get an insight into differences between three tested indicators a statistical comparison was of interest. We characterized these differences by means of nonparametric testing of statistical hypotheses. For unknown distributions of random quantities $coef_M$ (for MAMBO), $coef_L$ (for LOBBY), and $coef_{DSS}$ (for DSS), the Kolmogorov–Smirnov [22] nonparametric statistical test is suitable. The null hypothesis is formulated for each pair of these indicators as: ''*there is no difference between distribution functions of the quantities* $coef_M$, $coef_L$, *and* $coef_{DSS}$''.

The testing criterion (of our null hypothesis) is defined for samples taken from the tested distributions of coef, namely for the empirical distribution functions (cumulative histograms) $H_M(x)$ and $H_L(x)$ as a maximum difference: $D_{H-L} = \max_x |H_M(x) - H_L(x)|$. For a sample larger than 40 elements, the critical values can be approximated by the following formula (e.g., for a significance degree $\alpha =$

0.05):

$$D_{0.05} = 1.36\sqrt{\frac{n_1 + n_2}{n_1 n_2}}, \quad (6)$$

where n_1 and n_2 denote the numbers of elements in the sample sets. The null hypothesis is rejected, if $D_{M-L} \geq D_{0.05}$. In our evaluation study [14], we used 10 digitized gel images of one class (ethidium bromide). One hundred and sixty-eight lanes in total have been processed, in which 3280 generic band boundaries have been detected. Eighty-six percent of these detections were true and 14% were false. The expert added 290 extra detections (9% of true detections). For these detections, the values of coef_M (for MAMBO), coef_L (for LOBBY), and coef_{DSS} (for DSS) have been calculated. Based on cumulative histograms of the obtained coefficients, the values of the Kolmogorov–Smirnov statistics have been calculated for the pairs MAMBO–LOBBY (D_{M-L}), MAMBO–DSS (D_{M-DSS}), and LOBBY–DSS (D_{L-D}). For 110 elements of the sample and significance value $\alpha = 0.05$, the critical value $D_{0.05} = 0.1348$ has been calculated. The following values of the Kolmogorov–Smirnov statistics have been obtained: $D_{M-L} = 0.1000$, $D_{M-DSS} = 0.1364$, $D_{L-DSS} = 0.2000$. In case of pairs MAMBO–DSS and LOBBY–DSS, the values of the statistics obtained are greater than the critical value, and therefore the null hypothesis is rejected. This implies the mutual differences between the DSS operator and the other two operators are statistically significant. Based on this finding and the minimum mean value, we conclude that the DSS algorithm is significantly better than the other two algorithms.

4.9. Calibration of molecular weights

Calibration is the final principal stage of the gel image analysis. The aim of the calibration is to assign a priori known representative values such as molecular weights or number of base pairs to reference bands located in the so-called standard lanes (this quantity is called fragment size). Once this assignment has been made, all other unknown bands with mobility distances identified in the previous GIA stages can be assigned corresponding values of the fragment size. The assignment is calculated using approximation of the mobility distance function under the assumption that the mobility of fragments of the same size is the same in all lanes of one gel. Only bands with mobility distances from the range given by bands of a standard lane can be processed, i.e., no extrapolation of the approximation function is considered.

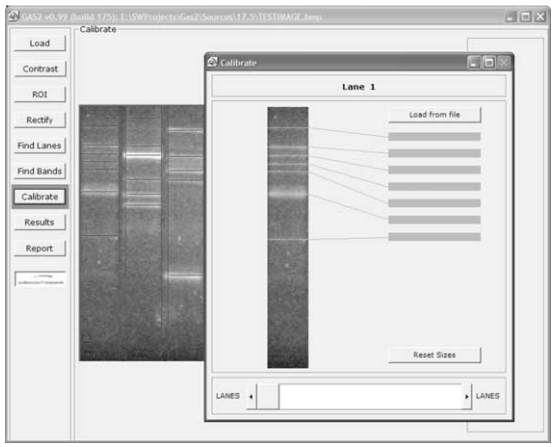

Fig. 18 The GelMaster calibration window. The user types the size and amount values for standard lanes in the corresponding boxes.

