John S. Gustafsson[1]
Anders Blomberg[2]
Mats Rudemo[1]

[1]Department of Mathematical Statistics, Chalmers University of Technology
[2]Department of Cell and Molecular Biology – Microbiology, Göteborg University, Sweden

Warping two-dimensional electrophoresis gel images to correct for geometric distortions of the spot pattern

A crucial step in two-dimensional gel based protein expression analysis is to match spots in different gel images that correspond to the same protein. It still requires extensive and time-consuming manual interference, although several semiautomatic techniques exist. Geometric distortion of the protein patterns inherent to the electrophoresis procedure is one of the main causes of these difficulties. An image warping method to reduce this problem is presented. A warping is a function that deforms images by mapping between image domains. The method proceeds in two steps. Firstly, a simple physicochemical model is formulated and applied for warping of each gel image to correct for what might be one of the main causes of the distortions: current leakage across the sides during the second-dimensional electrophoresis. Secondly, the images are automatically aligned by maximizing a penalized likelihood criterion. The method is applied to a set of ten gel images showing the radioactively labeled proteome of yeast *Saccharomyces cerevisiae* during normal and steady-state saline growth. The improvement in matching when given the warped images instead of the original ones is exemplified by a comparison within a commercially available software.

1 Introduction

Two-dimensional (2-D) gel electrophoresis is still the major technique in quantitative global protein expression analysis. A crucial step in this technique is image analysis where the protein spot pattern in different images are to be matched. Two main steps are common in the current methodology of 2-D gel image analysis: an initial spot detection and quantification step, and a subsequent matching step [1]. Both these steps are nontrivial and often need time-consuming manual interference. The matching step is a bottleneck in the gel analysis which has consequences for a comprehensive statistical analysis [2], especially in large-scale proteomics. One main cause that pushes the semiautomatic matching algorithms to their limits is geometric distortions of the protein patterns inherent to the casting, polymerization and running procedure of the gels.

Image warping is a function which deforms images; all positions in one image plane are mapped to positions in another plane. It has many applications in the field of image analysis [3]: for example, to remove optical distortions caused by camera lenses or viewing perspective, to register an image with a template, or to bring two or more images into alignment. In the registration and alignment cases, the choice of warp is a compromise between achieving a good match and introducing a smooth deformation. In the analysis of 2-D electrophoresis gel images, the warping technique has mainly been used to align images for visual comparison of gels [4–7]. In these examples, the smoothness of the warp is ensured with the usage of smooth functions, such as low-degree polynomials or thin-plate splines, and the function is specified to fit a set of manually added landmarks. Another example related to 2-D gels is a warping method to remove geometric distortions, so-called "smiles" and "frowns", in multitrack electrophoresis gels [8].

Image warping can also be used for automatic alignment of 2-D gel images. In the common spot matching approach the data in the images are initially reduced to lists of spot features in the detection step, and one of these features, the spot center, is used in the following matching step. With a pixel-based alignment method one is able to use all data in the gel, such as spot shape, streaks and reproducible background features. In an early example of this warping approach Conradsen and Pedersen [9] apply linear transformation at a progressively finer series of scales, to minimize the sum of squared pixel-by-pixel differences. However, no smoothness constraints are considered. A recent contribution to the methodologies in 2-D gel analysis that uses the warping technique for image alignment is the one presented by Smilansky

Correspondence: John S. Gustafsson, Department of Mathematical Statistics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden
E-mail: johng@math.chalmers.se
Fax: +46-31-7723508

[10]. A sequence of covering sub-rectangles in the gel image is constructed such that each sub-rectangle contains a small number of spots. A pixel-based alignment is then applied for each sub-rectangle, and the shift vectors of these sub-image alignments are combined into a global warping using Delaunay triangulation. The similarity measure in the sub-image alignment is made up by several components based on both contrast and gradient information in the images. Too rough distortions are avoided by restrictions when changing the shift vectors in the iteration loop.

We here present another warping method to correct for smooth geometric distortions of the spot patterns in 2-D gels, which applied prior to the common matching procedure should reduce problems due to these distortions. The method proceeds in two steps. Firstly, a warping of each gel image is applied to correct for what might be one main cause of distortion: current leakage along the sides of the gel in the second-dimensional gel electrophoresis. The effect of current leakage on the electric field in the gel and its consequences for the protein pattern is investigated with a simple physicochemical model. In the case of multitrack gels, it is possible to use the fact that bands are asymmetric to estimate orientation directions, and then to formulate a mathematical model of the warping function [8]. It is harder to extract similar directional information in 2-D gels. An easily identifiable object with a direction in a 2-D gel is the gel front at the bottom of the gel. In the presence of current leakage, the gel front is deformed from a straight line. We use the curve form of the gel front to estimate the amount of current leakage.

Secondly, the current leakage corrected images are automatically aligned to correct for other sources of geometric distortions. There are many approaches to image alignment [3]. We have adopted a statistical formulation proposed by Glasbey and Mardia [11] where the compromise between achieving a good match and introducing rough distortions is formulated as to maximize a penalized likelihood. The likelihood measures the similarity between aligned images and the penalty is a distortion criterion, uniquely minimized by a null-set of functions. We use a simple Gaussian model for the warped image after initial logarithmic intensity transformations of the images, and as distortion criterion we choose one that as null-set has the set of affine functions. Within this unifying statistical framework one has the flexibility to consider other image models and null-sets in the future.

The method is applied to a set of ten 2-D polyacrylamide gel electrophoresis (2-D PAGE) images from an analysis of protein expression changes in the yeast *Saccharomyces cerevisiae* during steady-state saline growth (Alipour *et al.,* unpublished). The protein pattern has already

been manually matched in this analysis and the list of matched spot centers serves as reference in a geometric similarity measure of the protein patterns in the gel images. Finally, the improvement in matching when given the warped images instead of the original ones is exemplified by comparison within a commercially available software.

2 Materials and methods

2.1 Image warping

A warping function is a function that maps positions $x = (x_1, x_2)$ in one image domain to positions $f(x) = (f_1(x), f_2(x))$ in another image domain. Let Y be an image and let Y_x denote the pixel value at position x. The warped version of Y under f, denoted Y_f, is specified by

$$(Y_f)_x = Y_{f(x)}, \text{ for } x \in X \tag{1}$$

where $X = \{1, \ldots, n_1\} \times \{1, \ldots, n_2\}$ is the domain of Y_f. This means that the pixel value at position x in the warped image Y_f is defined to be the value at position $f(x)$ in the original image Y. If $f_1(x)$ and $f_2(x)$ are not integers, we use bilinear interpolation to calculate the pixel value and if $f(x)$ lies outside the domain of Y we define $Y_{f(x)}$ to be a constant value. Ideally this constant value would be zero, no protein contents, but we use an estimate of the background level defined as the median value of the right and left-most edges of Y.

2.2 Gel images

The ten gel images are from an analysis of protein expression changes in the yeast *Saccharomyces cerevisiae* during steady-state saline growth (1 M NaCl) (Alipour *et al.,* unpublished). Normal and saline growth conditions were replicated in five individual samples. The proteins were radioactively labeled with [^{35}S]methionine during growth. Protein samples were extracted from harvested cells and 2×10^6 DPM applied per gel. The 2-D PAGE was run on an Investigator™ system (Oxford Glycosystem Ltd) as described in [12]. In brief, each protein sample was subjected to isoelectric focusing in a gel strip containing an immobilized pH gradient (pH 3–10). The second-dimensional polyacrylamide gels were cast between two glass plates separated from each other by two thin plastic spacers glued between the plates along each vertical side. The width of the plastic spacers was $\Delta w = 25$ mm, and the gel width and height were $w = 235$ mm and $h = 215$ mm. The first-dimensional strips were equilibrated in a sodium dodecyl sulfate (SDS) buffer and mounted on the SDS-PAGE slab gels. The second-dimensional gel electrophoresis was run on a vertical system at a limiting power

of 16 GW (maximum voltage 500 V) and with gels fully submerged in running buffer for efficient cooling. Immediately after the run the gels were dried on filter paper. Dried gels were exposed to phosphor image plates, which were scanned in a PhosphorImager™ (Molecular Dynamics, Sunnyvale, CA, USA) with a pixel size of 176 × 176 μm and stored as 16-bits TIFF-images. The images were processed as described in [13] and 1028 spots were manually matched in all ten gel images. The gels were numbered such that gels 1–5 were from normal growth conditions and gels 6–10 were from steady-state saline growth.

