

# Protein Image Alignment via Piecewise Affine Transformations

FLORIAN A. POTRA,<sup>1</sup> XING LIU,<sup>1</sup> FRANÇOISE SEILLIER-MOISEIWITSCH,<sup>2</sup>  
ANINDYA ROY,<sup>1</sup> YAMING HANG,<sup>1</sup> MARK R. MARTEN,<sup>3</sup> BABU RAMAN,<sup>3</sup>  
and CAROL WHISNANT<sup>4</sup>

## ABSTRACT

We present a new approach for aligning families of 2D gels. Instead of choosing one of the gels as reference and performing a pairwise alignment, we construct an ideal gel that is representative of the entire family and obtain a set of piecewise affine transformations that optimally align each gel of the family to the ideal gel. The coefficients defining the transformations as well as the ideal landmarks are obtained as the solution of a large-scale quadratic programming problem that can be solved efficiently by interior-point methods.

**Key words:** image alignment registration, two-dimensional gels, affine transformations, interior-point methods, quadratic programming.

## 1. PROTEOMICS AND 2D-PAGE

**P**ROTEOME ANALYSIS INVOLVES THE SEPARATION, visualization and analysis of complex mixtures containing as many as several thousand proteins obtained from whole cells, tissues, or organisms. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), first introduced by O'Farrell and Klose in 1975, remains a core technology for separating complex protein mixtures in the majority of proteome projects. The main goal of protein separation methods is to detect differentially expressed proteins across treatment groups. However, a major bottleneck is the misalignment of gels due to warping, thus confounding biological variation with nonbiologically relevant distortions. This paper provides a computationally feasible gel alignment method based on powerful optimization techniques such as interior point methods.

With 2D-PAGE technique, proteins are separated orthogonally according to their charge and size. The separated proteins are then stained so that they are readily detectable, and the gels are digitally scanned into a database for storage. Gel images are analyzed with special software packages like Z3 (Compugen, [www.2dgels.com](http://www.2dgels.com)), MELANIE (Geneva Bioinformatics [GeneBio] S.A., [www.genebio.com](http://www.genebio.com)) and PD Quest (Bio-Rad Laboratories Inc., [www.biorad.com](http://www.biorad.com)).

Often we want to compare the effects of different treatments, or to see what biological changes happened to a species when exposed to a specific environment. Replicate gels are generated from tissue or blood

---

<sup>1</sup>Department of Mathematics and Statistics, University of Maryland, Baltimore, MD 21250.

<sup>2</sup>Department of Biostatistics, Bioinformatics, and Biomathematics, Georgetown University, Washington, DC 20057.