The GelMaster also allows for calibration and computation of *fragment amount*, the quantity that represents the amount of substance deposited in one band. The underlying assumption is that this quantity is reflected by the integral brightness of the band and that two equally bright bands (with the brightness measured as area integral of the pixel intensities in the band box) feature equal amount of the deposited substance (DNA fragments). Thanks to the template approach developed for band boundary detection the shape of each individual band is available. This makes the calculations of the *fragment amount* more accurate.

In Fig. 18, a calibration panel is displayed. This panel appears after pressing the ''Calibrate'' button. The default position is the first lane (which is usually the standard lane). The scrollbar in the bottom part of the panel serves for transition to another lane. The panel contains a lane identification label on the top similar to the BFP. It shows an image of the given lane overlaid by band detection markers. The representative positions of the bands are taken from the previous band finding stage. A rectangular textbox is attached to each band. At the beginning of the calibration mode, the textboxes are empty. The GelMaster offers two options how to enter the values of fragment sizes or amounts: manual editing or loading from a file.

4.10. Displaying results

Final results of the gel image analysis are displayed (Fig. 19) after pressing the ''Results'' button in the left panel of the main menu. The main im-

Fig. 19 The positions of the detected bands are represented graphically as horizontal bars.

age display is overlaid with green and red band bar-markers. The red markers represent the reference bands, the green markers display the normal (computed) values. A small panel with information about fragment size and fragment weight (amount) is displayed at the right of the cursor when prompted.

4.11. Generating reports

The information acquired during the process of gel image analysis can be either printed out or recorded in files, which can in turn be used for further work, e.g., as an input for another software package or for documentation. There are two options: a simple ASCII text file and a graphical report file. There is a possibility to print, or save the report in EPS or EMF formats. The main purpose of this kind of report is to generate illustration of the gel image analysis process which can be either printed out directly or edited in a graphical editor and imported into another document such as research report or paper. An example of the graphical report is given in Fig. 20.

5. Status report

In the course of development of the GelMaster system a number of tests have been carried out by biologists in laboratory environment. The performance of the system has been permanently evaluated and detailed comments from several users