2.3 Warping step I: current leakage correction

Despite the wide application of SDS-PAGE considerable uncertainty still exists about the migration behavior of the SDS-protein complexes [14, 15]. In an analysis of electrophoretic transport in porous media, of which SDS-PAGE is an example, four major factors have to be considered: (i) the structure of the media, the polyacrylamide net; (ii) the characteristics of the transporting solute, the SDS-protein complexes; (iii) the solvent conditions, the buffer; and finally (iv) the nature of the electric field. The role of the gel and the size and shape of the SDS-protein complexes are the factors considered in the two dominant models of gel migration of SDS-protein complexes: the standard Ogston-Morris-Rodbard-Chrambach model [16] and the reptation model [17] – reptation as in reptile-like motion. These two factors are also the ones investigated in a recent contribution to the general gel electrophoresis theory [18]. Buffer conditions in different electrophoresis techniques are modeled in [19]. The effect of the geometry and electric conductivity of gel fibers on the local electric field and its consequences for the effective mobility has recently been considered [20, 21]. The major factor in the geometric distortions of the protein pattern due to current leakage is a global change of the electric field. Therefore, we focus on the global electric field in our model and a simple approach is applied to the other three of the four factors listed above. The experimental setups considered are gels that are not perfectly isolated along the sides, and the electric potential outside this incomplete isolation is the same as at the anode. This situation arises in gels, where there may be imperfections in the gluing of the spacers and the gels are run fully submerged in running buffer.

2.3.1 Current leakage model

2.3.1.1 General assumptions

The electrostatic potential in the gel, Ψ (V), satisfies Poisson's equation

$$-\nabla \cdot (\varepsilon \nabla \Psi) = \rho \tag{2}$$

where ∇ (1/m) is the gradient operator, ρ (C/m³) is the net charge volume density and ε (As/Vm) is the permittivity of the medium. In general, the species continuity equations for all charged species in the solute and solvent have to be considered and coupled to Eq. (2), but when the concentration of the solute is much lower than that of the current carrying ions in the solvent, electro-neutrality, that is $\rho = 0$, can be assumed and a decoupling of the electric potential from the migration of the solute and solvent is possible [22]. The media and solvent is looked upon as a single homogeneous unit, referred to as the gel. The migration velocity, v (m/s), of the SDS-protein complexes in the gel is assumed to be a linear function of the electric field, the negative gradient of the electric potential:

$$v = -\nu \varepsilon \nabla \Psi \tag{3}$$

where ν (m²/Vs) is the mobility coefficient of the SDS-protein complex. We assume that the mobility coefficient is spatially uniform and independent of the electric field strength. The spatial dimensions are denoted in terms of a gel run on a vertical system. The geometry can be seen in Fig. 1, where the cathode and anode are the top and bottom boundaries, respectively. The first-dimensional strip with the isoelectric focused proteins is put on top of the gel, and the proteins migrate downwards in the second-dimensional gel electrophoresis. The upper left corner of the gel is defined as the origin, with the horizontal dimension denoted x_1 and the vertical dimension x_2. Any dependency of the third dimension is disregarded, and all spatial quantities are considered to be two-dimensional.

2.3.1.2 Electrostatic potential in gel

With assumptions of electro-neutrality and a uniform permittivity, the electrostatic potential in the gel is given by Laplace's equation

$$\nabla^2 \Psi = 0 \tag{4}$$

with boundary conditions

$$\Psi = 0 \text{ at the anode (bottom)}, \tag{5}$$

$$\frac{\partial \Psi}{\partial n} + \gamma_l \Psi = 0 \text{ at the left boundary}, \tag{6}$$

$$\Psi = -V_0 \text{ at the cathode (top)}, \tag{7}$$

$$\frac{\partial \Psi}{\partial n} + \gamma_r \Psi = 0 \text{ at the right boundary}. \tag{8}$$

Here, V_0 (V) denotes the absolute value of the applied voltage and $\partial\Psi/\partial n$ is the outward normal derivative of Ψ at the boundary.

At the anode and cathode boundaries Ψ is constrained by Dirichlet conditions. The generalized Neumann condition at the two vertical boundaries is derived from Ohm's
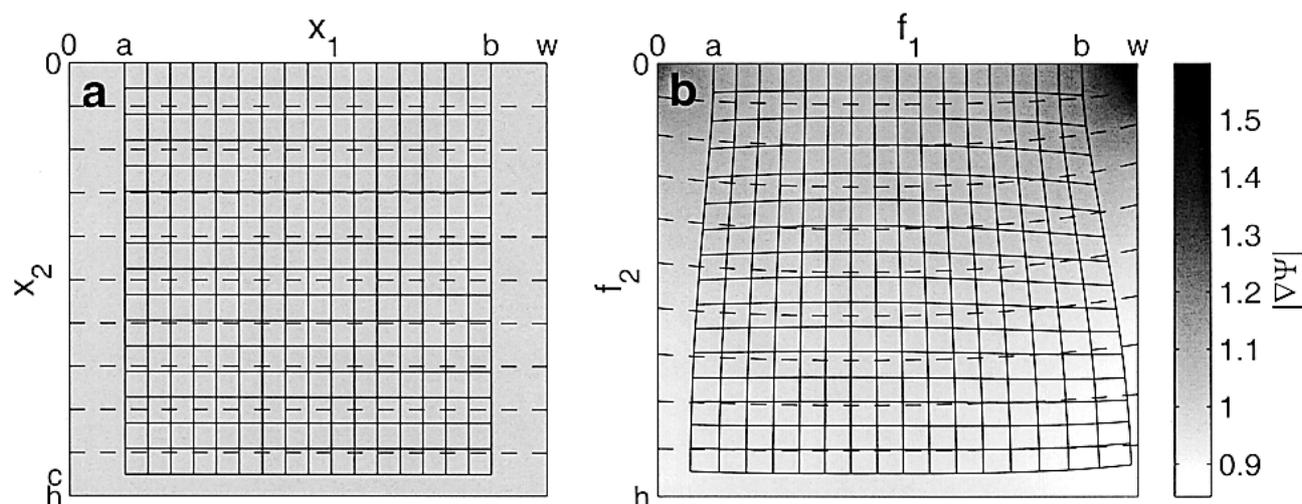
Figure 1. Warping of the protein pattern due to current leakage. (a) An ideal gel and (b) a gel with current leakage across the sides. The electric potential is visualized by equipotential lines drawn as dashed curves and the electric field strength is indicated with a grey scale intensity according to the scale shown in the bar to the right, where it is given relative to the undisturbed field strength in the ideal gel. The grid of deformation, shown with solid lines, illustrates the current leakage warping function: the map between the position *x* a protein would have had in an ideal gel to the position *f* in the gel with current leakage. The dimensionless conductivity parameters at the left and right boundaries are in this example $g_l = 0.2$ and $g_r = 0.4$, respectively.

law and the law of charge conservation. In brief, the first term in the condition originates from the component of the current field in the gel that is orthogonal to the boundary, and the second term originates from an expression for the leakage current in the imperfect gluing of the spacer, where it is assumed that the spacer width Δw is small in comparison with the spacer length. The quantity that determines the amount of current leakage is

$$\gamma = \frac{\sigma_{\text{side}}}{\sigma_{\text{gel}}\Delta w} \qquad (9)$$

where σ_{gel} and σ_{side} (A/V) are the conductivities in the gel and in the gluing of the spacer, respectively. The larger γ is, the larger is the current leakage. Note that γ_l and γ_r are functions of the second dimension coordinate along the boundaries.