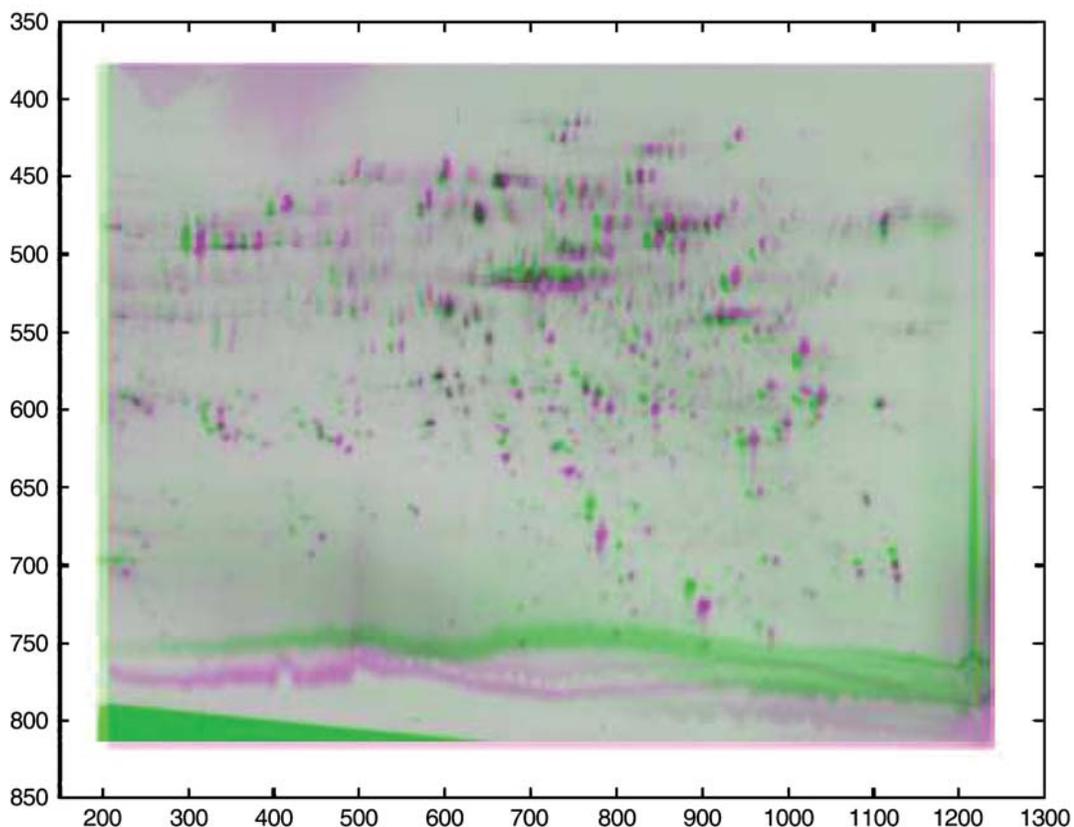
<sup>3</sup>Department of Chemical and Biochemical Engineering, University of Maryland, Baltimore, MD 21250.

<sup>4</sup>Research Triangle Institute, Research Triangle Park, NC 27709.

serum sampled of subjects belonging to different treatment groups and comparisons are made between the groups via a study of protein abundance on scanned images of the gels. A standard gel analysis system includes the following steps: image smoothing, spot detection and quantification, image alignment and spot matching, synthetic gel creation, and variation analysis. Fluctuation in location, size, and intensity of protein spots comes from three different sources: experimental variation, biological variability, and treatment effects.

Due to the complexity of the 2D-PAGE procedure, gel-to-gel experimental variation is substantial. Even under strictly controlled laboratory conditions, aliquots from the same sample can yield somewhat different protein maps. A spot that appears at one location on a given image may appear at a different location on another image. Gel images may be distorted from global shifts or local image warping. Evaluating a set of high-resolution 2-D gels by manual comparison is impossible. It is indeed very laborious to detect the emergence of a few new spots or the disappearance of a single spot among several hundreds of spots of different sizes and intensities when local distortions are present in each gel. Figure 1 is a superimposition of two gel images from two different treatment groups: overall location shift and nonuniform local distortions are obvious. To account for experimental variation, it is desirable that each subject have several profiles. Then we will be better able to detect the biological variation and the variation attributable to a treatment effect or to an exogenous stimulus like environmental changes. Thus gels have to be aligned first for spot matching, and a master gel representing a subject can then be created.

In the literature (see the survey by Dowsey *et al.* [2003]), one gel of the family is chosen as a reference gel and the other gels are aligned to it by means of appropriate warping functions. The choice of the reference gel has to be done by an operator, which introduces a bias in the process. In the present paper, we construct an ideal master gel that is representative of the entire family and obtain a set of geometric transformations that optimally align each gel of the family to the master gel. The coefficients defining



**FIG. 1.** Two gels from different treatment groups are superimposed. Overall location shift and local distortions can be clearly seen here.

the transformations as well as the ideal landmarks are obtained as the solution of a large-scale quadratic programming problem that can be solved efficiently by interior-point methods.