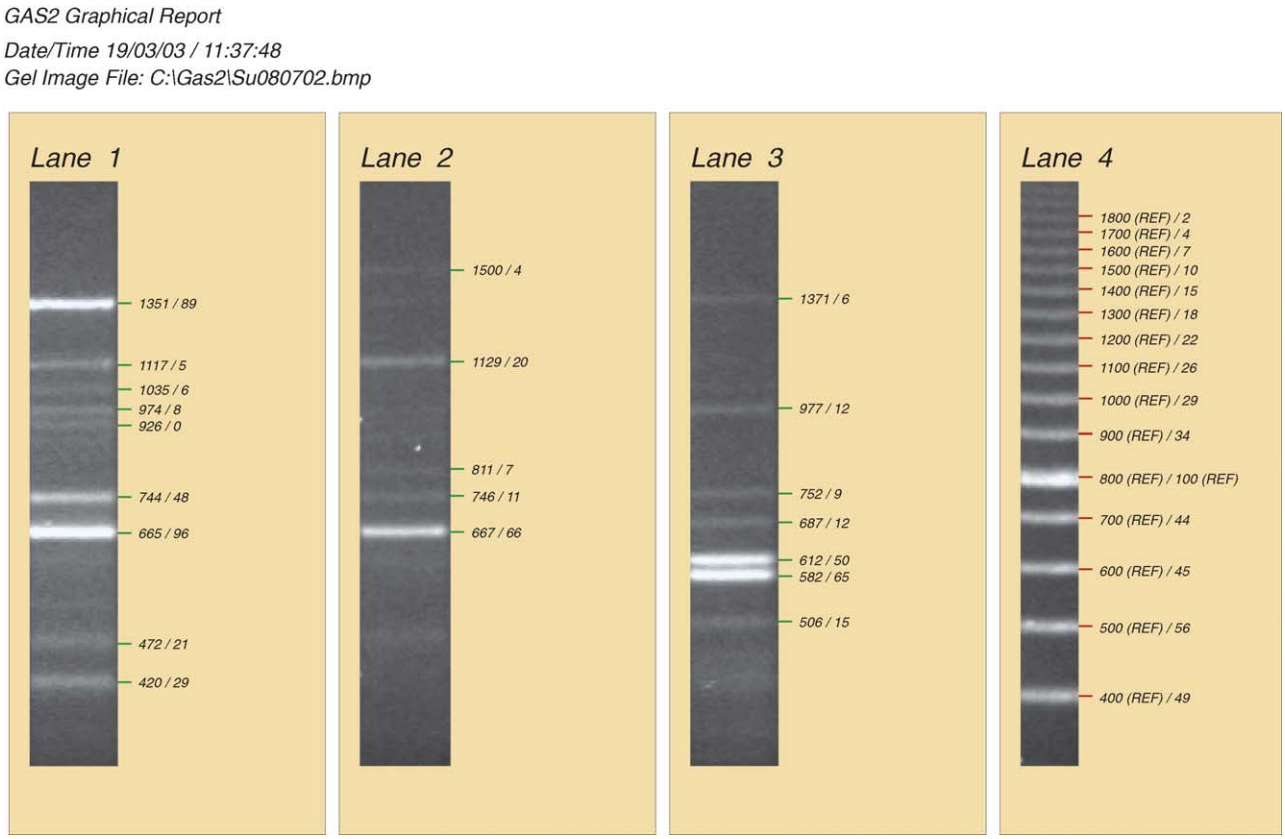

Fig. 20 The report example.

have been collected. In particular, scrupulous comparative tests between GelMaster and one of the mentioned commercial gel image analysis systems have been performed. Average analysis time (excluding a part of calibration which was not crucial) for test images with approximately 20 lanes and 10 bands per lane is two to three times shorter in case of the GelMaster system. Based on this feedback many modifications, improvements and add-ons to the system have been done. In 2002, we reported in our paper [23], the status of the system development including the improvements achieved by the template approach to the band detection problem. The GelMaster system has been used in routine laboratory practice in ARC Seibersdorf research GmbH since September 2003.

The general conclusion of the users is that even with bad quality gel pictures, lane and band detection is fairly easy—needing little interaction. The band detection method using a slider for moving threshold is easy to understand and straightforward (that is not the case in commercial systems). Controlling image analysis using keyboard or mouse (without the burden of adjusting numerical parameter values in other systems) facilitates the work a lot. GelMaster was distributed to selected beta testers in Europe in January 2004. We expect to collect feedback by the end of 2004.

Minimal system requirements: PC 300 MHz, 64 RAM, 5 MB disk space, operating system Win98, Win2K or WinXP.

6. Lessons learned

Based on the theoretical and simulation research and using the pilot implementation GAS1, the GelMaster system has been developed. The developed know-how comprises a number of novelties in each of the three important parts of a complex approach: (i) philosophy of analysis, (ii) mathematical solutions, and (iii) image data processing algorithms, which can be summarized as follows:

- Preprocessing of gel images by geometry-driven diffusion filters.
- Modified approach to rectification (correction) of geometrical distortions.
- Genuine algorithm for lane separation.
- Replacement of one-dimensional paradigm of band detection by two-dimensional paradigm.
- Replacement of the band detection principle based on density curve peak identification by the detection of band boundaries using the band boundary template approach.
- The band detection task has been splitted explicitly into two different subtasks (stages): I—semi-automatic detection of all true band boundaries and automatic rejection of the main part of false detections; II—semi-automatic checking of the band boxes for false detections.
- Particular algorithms have been developed for both band detection stages: I—TEDEBBY, TOODIS; stage II—MAMBO, LOBBY, DSS.
- Development of the original GUI for band detection which includes three vertical copies of the given lane located in separate panes of the window (each having specific function in user interaction within two stages).
- The control of band detection algorithms is implemented exclusively by graphical objects; no numerical parameters are used.

7. Future plans

The GelMaster software system for semi-automatic gel image analysis proved to be fully acceptable for everyday usage in molecular biology labs. A number of tests have been carried out with various types of real gel images which suffered by different drawbacks. Based on the encouraging results achieved, we intend to prepare a commercial product based on this system. The know-how used in the design of the GelMaster has been included in two proposals for the Austrian patents (1026/2002, 1701/2002). There are further possibilities how to modify the individual functions of the GelMaster to serve for specific applications of the GIA like denaturated gradient gel electrophoresis (DGGE), antidoping testing for recombinant EPO presence based on isoelectric focusing in gels, etc.