2.3.1.3 Warping of the protein pattern

The geometric distortions of the protein pattern due to current leakage is a function that maps the position *x* that a protein would have had in an ideal gel with perfect isolation along the sides to the corresponding position *f(x)* in a gel with leakage. This current leakage warping function *f(x)* is given by a set of ordinary differential equations in x_2, one for each x_1. Each differential equation determines the map of positions along a migration path in the ideal gel to positions along the corresponding path in the gel with current leakage:

$$\frac{\partial}{\partial x_2}f(x) = -\frac{h}{V_0}\nabla\Psi(f(x)), \quad \text{for } x_2 \in (0, c) \qquad (10)$$

with initial value

$$f((x_1, 0)) = (x_1, 0), \text{ for } x_1 \in (a, b). \qquad (11)$$

Here, *a* (m) and *b* (m) are the left-most and right-most x_1-coordinate of the first-dimensional strip, and *c* (m) denotes the x_2-position of the gel front in the ideal gel; the gel front consists of proteins and protein fragments that are too small to be resolved in the gel. Hence, these numbers delimit the area that contains proteins in the ideal gel: $(a, b) \times (0, c)$.

The differential equation is derived by considering the velocity equation (3) for a protein in both gels, noticing that time can be expressed in terms of settled distance in the second dimension in the ideal gel and using this substitution in the differential equation for the migration in the gel with leakage. Note that the mobility coefficient has canceled out in Eq. (10).

2.3.1.4 Governing dimensionless quantities

By putting the model equations on dimensionless form it is possible to identify dimensionless combinations of the physical quantities that alone govern the result. A dimensionless formulation also provides proper scaling of the quantities for numerical treatment. A natural choice is to

express all length quantities in terms of the gel height h and the electric potential in terms of the absolute value of the applied voltage, V_0.

There are two governing dimensionless combinations of the physical quantities: the ratio of the gel width and the gel height, w/h, and a dimensionless version of the conductivity expression in Eq. (9), $h\gamma$. Hence, the scaled current leakage warping function is fixed by specifying the geometry ratio and the dimensionless conductivities along the two vertical sides with two arbitrary functions.

2.3.2 Parameter estimation

To be able to use the current leakage model one has to consider two tasks. Firstly, one has to know the map between the gel image domain and the model gel domain. Secondly, one has to estimate the governing dimensionless quantities in the model to calculate a current leakage warping function. In order to solve these tasks, we extract two features in the gel image: a straight line is fitted to the top edge and a low-degree polynomial curve is fitted to the gel front. We have done this manually by specifying two points on the cathode gel edge and 6–8 points on the gel front, but the procedure could easily have been automatized.

The map between the gel image domain and the model gel domain is a transformation consisting of a dilation, a rotation and a translation. The scanning resolution and physical gel size determine the scale factor in the transformation. The slope of the top edge line gives the angle in the rotation. The vertical shift in the remaining translation is determined by the vertical position of the top edge line. However, we do not know the sideways position of the physical gel in the gel image: for example, the position of the upper left corner of the physical gel along the extension of the top edge line in the image. Hence, we cannot determine the horizontal shift in the translation step in the transformation from the gel image domain to the model gel domain. We therefore introduce a horizontal translation parameter ξ defined as the unknown position of the upper left corner of the gel along the top edge line.

We parameterize the dimensionless conductivity functions with constant values g_l and g_r, *i.e.* $h\gamma_l \equiv g_l$ and $h\gamma_r \equiv g_r$, one constant for each side. The partial differential equation for the dimensionless electrostatic potential is numerically solved in the PDE Toolbox in MATLAB® (MathWorks Inc., [23]). A grid of points is used to specify the current leakage warping function: the area of interest $(a, b) \times (0, c)$ is partitioned into a 16×16 grid, where the grid size is chosen *ad hoc*. The ordinary differential equation for the warping of the 17 migration paths, one for

each initial position in the partitioning of the interval (a, b), is solved with a simple forward iteration after a partitioning of the x_2-dimension, and the 17 uniformly spaced points in the interval $(0, c)$, 17 for each initial value, then specify the grid.

The parameters are estimated by fitting the model gel front to the image gel front. The model gel front curve is specified by three parameters: the two conductivity parameters g_l and g_r and the second-dimensional position of the gel front in the ideal gel, c. The image gel front curve is mapped into the model coordinate system, where the mapping needs the specification of the horizontal translation parameter, ξ. It is these four parameters that are estimated by a least square fitting of the two curves. This was implemented with an initial preprocessing step where the model is simulated for a discrete number of conductivity parameters, and a low-degree polynomial curve is fitted to a discrete number of gel front positions c. This library of curves for different g_l, g_r and c parameters is then searched together with the parameter space for ξ to get an optimal fit of the two curves.

2.3.3 Warping gel image

The model is first simulated with the estimated conductivities $\hat{g}_l$ and $\hat{g}_r$. In order to include a maximum defined area in the warping of the gel image, the delimiters of the first dimension, $\hat{a}$ and $\hat{b}$, are chosen by backward iteration from the lower left and lower right corners of the gel, respectively. The grid of deformation of the scaled current leakage warping function is calculated for the area of interest $(\hat{a}, \hat{b}) \times (0, \hat{c})$ and then mapped into the gel image with the estimated translation parameter $\hat{\xi}$. The warped version of the gel image is finally obtained with Eq. (1) where bicubic spline interpolation in the algorithm interp2 in MATLAB® is used to interpolate the warping function between the grid points.

2.4 Warping step II: image alignment

We automatically align the current leakage corrected images one-by-one to a reference image. For reference image we choose one of the five current leakage corrected gel images from normal growth conditions: the one with smallest maximum of the two estimated conductivity parameter. In order to use only the spot pattern in the alignment, the reference image is cropped such that all spots in the earlier analysis (Alipour *et al.,* unpublished) are present in the smaller version, while the irreproducible top part and gel front are excluded. We denote the cropped reference image with μ and let Y be the image that we wish to align with μ.

First a top-hat transformation [24] is applied to all gel images to subtract large-scale trends in the background level. The diameter of the cylinder in the top-hat transformation is set to about one tenth of the width of the spot containing area, far larger than any spot size but small enough to capture large-scale trends. We use a diameter of 100 pixels in our images. The data in 2-D gel images range over several orders of magnitude, and since we will use an intensity based similarity criterion, see Eq. (15) below, it is convenient to have the data on the same order of magnitude. The intensity in all images is therefore transformed according to

$$\tilde{Y}_x = \log(Y_x + 1), \text{ for } x \in X \tag{12}$$

A tilde mark is used to denote the log-transformed versions of the images. These versions are only used in the alignment procedure and not in the final warping of the images.

2.4.1 Penalized likelihood

We estimate the warping function that aligns $\tilde{Y}$ with $\tilde{\mu}$ by maximizing a penalized likelihood for the warping [11]. The penalized likelihood is a functional, P, composed of two components:

$$P(\tilde{Y}|\tilde{\mu}, f, \lambda, C) = L(\tilde{Y}|\tilde{\mu}, f) - \lambda D(f, C). \tag{13}$$

The first component L is the log-likelihood for $\tilde{Y}$, conditional on $\tilde{\mu}$ and the warping function f, and measures the similarity between the warped image $\tilde{Y}_f$ and $\tilde{\mu}$. In the second component D is a distortion criterion, chosen to be zero if f belongs to C, a null-set of functions, and λ is a non-negative constant that gives the relative weighting between L and D.

We use a simple Gaussian image model with independent pixel values for $\tilde{Y}$ after warping:

$$\tilde{Y}_{f(x)} \sim N(\tilde{\mu}_x, \sigma^2), \text{ for } x \in X \tag{14}$$

Disregarding additive and scaling constants, the log-likelihood for $\tilde{Y}$ is then the negative sum of squared differences:

$$L(\tilde{Y}|\tilde{\mu}, f) = -\frac{1}{n_1 n_2} \sum_{x \in X} (\tilde{Y}_{f(x)} - \tilde{\mu}_x)^2 \tag{15}$$

where we have normalized with the image size because we will shift between different image resolutions in the optimization algorithm.