In Section 2, we discuss finer issues involved in gel alignment and spot matching steps and describe some of the existing methodologies. In Section 3, we introduce our new approach for aligning a family of gels to an ideal master gel, and we analyze the possibility of implementing this approach via a global affine transformation. The global transformation has limited applicability as the constraints are too stringent. To expand the scope of application, in Section 4, we describe a piecewise affine transformation method for gel alignment based on a nonuniform segmentation depending on the given data. In Section 5, we consider a piecewise affine transformation based on a uniform hierarchical grid that is independent of the given data. Section 6 evaluates the performance of the proposed method through data analysis. We summarize our results in Section 7.

## 2. ALIGNMENT AND SPOT MATCHING

The task of spot matching is to identify the spots corresponding to identical proteins across two or more different gels. Along with other difficulties, spot matching needs extensive human interaction, which is very time consuming. In order to match spots efficiently, one needs first to align the gel images to correct for the global and local distortions. Image alignment is performed by image warping. Warping is a transformation applied to the coordinate domain and aims at recovering the images from deformations.

The coordinates of the protein spots on the image are usually obtained on a pixel-by-pixel basis. In general, it is impossible to find a single transformation that can describe the change for all spots as there could be thousands of spots on a single image. Moreover, it is computationally extremely expensive. Efficiency is improved by preassigning control points or landmarks. Landmarks are a relatively small group of spots present in all gels being compared. They should be located across the entire gel image. The number of landmarks should be large enough so that they can carry enough information about the deformation. However, selecting too many landmarks will result in overfitting or an inefficient algorithm. In the remaining sections of the paper, we give different prescriptions to find transformations for the source gel images based on a given set of landmarks so that corresponding landmarks for the same protein on different images will gather together in a small local area after transformation. The size of these protein spots could vary a lot in different applications. In the remainder of the this paper, we consider a landmark as the set of  $(x, y)$ -coordinates for its center.

Many algorithms for warping have been implemented (Conradsen and Pedersen, 1992; Appel *et al.*, 1997; Gustafsson *et al.*, 2002) to align gel images. In all previous work, one of the gels is chosen as the reference gel (target gel) and the others (source gels or input gels) are transformed to provide a best match to the reference gel by using least-square minimization. In some applications, we do not actually need to minimize the sum of distances between the source landmarks and the transformed landmarks. As long as the distances are not bigger than some threshold, or the transformed landmarks gather together in some small local area, we would rather leave more flexibility to the parameters so that the model could fit better globally. For this reason, we will formulate this as a constraint in our model rather than adding it as a penalty to the objective function as most of the algorithms do.

Several warping methods have been implemented (see, for example Glasbey and Mardia [1998]): non-parametric transformations like elastic deformations and thin plate splines, and parametric transformations like the polynomial transformations

$$\begin{cases} u = \sum_{i=0}^l \sum_{j=0}^{l-i} a_{ij} x^i y^j \\ v = \sum_{i=0}^l \sum_{j=0}^{l-i} b_{ij} x^i y^j. \end{cases} \quad (1)$$

Some software packages (Appel *et al.*, 1997) assume that the distortion is direction independent in the sense that warping along the horizontal axis depends only on the horizontal position and warping along

the vertical axis depends only on vertical position:

$$u = f(x), \quad v = g(y).$$

However, we found that this assumption is not always realistic and that it does not lead to significant decrease in computational cost. In the present paper, we consider the more general form (1).

### 3. GLOBAL AFFINE TRANSFORMATIONS

Given a set of gel samples with selected landmarks, our goal is to find a transformation for each gel sample so that the corresponding landmarks after transformation will gather together in a relatively small area on a master gel image. The problem can be stated as follows: We are given  $M$  gel samples with  $N$  landmarks on each sample. The positions of landmark centers on each gel are given. In order to analyze the samples, these  $N$  gels have to be comparable. We want to find transformations  $T^{(i)} : R^2 \rightarrow R^2$ , for gel sample  $i$ ,  $i = 1, \dots, M$ , so that every set of deformed landmarks will be superimposed on an ideal landmark on a master gel, with some allowed error tolerance. The ideal landmarks on the master gel are considered as unknowns as are the transformation coefficients.