References

[1] A.T. Andrews, Electrophoresis. Theory, Techniques, and Biochemical and Clinical Applications, Clarendon Press, Oxford, 1986.
[2] A. Chrambach, M.J. Dunn, B.J. Radola (Eds.), Advances in Electrophoresis, vol. 1, VCH Verlagsgesellschaft, Weinheim, 1989.
[3] P. Richards, Protein electrophoresis, in: R. Rapley, J.M. Walker (Eds.), Molecular Biomethods Handbook, Humana Press, Totowa, NJ, 1998, pp. 413—462.
[4] D. Tietz (Ed.), Nucleic Acid Electrophoresis, Springer-Verlag, Berlin, 1998.
[5] D.-P. Häder (Ed.), Image Analysis, Methods and Applications, second ed., CRC Press, Boca Raton, 2001.
[6] J.E. Celis, R. Bravo, Two-Dimensional Gel Electrophoresis of Proteins, Academic Press, New York, 1984.
[7] A. Görg, W. Postel, S. Günther, The current state of two-dimensional electrophoresis with immobilized pH gradients, Electrophoresis (1988) 531—552.

[8] P.M. Nugues, Two-dimensional electrophoresis image interpretation, IEEE Trans. Biomed. Eng. (1993) 760—767.
[9] D.-P. Häder, Computer-assisted image analysis in biological sciences, Ind. Acad. Sci. (Plant Sci.) (1988) 227—238.
[10] Z.A.M. Boniszewski, J.S. Compley, B. Hughes, C.A. Read, The use of charge-coupled devices in the quantitative evaluation of images, on photographic film or membranes, obtained following electrophoretic separation of DNA fragments, Electrophoresis (1990) 432—451.
[11] J. Vohradsky, J. Panek, Quantitative analysis of gel electrophoretograms by image analysis and least squares modeling, Electrophoresis (1993) 601—610.
[12] G.W. Horgan, C.A. Glasbey, Uses of digital image analysis in electrophoresis, Electrophoresis (1995) 298—309.
[13] P.Ch. Hansen, Numerical aspects of deconvolution, Lecture Notes, Department of Mathematical Modelling, Technical University of Denmark, http://www.imm.dtu.dk/~pch).
[14] I. Bajla, I. Holländer, K. Burg, Improvement of electrophoretic gel image analysis, Technical report OEFZS-IT—0008, ARC Seibersdorf research GmbH, Seibersdorf, 2001.
[15] I. Bajla, I. Holländer, K. Mayer, Algorithms for degradation correction and band detection in electrophoretic gel images, Technical report ARC-IT—0049, ARC Seibersdorf research GmbH, Seibersdorf, 2002.
[16] I. Bajla, I. Holländer, Nonlinear filtering of magnetic resonance tomograms by geometry-driven diffusion, Mach. Vis. Appl. (1998) 243—255.
[17] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Mach. Intell. (1990) 629—639.
[18] F. Catté, P.-L. Lions, J.-M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal. (1992) 182—193.
[19] J. Weickert, Anisotropic Diffusion in Image Processing, Teubner, Stuttgart, 1998.
[20] R.C. Gonzales, P. Wintz, Digital Image Processing, second ed., Addison-Wesley, Reading, MA, 1987.
[21] I. Bajla, I. Holländer, K. Burg, A study on DNA gel image analysis improvement, in: Proceedings of the Third International Conference on Measurement, Smolenice, Slovak Republic, 14—17 May 2001, pp. 223—226.
[22] P.K. Sen, M.L. Puri, Nonparametric Methods in Multivariate Analysis, Wiley, New York, 1971.
[23] I. Bajla, I. Holländer, K. Burg, S. Fluch, A novel approach to quantitative analysis of electrophoretic gel images of DNA fragments, in: Proceedings of the IEEE International Symposium on Biomedical Imaging, Washington, USA, 7—10 July 2002, pp. 899—901 (CD-ROM).