As null-set in the distortion criterion we choose the set of affine functions, A, and as the distortion criterion the bending energy of thin plate splines in a finite window [26]:

$$D(f, A) = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} \int_{\square} \left(\frac{\partial f_i}{\partial x_j \partial x_k} \right)^2 dx \tag{16}$$

integrated over the image domain $\square = (0, n_1) \times (0, n_2)$.

2.4.2 Optimization algorithm

To assess the alignment we followed the method in [11], where the warping function is specified on a $q_1 \times q_2$ grid, a piecewise bilinear function interpolates elsewhere and a conjugate gradient algorithm is used to find the maximum penalized likelihood. Since the distortion criterion in Eq. (16) involves second order derivatives, and a piecewise bilinear function is a first-degree polynomial with discontinuous first order derivatives, $D(f, A)$ is approximated in the method. The first partial derivatives of P required by the conjugate gradient algorithm are computed using difference methods, taking advantage of the fact that the piecewise bilinear function has local support; changes in a point in the grid of deformation only affect a subset of terms in $\tilde{Y}_f$ and D.

All implementations were done in the MATLAB® software. The algorithm fminunc in the Optimization Toolbox in MATLAB® was used to maximize the penalized likelihood with respect to the $2(q_1 + 1)(q_2 + 1)$ parameters in the grid of deformation. We adopted a multiresolution strategy to avoid being trapped in local sub-optima, see Table 1. The strategy operates on different image resolutions, a coarse resolution in the beginning and finer scales as iterations advance, simultaneously refining the grid size $q = (q_1, q_2)$ and allowing more distortions by decreasing λ. The different image resolutions were obtained with a sequence of image reductions by 2×2 block averaging [24], and a subsequent smoothing with a Gaussian filter. The original size of the reference image μ was 816×1168 pixels. The final grid size q and the permitted distortion *via* the parameter λ_0 were chosen to maximize the number of adjacent spot pairs in the aligned images, see Section 3.2.3.

Table 1. The multiresolution strategy in the optimization algorithm to assess the image alignment in warping step II

Iteration	Size of reference image μ [pixels]	Grid size q	Penalty weight λ
1	51×73	$(1, 1)$	$30 \lambda_0$
2	102×146	$(2, 3)$	$10 \lambda_0$
3	102×146	$(4, 6)$	$3 \lambda_0$
4	204×292	$(8, 12)$	λ_0
5	204×292	$(16, 24)$	$0.3 \lambda_0$

The original size of the reference image μ is 816×1168 pixels.

2.5 Evaluation

2.5.1 Spot displacement vectors

We have used the list of 1028 manually matched spots from the earlier analysis of the gels (Alipour *et al.,* unpublished) to measure the geometric similarity between the spot patterns in two images. Let $\left(x_1^{(j)}, x_2^{(j)}\right)$ and $\left({}^{\text{ref}}x_1^{(j)}, {}^{\text{ref}}x_2^{(j)}\right)$, where $j = 1, \ldots, 1028$ is the protein identification number, denote the coordinates of the spot centers in the image Y and the reference image μ, respectively. The images can be either the original images, the current leakage corrected versions or the aligned versions. When superimposing the original and current leakage corrected versions, the image Y is first translated such that the most dominant spot (protein Tdh3 [26]) match. The displacement vector between the two spot centers in the spot pair that correspond to protein j is given by

$$\Delta x^{(j)} = x^{(j)} - {}^{\text{ref}}x^{(j)} \tag{17}$$

We define the spot pair to be an adjacent spot pair if the spot displacement is less than or equal to a mean spot radius,

$$\text{adjacent spot pair} \Leftrightarrow \left|\Delta x^{(j)}\right| \leq \bar{\sigma} \tag{18}$$

We let the mean spot radius $\bar{\sigma}$ be the mean value of the spot half-widths [28] of all spots in all ten gel images and in both directions from the earlier analysis. In our gels this radius is equal to 1.1 mm or equivalently 6.3 pixels.

2.5.2 Gel image analysis in PDQuest™

We have adopted the following strategy to create a list of scattered landmarks for the matching step in the comparison within PDQuest™ (Version 6.1.1) (Bio-Rad Laboratories, Hercules, CA, USA [28]). The reference image μ is first divided into 2×3 main rectangles. Each main rectangle is then split into 3×3 sub-rectangles. Hence there are totally 54 sub-rectangles. The most intense spot in each sub-rectangle is picked as a landmark candidate and the nine candidates in each main rectangle are ordered in descending magnitude. The full list of 54 landmarks is then obtained by starting with the most intense spot in the gel and then walking clockwise around the main rectangles, successively picking landmarks from the ordered candidate list.

The preprocessing step in PDQuest™ was done with a background subtraction with a floating ball of radius 25 pixels and a smoothing with a 5×5 kernel (the so-called contra mean). Of the detection parameters the size scale was set to 5 and the sensitivity parameter was adjusted in each gel such that approximately 1250 spots

were detected in a preliminary test run. Spots in the gel fronts were removed manually after the spot detection step. Landmarks were then added according to the list above and an auto-match of all gels was done after each added landmark. After all 54 landmarks had been marked, further landmarks were manually added until all spots detected in all ten gels were matched.

3 Results

3.1 Warping step I: current leakage correction

3.1.1 Numerical simulation of the current leakage model

The result of a simulation of the current leakage model is shown in Fig. 1. The dimensionless conductivity parameters are in this example set to $g_l = 0.2$ and $g_r = 0.4$. Hence, in this simulation we consider a case where the conductivity in the imperfect gluing of the spacer on the right hand side is twice that on the left hand side. The dashed lines show equipotential lines and the grey scale intensity corresponds to the electric field strength in terms of V_0/h (V/m) according to the gray scale in the bar to the right. The grids of deformation shown with solid lines illustrate the current leakage warping function: the map between the position $x = (x_1, x_2)$ a protein would have had in the ideal gel to the position $f(x) = (f_1(x), f_2(x))$ it has in the gel with current leakage. When interpreting the current leakage warping function the positions in the vertical dimension can either be seen as positions of proteins with different molecular weight at a certain time: a snapshot of the gel, or as positions of proteins with the same molecular weight but at different times: a superimposition of several snapshots.

Studying the vertical lines in Fig. 1b, one can see that the SDS-protein complexes migrate in the electric field direction which slightly pulls them towards the sides of the gel. In the upper parts of the gel, the potential over the glued spacer is the largest, therefore the current leakage is the largest and the electric field deviates the most towards the sides. However, at the very top, the electric field points straight down due to the constant potential along the cathode boundary. Further down the gel, the potential is smaller and thus the leakage is also smaller, and the field reverts towards the vertical direction.