One of the most fundamental transformations is the affine transformation. It has been widely used in image warping and point matching (see, for example, Glasbey and Mardia [1998]). Let us consider a global affine transformation for each sample:

$$T^{(i)}(l) = A^{(i)}l + b^{(i)}, \tag{2}$$

where

$$l = \begin{pmatrix} x \\ y \end{pmatrix}, \quad A^{(i)} = \begin{pmatrix} \alpha^{(i)} & \beta^{(i)} \\ \delta^{(i)} & \varphi^{(i)} \end{pmatrix}, \quad b^{(i)} = \begin{pmatrix} \gamma^{(i)} \\ \psi^{(i)} \end{pmatrix}, \quad i = 1, \dots, M.$$

We denote the  $j$ th landmark on the  $i$ th gel sample by

$$l_{ij} = \begin{pmatrix} x_{ij} \\ y_{ij} \end{pmatrix}, \quad i = 1, \dots, M, j = 1, \dots, N,$$

and denote the landmark after transformation by

$$T^{(i)}(l_{ij}) = \begin{pmatrix} \alpha^{(i)}x_{ij} + \beta^{(i)}y_{ij} + \gamma^{(i)} \\ \delta^{(i)}x_{ij} + \varphi^{(i)}y_{ij} + \psi^{(i)} \end{pmatrix}.$$

Since we do not know the coordinates of the landmarks on the master gel, we assign an ideal position

$$l_j = \begin{pmatrix} x_j \\ y_j \end{pmatrix}$$

for the  $j$ th landmark on the master gel. Thus, we assume that  $l_{ij}$  is the deformation of  $l_j$  on sample  $i$ .

In many cases, gel samples are deformed only slightly from the master gel. We thus search for transformations that will keep the changes for the landmarks as small as possible while satisfying the constraints we describe later. This will be implemented in our model by adding to the objective function a term of the form

$$\sum_{i=1}^M (\|A^{(i)} - I\|_F^2 + \|b^{(i)}\|_F^2) = \sum_{i=1}^M ((\alpha^{(i)} - 1)^2 + (\beta^{(i)})^2 + (\gamma^{(i)})^2 + (\delta^{(i)})^2 + (\varphi^{(i)} - 1)^2 + (\psi^{(i)})^2).$$

This will ensure the fact that from all possible affine transformations satisfying the constraints, we will select the one that is closest to the identity transformation in Frobenius norm.

Since we assume that  $l_{ij}$  is the deformation of  $l_j$  on sample  $i$ , we expect that the ideal landmark  $l_j$  is close to the average  $\bar{l}_j = \begin{pmatrix} \bar{x}_j \\ \bar{y}_j \end{pmatrix}$  of  $l_{1j}, l_{2j}, \dots, l_{Mj}$ . We address this concern by adding the terms

$$\|l_j - \bar{l}_j\|_2^2 = (x_j - \bar{x}_j)^2 + (y_j - \bar{y}_j)^2$$

to the objective function and impose the constraints

$$\|l_j - \bar{l}_j\|_\infty \leq E_j,$$

where the  $E_j$ 's are some constant tolerances that may be landmark dependent.

Ideally, the transformed landmarks should perfectly match the ideal landmark on the master gel:  $T^{(i)}(l_{ij}) = l_j$ . However, the coordinates of these landmarks are usually just approximations of the centers of the chosen protein spots on the images. As long as the distance between the transformed point and the ideal landmark does not exceed some error tolerance, we consider that the transformed points cover the ideal landmark. Let  $\epsilon_{ij}$  be the error tolerance for the difference between the transformed  $j$ th landmark on the  $i$ th sample and the ideal  $j$ th landmark  $l_j$ . We obtain the following constraints

$$\|T^{(i)}(l_{ij}) - l_j\|_\infty \leq \epsilon_{ij},$$

which are equivalent to

$$\begin{aligned} -\epsilon_{ij} &\leq \alpha^{(i)}x_{ij} + \beta^{(i)}y_{ij} + \gamma^{(i)} - x_j \leq \epsilon_{ij}, & i = 1, \dots, M \\ -\epsilon_{ij} &\leq \delta^{(i)}x_{ij} + \varphi^{(i)}y_{ij} + \psi^{(i)} - y_j \leq \epsilon_{ij}, & j = 1, \dots, N. \end{aligned}$$

We thus have the following optimization problem:

$$\min_{A^{(i)}, b^{(i)}, l_j} \sum_{i=1}^M \left( \|A^{(i)} - I\|_F^2 + \|b^{(i)}\|_F^2 \right) + \omega \sum_{j=1}^N \|l_j - \bar{l}_j\|_2^2 \quad (3)$$

$$\text{s.t. } \|T^{(i)}(l_{ij}) - l_j\|_\infty \leq \epsilon_{ij}, \quad i = 1, \dots, M$$

$$\|l_j - \bar{l}_j\|_\infty \leq E_j, \quad j = 1, \dots, N$$

where  $\omega$  is a constant weight to capture the relative importance of each part of the objective function and  $T^{(i)}$  is a transformation defined by  $A^{(i)}$  and  $b^{(i)}$  in (2). Notice that (3) is a quadratic programming problem (QP) and may be infeasible unless the inequalities are sufficiently relaxed (for example, by taking  $\epsilon_{ij}$  or  $E_j$  large enough). However, if we have only three points on each sample ( $N = 3$ ), then the linear system

$$\alpha^{(i)}x_{ij} + \beta^{(i)}y_{ij} + \gamma^{(i)} = x_j, \quad j = 1, 2, 3$$

$$\delta^{(i)}x_{ij} + \varphi^{(i)}y_{ij} + \psi^{(i)} = y_j,$$

has a solution for each  $i$ , which implies (3) is always feasible (even when  $\epsilon_{ij} = 0$ ). For a given tolerance, the fewer landmarks on each sample, the more likely it is to obtain a feasible QP. On the other hand, if the number of landmarks is too small, we do not have enough information for constructing a suitable warping function capable to compensate for all the deformations. In order to address these conflicting demands, in the following sections we consider global transformations constructed from local affine transformations which allow more degrees of freedom, and we will show that a feasible solution can always be found *for any given tolerance*.

#### 4. PIECEWISE AFFINE TRANSFORMATIONS USING NONUNIFORM SEGMENTATION

The first model (3) based on a global affine transformation is likely to be infeasible if the  $\epsilon_{ij}$ 's are small. We next consider a global transformation constructed from segmented local affine transformations in the sense that we divide the gels into segments and perform the affine transformation on each segment. Similar ideas have been addressed before: Goshtasby partitioned the image into triangles having the landmarks as vertices (Goshtasby, 1986). On each triangle he defined an affine transformation in such a way that the resulting global transformation is continuous on the convex hull of the landmarks. However, the global transformation may be discontinuous outside this convex hull. In the present paper, we consider partitioning the gels into rectangles and defining an affine transformation on each rectangle such that the resulting global transformation is continuous on the *entire gel*. Moreover, the global transformation is optimal in the sense that it has minimal curvature. This means that among all piecewise affine transformations that align the landmarks, our transformation accomplishes this with minimal distortion.