Looking at the horizontal lines in Fig. 1b, which indicate proteins with the same molecular weight, one can see that the inhomogeneous electric field strength in the upper part of the gel causes proteins with the same molecular weight to move faster nearer the sides than in the middle part. This lead is lost further down the gel when the electric field becomes weaker near the sides in com-
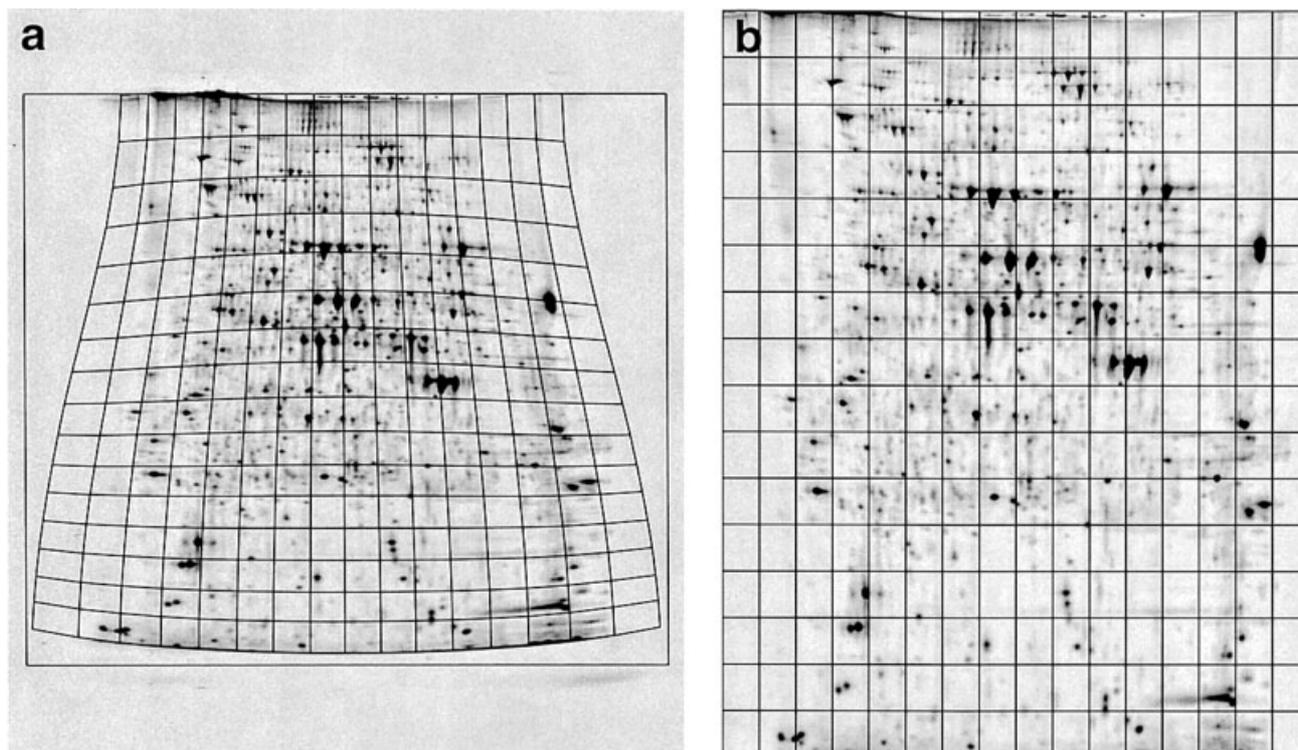
Figure 2. Warping step I: Current leakage correction of a gel image. (a) Original image (gel 1). (b) Warped and current leakage corrected version of the gel image. The grids of deformations according to the model are superimposed in the images. The estimated dimensionless conductivity parameters are $\hat{g}_l = 0.65$ and $\hat{g}_r = 0.7$.

parison with the middle part. Finally at the very bottom, the gel front appears as a "smile" in the gel with current leakage. The curve form of this "smile" was used to estimate the parameters in the model.

3.1.2 Current leakage corrected images

A result of the first warping step is shown in Fig. 2, where gel image 1 is warped to correct for the effect of current leakage. The dimensionless conductivity parameters for this gel were estimated to $\hat{g}_l = 0.65$ and $\hat{g}_r = 0.7$. The estimated parameters in the other gel images ranged between 0 and 1.175 with a mean value of 0.50. Considering the definition of the conductivity expression in Eq. (9), this corresponds in rough numbers to that the average conductivity in the gluing of the spacers was 0.06 times the one in the gel.

3.2 Warping step II: image alignment

3.2.1 Aligned images

The second warping step where the current leakage corrected images are aligned is illustrated in Fig. 3, where an alignment of the current leakage corrected gel 6 with the

reference image (cropped current leakage corrected gel 1) is shown. The distortion penalty weight λ_0 was in this example chosen to be 10^{-3} and iterations were stopped after iteration four when the size of the grid of deformation was $q = (8, 12)$, see Table 1. Note the log-transformed intensities in the gel images that enhances the low-intensity parts; compare Fig. 3c with Fig. 2b. When we tried to align the gel images without this initial intensity transformation, the algorithm primarily considered the most prominent spots due to the intensity-based similarity criterion and the fact that the data in 2-D gel images range over several orders of magnitude, and the alignment was not successful in areas with no dominant spots.

There are two sources of variation between the two aligned gel images in Fig. 3. Firstly, they show proteins from cells grown under different conditions: image Y (gel 6) from saline growth conditions and reference image μ (gel 1) from normal conditions. Secondly, the gels were run in two different electrophoresis tanks. Despite these differences the spot patterns and the reproducible background features are sufficiently similar for the alignment to succeed. However, a weakness in the intensity-based similarity criterion can be seen in the lower left corner of the images. The irreproducible vertical band to the left in Y is not present in μ.
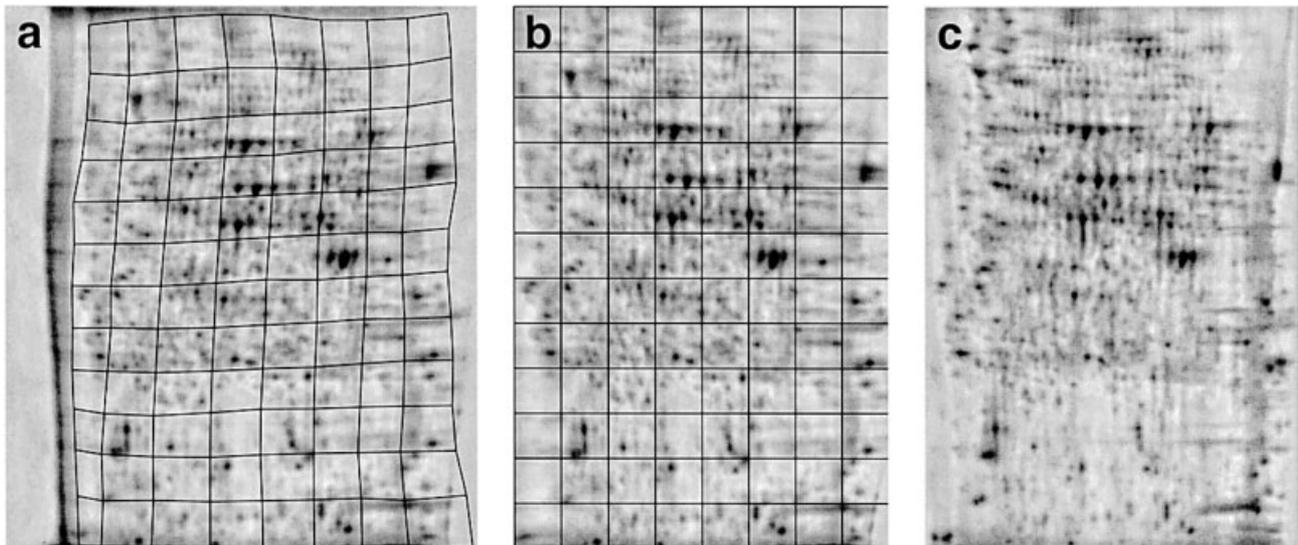
Figure 3. Warping step II: Image alignment of two current leakage corrected images. (a) Image $\tilde{Y}$ (current leakage corrected gel 6). (b) Warped version $\tilde{Y}_f$ to align with (c) reference image $\tilde{\mu}$ (cropped current leakage corrected gel 1). The grid of deformation for the alignment is superimposed in the images. Note the log-transformed intensities in the images.

Therefore, the method avoids this area and is unable to align the spots that are embedded in this band in image *Y*.

The image alignment can also be seen in the color-superimposition in Fig. 4a. As in the color-superimposition of 2-D gels by Horgan *et al.* [4], the reference image μ is colored red and the warped image Y_f is colored blue. When the spots in the images coincide the two complementary colors add up to grey or black. An unmatched spot pair shows up as a blue and a red spot near each other; the reason for this can either be a misaligned area, as in the lower left corner, or a protein specific change in position due to for example a phosphorylation. A single red or blue spot indicates that the protein is downregulated or upregulated during steady state saline growth, respectively. There are quite a number of examples of the latter in Fig. 4a.

3.2.2 Spot pattern similarity

We have used a list of 1028 manually matched spots from an earlier analysis of the gel images (Alipour *et al.*, unpublished) to measure the similarity of the spot patterns in the images. Figure 4b shows the spot displacement vectors in the color-superimposition of the aligned images in Fig. 4a. The larger the vector is, the larger is the arrow head. Figure 4c shows another representation of these displacement vectors. The vectors are translated with their starting points to the origin and their endpoints, *i.e.* the arrow tips, are marked with dots. The

circle shows the criterion for an adjacent spot pair. The circle diameter is the mean value of the spot widths of all spots in all images. The number of dots inside the circle is the number of adjacent spot pairs. Among the largest vectors we find the misaligned spots in the lower left corner mentioned above, and horizontal displacements of alkaline proteins in the rightmost part of the gels; these proteins are sometimes hard to get properly focused in the first dimension. There are also a few number of cases where one might suspect a manual error in the earlier analysis: for example the almost vertical displacement vector in the uppermost part of Fig. 4b.

3.2.3 Choice of alignment algorithm parameters

Figure 5a shows the successive improvement of the spot pattern similarity, measured as percentage of adjacent spot pairs, during the iterations in the alignment algorithm; the figure shows the case where the distortion penalty weight λ_0 was 10^{-3}. The percentage of adjacent spot pairs is plotted as a function of the number of parameters in the successively refined grid of deformation, see Table 1. The number of parameters in the grid is a natural choice to characterize each iteration step since it is one of the crucial properties of a multidimensional optimization problem. The circles in Fig. 5 mark the mean values and the error bars the standard deviations in the set of the nine images aligned to the reference image. The percentage of adjacent spot pairs was increased with refined grid size until a saturated value was reached
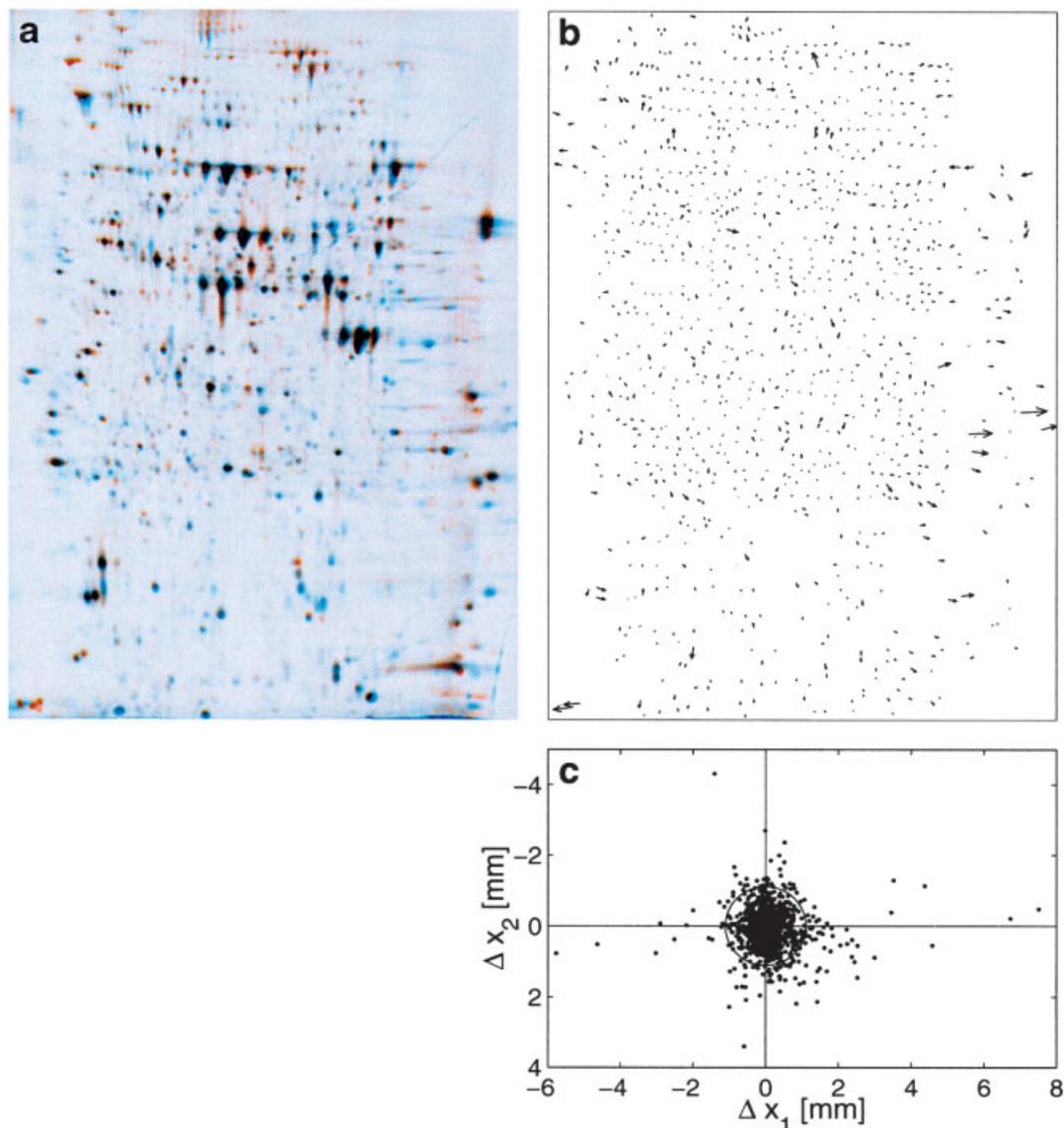
Figure 4. Similarity of the spot patterns in the aligned images in Fig. 3 (a) A color-superimposition of the two images: the reference image μ is colored red and the warped image Y_f is colored blue. (b) Spot displacement vectors in the image superimposition; the larger the vector the larger the arrow head. (c) The displacement vectors are translated with their starting points to the origin and their end-points are marked with dots. The circle shows the criterion for an adjacent spot pair.

after the fourth iteration step where a 8 × 12 grid was used; further refinement of the grid in the fifth iteration did not improve the spot pattern similarity.

Figure 5b shows the effect of using different distortion penalty weights in the alignment algorithm. The attained percentage of adjacent spot pairs after iteration four is plotted as a function of λ_0. Hard distortion penalty *via* large λ_0 corresponds to a warping function that is close to the null-set of the distortion criterion: the set of affine functions. An affine function is a combination of a dilation,

a translation, a rotation and a shearing of the image. It is unable to align the images and the percentage of adjacent spot pairs was thus low for large penalty weights. The percentage was increased as more distortions were allowed until a saturated value was reached for $\lambda_0 = 10^{-3}$; further decrease of the distortion penalty weight did not improve the spot pattern similarity.

The final grid size, *i.e.* the number of iterations, and the distortion penalty weight for the final warping of the images were chosen to be the ones that achieved the
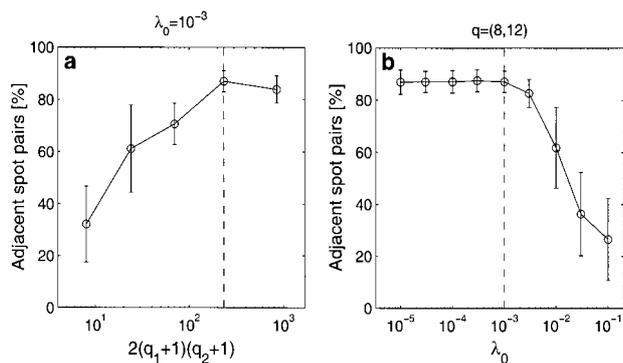
Figure 5. Spot pattern similarity in aligned images after each iteration step in the optimization algorithm and for different penalty weights. (a) The percentage of adjacent spot pairs obtained after each iteration step. Each iteration is illustrated with the number of parameters in the grid of deformation. (b) The effect of using different values of the distortion penalty weight λ_0 on the percentage of adjacent spots pairs. The circles mark the mean values and the error bars the standard deviations in the set of the nine images aligned to the reference image. The vertical dashed lines show the chosen final grid size and penalty weight.

saturated number of adjacent spot pairs, and at the same time was computationally least time-consuming and introduced the least distortion. The choices, $\lambda_0 = 10^{-3}$ and $q = (8, 12)$, are shown with vertical dashed lines in Fig. 5. The finer the grid of deformation was, *i.e.* the larger the parameter space for the optimization algorithm to search, the longer was the computation time. The lower the distortion penalty, the larger was the chance that the algorithm produced spurious transformations in parts of the image where there were little information for the similarity criterion. This sometime happened near the image edges for low distortion penalties (data not shown).

3.3 Comparing the two warping steps

The successive improvement of the spot pattern similarity after each one of the two warping steps is shown in Fig. 6. The length distribution of spot displacement vectors is plotted for the three sets of images: the original, the current leakage corrected (warping step I) and the aligned images (warping steps I+II). The current leakage correction step slightly improved the spot pattern similarity. The step had mainly corrected for large displacements in the lower part of the gel image that were caused by the wide "skirt" in the geometric distortion caused by current leakage. However, it was the second alignment step that clearly made the largest improvement.
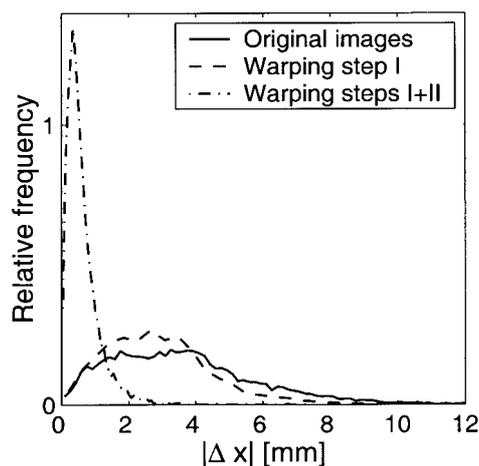

Figure 6. Comparison of the spot pattern similarity before, after one and after the two warping steps. The figure shows the length distributions of the spot displacement vectors in three sets of images: original images (solid line), current leakage corrected images (warping step I) (dashed line) and aligned images (warping steps I+II) (dash-dot line).

The result when only applying the alignment step (warping step II) on the original images in comparison to the two-step warping (warping steps I+II) is presented in Table 2. We used the same reference image, *i.e.* the cropped current leakage corrected gel 1, since it did not contain the irreproducible top part and gel front. Both the image similarity measure L and the spot pattern similarity measure are very much the same in the two cases. Actually, the length distribution of spot displacement vectors almost coincided with the one of the two-step warped images in Fig. 6 (data not shown). The value of the distortion criterion D was slightly larger, *i.e.* the warping functions for alignment of the original images were further different from an affine function than the warping functions for alignment of the current leakage corrected images.

Table 2. Result of only applying the alignment step (warping step II) on the original images in comparsion with the two-step warping (warping steps I+II)

	Warping steps I+II		Only warping step II	
Similarity criterion L	−0.158	(0.085)	−0.151	(0.066)
Distortion criterion D	3.88	(1.61)	4.80	(1.20)
Computation time (min)	74	(16)	129	(28)
Adjacent spot pairs (%)	87.1	(4.1)	86.8	(4.4)

The numbers in the table are the means in the set of nine images and with standard deviations given in parentheses.

The warping due to current leakage that the alignment step now also had to deal with is far different from an affine function. However, the most significant difference was in computation time: applying the alignment step on the original images instead of the current leakage corrected versions took almost twice as much computation time per gel (CPU time on a SUN Ultra 10). The first current leakage correction step merely took a minute or two per gel.

3.4 Comparison within PDQuest™

The improvement of the matching efficiency when applying the warped images to an existing matching algorithm is exemplified with the PDQuest™ software [27]. The global matching efficiency when processing the sets of original (dashed line) and two-step warped images (solid line) is shown in Fig. 7a. The number of automatically matched spots in all ten gels is plotted as a function of the number of manually added landmarks, see Section 2.5.2. The number of spots is expressed as a percentage of the number of spots that was detected in all ten gels. The matching efficiency was clearly improved when using the warped images instead of the original versions: the number of matches was larger for each number of landmarks and the total number of possible matches was reached earlier.

However, it was a limited number of spots that was detected in all ten images. Figure 7b illustrates the result of the initial spot detection in the set of warped images.
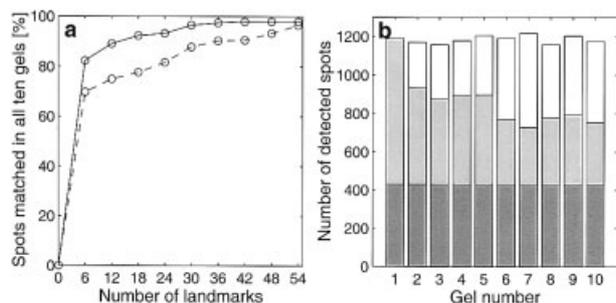

Figure 7. Gel image analysis in PDQuest™. (a) Global matching efficiencies when processing the two-step warped images (solid line) and the original images (dashed line). The number of automatically matched spots is expressed as percentage of spots detected in all ten gel images, and shown as a function of manually added landmarks. (b) Analysis of the spot detection. The bars show the number of detected spots in each gel image in the set of warped images: the spots detected in all ten images (dark gray), spots additionally in common with the reference image (gel 1) (light gray) and spots not detected in the reference image (white).

Although the mean number of detected spots in each gel was 1194, the number of spots that each gel had in common with the reference gel (dark + light grey) was reduced to on average 826 spots. This resulted in that there were merely 430 spots detected in all ten gels (dark grey) that could be matched and then be used in a subsequent comprehensive statistical analysis. The situation was similar in the set of original images (data not shown). Also note the slightly higher number of spots in common with the reference image (gel 1) within the set of gels from normal growth conditions (gels 1–5) compared with the gels from steady state saline growth (gels 6–10).

4 Discussion

4.1 Current leakage correction

The theoretical studies that model the migration behavior of SDS-protein complexes [16–18, 20, 21] aim at deriving a relation between the size of a protein and its mobility coefficient. They all consider the migration of the SDS-protein complexes in the gel fiber net at a microscopic scale. Most studies assume a constant electric field, whereas Locke [20] and Mercier *et al.* [21] investigate curved electric field lines at this size scale. We have worked on a macroscopic scale to investigate the effect that an inhomogeneous global electric field caused by current leakage has on the protein spot pattern. Since we are interested in relative position changes due to the altered global electric field, the mobility coefficient cancels out in the differential equation for the warping function in Eq. (10).

The warping function produced by the current leakage model seems similar to the warping function in the method to correct for "frowns" and "smiles" in multitrack electrophoresis gels proposed by Glasbey and Wright [8]. They also treat the images individually, but they can use the fact that bands in these gels are asymmetric to estimate orientation directions and to mathematically derive a warping function of the whole image. We have merely used information in the top and bottom part of the gel images and have had to rely on the physicochemical model in between. When evaluating the result of the current leakage correction step in Fig. 6, one should have the following in mind: the images are treated individually with no knowledge of the spot patterns in the other gels and the choice of a constant functions to parameterize the unknown conductivity function along the side is clearly too simple; the imperfections in the gluing of the spacers are likely to be more heterogeneous.

4.2 Automatic image alignment

We agree with Smilansky [10] that spot pattern features, such as spot shapes, streaks and smears, and background structures are reproducible and can be used by pixel-based methods to align the protein spot pattern in gel images. The image warping technique for automatic alignment of images has been used extensively in the field of digital image analysis [3]. With the exception of an early example by Conradsen and Pedersen [9], the technique seems not to have been applied to 2-D gels until very recently [10]. We find that the penalized likelihood formulation proposed by Glasbey and Mardia [11] is a powerful tool for the automatic alignment of gel images.

Our similarity criterion – the log-likelihood based on a Gaussian image model for the images after warping – is the well-known negative sums of squared differences in machine vision literature. This intensity matching measure was also used by Conradsen and Pedersen [9]. However, we applied an initial log-transformation of the intensities in the images. This essentially corresponds to that we have used a log-normal image model for the images after warping without an initial log-transformation. The similarity measure used by Smilansky [10] is described in general terms and consists of several components: two intensity matching measures, and an edge matching measure based on the scalar product of the gradient vectors in the two images. In a combination of different type of measures, like in Smilanskys' method, it is necessary to specify the weights required to combine the components. An alternative is to use the Fourier-von Mises image model proposed by Glasbey and Mardia [11] which combines intensity and edge matching in one measure.

The largest conceptual difference of our method compared with the earlier warping methods [9, 10] is that we keep explicit control of the distortions introduced in the warping steps: in the first warping step the distortion is a smooth function derived from a physicochemical model, and in the second alignment step rough distortions are penalized by a criterion consisting of the bending energy of thin-plate splines [26] which is minimized by the set of affine functions. Conradsen and Pedersen [9] do not consider any distortion penalty at all and the assessed warping function that aligns their gel images seems rough. Smilansky [10] keeps implicit control of the introduced distortions by restrictions in the iteration loop. Given an undistorted reference gel, the value of our distortion criterion can serve as a measure of the geometric distortion of the spot pattern in the gel image that is aligned with the reference.

There is also a difference in computation time between the method proposed by Smilansky [10], which is based on a large number of alignments of small subimages, and the methods proposed by Conradsen and Pedersen [9] and us, which both apply an alignment of the whole image. Smilansky's method is much faster in computation time, seconds or minutes per gel, while our implementation took about an hour per gel. However, our method seems robust; there was no need for any manual inteference to guide the alignment as sometimes is the case in Smilansky's faster method.

From the comparison when the alignment step was applied on the original images instead of the two-step procedure, see Table 2, we can conclude that the final result is essentially the same but that the current leakage correction step clearly gives the automatic alignment step more similar images and consequently a shorter computation time. When analyzing gels where there have been no current leakage during the second-dimensional electrophoresis and therefore less geometric distortions, one can probably skip the first current leakage correction step. Given an undistorted reference gel image, a next step might be to combine the two warping steps into one single alignment step by formulating a distortion criterion that has as null-set the functions that satisfy the differential equations in the physicochemical model.

4.3 Image warping methods in 2-D gel image analysis

The geometric distortions of the spot pattern seems smooth, which makes them correctable, *e.g.*, by using piecewise low-degree polynomials on relatively sparse grids. This deduction is supported by the result when a piecewise bilinear function was used in the image alignment and the effect of refining the grid and allowing more distortions was examined: a saturated number of adjacent spot pairs was reached, see Fig. 5, and most of the remaining pattern dissimilarities seem to be spot-specific displacements relative to the overall spot pattern, see Fig. 4b. We therefore conclude that one can divide the spot matching problem into two separate parts: a primary global correction for smooth nonspot-specific distortions where one regards the whole image, and a second more traditional spot-specific matching step where one can use a feature-based or pixel-based matching algorithm.

Voss and Habel [2] pointed out two major problems that push matching algorithms to their limit and have consequences for a comprehensive statistical analysis: geometric distortions of the spot patterns and irreproducibilities in spot intensities. In their study of a set of 49 gels, the global matching efficiency seems hampered primarily because of these two problems and only a limited number of spots is matched in all gels. Although we think that we have corrected for smooth geometric dis-

tortions, we see a similar limited number of matched spots in all gels in Fig. 7b. The reason for this is problems with the crucial spot detection step; there is a large number of spots that are near the detection threshold and because the images are treated individually, spots corresponding to the same protein are above this threshold in some images and below in other. Despite the fact that the warping technique proposed in this article is computationally time-consuming we think one should weigh, especially in the case of large-scale proteomics, this cost against the fact that the method is automatic and seems to be robust.

With the implementations in the present study it is not possible to quantify the spot contents in the warped images. The image warping as defined in Eq. (1) does change the protein spot content in the images (data not shown). Mathematically, a change in spot contents can be avoided by compensating for the area scale in the warping function and using a slightly different definition of the warped image compared with the one in Eq. (1): it reads

$$(Y_f)_x = Y_{f(x)} \left| \frac{\partial f}{\partial x} \right| \quad \text{for } x \in X \tag{19}$$

where the area scale $|\partial f / \partial x|$ is the Jacobian determinant of the warping function. We have not implemented this area scale compensation in this study since we in the second warping step have used a piecewise bilinear function that has discontinuous first order derivatives. A solution might be to use bicubic B-splines instead: a piecewise third-degree polynomial that has continuous second order derivatives and a local support [29]. In the present study we have started by considering the use of one of the simplest numerical approximations of arbitrary functions: the piecewise bilinear function, and have not taken a step to the B-spline in the hierarchy of function complexity. The presented work is primarily a method study, and our aim has not been to produce a general and user-friendly form that can be available to others.

5 References

[1] Pleissner, K.–R, Oswald, H., Wegner, S., in: Pennington, S. R., Dünn, M. J., (Eds.), *Proteomics*, BIOS Scientific Publishers, Oxford 2001, pp. 131–149.

[2] Voss, T., Haberl, P., *Electrophoresis* 2000, *21*, 3345–3350.

[3] Glasbey, C. A., Mardia, K. V., *J. Appl. Stat.* 1998, *25*, 155–171.

[4] Horgan, G., Creasey, A., Fenton, B., *Electrophoresis* 1992, *13*, 871–875.

[5] Appel, R. D., Vargas, J. R., Palagi, P. M., Walther, D., Hochstrasser, D. F., *Electrophoresis* 1997, *18*, 2735–2748.

[6] Lemkin, P. E., *Electrophoresis* 1997, *18*, 461–470.

[7] Pleissner, K.-P, Sander, S., Oswald, H., Regitz-Zagrosek, V., Fleck, E., *Electrophoresis* 1997, *18*, 480–483.

[8] Glasbey, C. A., Wright, F., *Electrophoresis* 1994, *15*, 143–148.

[9] Conradsen, K., Pedersen, J., *Biometrics* 1992, *48*, 1273–1287.

[10] Smilansky, Z., *Electrophoresis* 2001, *22*, 1616–1626.

[11] Glasbey, C. A., Mardia, K. V., *J. Roy. Stat. Soc. B* 2001, *63*, 465–514.

[12] Norbeck, J., Blomberg, A., *Yeast* 1997, *13*, 1519–1534.

[13] Blomberg, A., Blomberg, L., Norbeck, J., Fey, S., Larsen, P., Larsen, M., Roepstorff, P., Degand, H., Boutry, M., Posch, A., Görg, A., *Electrophoresis* 1995, *16*, 1935–1945.

[14] Westerhuis, W. H. J., Sturgis, M. N., Niederman, R. A., *Anal. Biochem.* 2000, 284, 143–152.

[15] Kozulic, B., *Anal. Biochem.* 1995, *231*, 1–12.

[16] Rodbard, D., Chrambach, A., *Proc. Natl. Acad. Sci. USA* 1970, *65*, 970–977.

[17] Guo, X. H, Chen, S. H., *Chem. Phys.* 1990, *149*, 129–139.

[18] Boileau, J., Slater, G., *Electrophoresis* 2001, *22*, 673–683.

[19] Bier, M., Palusinski, O. A., Mosher, R. A., Saville, D. A., *Science* 1983, *219*, 1281–1287.

[20] Locke, B. R., Trinh, S. H., *Electrophoresis* 1999, *20*, 3331–3334.

[21] Mercier, J.-F., Tessier, F., Slater, G. W., *Electrophoresis* 2001, *22*, 2631–2638.

[22] Locke, B. R., *Ind. Eng. Chem. Res.* 1998, *37*, 616–625.

[23] http://www.mathworks.com/.

[24] Glasbey, C. A., Horgan, G. W., *Image Analysis for the Biological Sciences*, Wiley, Chichester 1995.

[25] Bookstein, F. L., *IEEE Trans. Pattern Anal.* 1989, *11*, 567–585.

[26] http://yeast-2dpage.gmm.gu.se/.

[27] Garrels, J. I., *J. Biol. Chem.* 1989, *264*, 5269–5289.

[28] http://www.bio-rad.com/.

[29] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., Hawkes, D. J., *IEEE Trans. Med. Imaging* 1999, *18*, 712–721.