We sort the landmarks in ascending order according to their  $x$ -coordinates, so that  $x_{i1} \leq x_{i2} \dots \leq x_{iN}$ . For each sample  $i$ , we assume that the image is a rectangle  $[\check{t}_i^x, \hat{t}_i^x] \times [\check{t}_i^y, \hat{t}_i^y]$ . We compute  $s_{ij} = (x_{i,j-1} + x_{ij})/2$  for  $j = 2, \dots, N$ , and we define  $s_{i1} = \check{t}_i^x$  and  $s_{i,N+1} = \hat{t}_i^x$ . Let us consider the segment

$$\Pi_j^{(i)} = \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \in R^2, s_{ij} \leq x \leq s_{i,j+1} \right\}, \quad i = 1, \dots, M, \quad j = 1, \dots, N,$$

and define the corresponding local affine transformation  $T_{x,j}^{(i)} : \Pi_j^{(i)} \rightarrow R^2$  as

$$T_{x,j}^{(i)}(l) = A_{x,j}^{(i)}l + b_{x,j}^{(i)} = \begin{pmatrix} \alpha_{x,j}^{(i)}x + \beta_{x,j}^{(i)}y + \gamma_{x,j}^{(i)} \\ \delta_{x,j}^{(i)}x + \varphi_{x,j}^{(i)}y + \psi_{x,j}^{(i)} \end{pmatrix},$$

where

$$A_{x,j}^{(i)} = \begin{pmatrix} \alpha_{x,j}^{(i)} & \beta_{x,j}^{(i)} \\ \delta_{x,j}^{(i)} & \varphi_{x,j}^{(i)} \end{pmatrix}, \quad b_{x,j}^{(i)} = \begin{pmatrix} \gamma_{x,j}^{(i)} \\ \psi_{x,j}^{(i)} \end{pmatrix}.$$

Figure 2 shows such a segmentation. In order to have a global continuous transformation, we need to add the following constraints to ensure continuity on the common boundaries of adjacent segments:

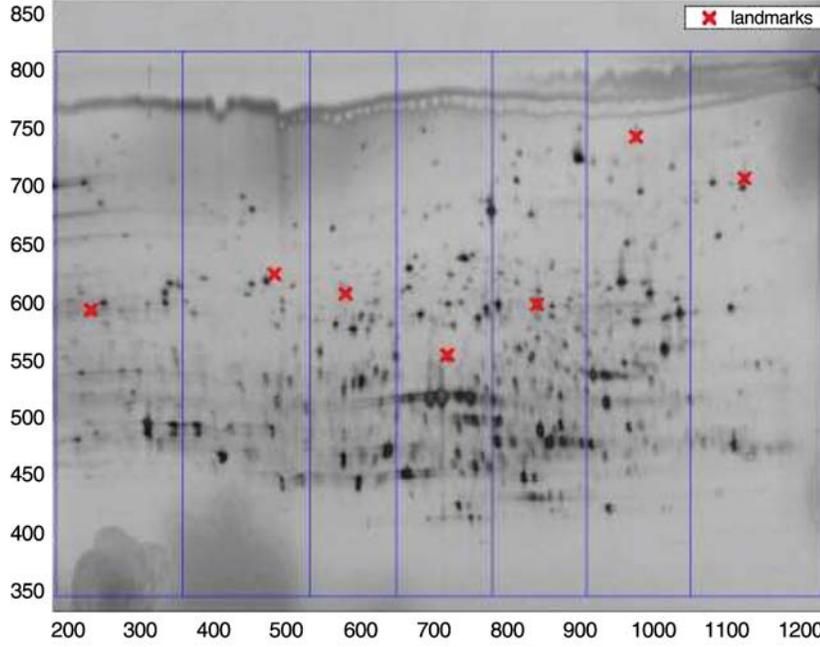
$$T_{x,j-1}^{(i)}(p_{ij}) = T_{x,j}^{(i)}(p_{ij}), \quad j = 2, \dots, N, \quad (4)$$

where  $p_{ij} = \begin{pmatrix} s_{ij} \\ y \end{pmatrix}$  is an arbitrary point on the common boundary of  $\Pi_{j-1}^{(i)}$  and  $\Pi_j^{(i)}$ . Since the equality holds for any  $y$ , (4) yields the following continuity constraints:

$$\begin{aligned} \alpha_{x,j-1}^{(i)}s_{ij} + \gamma_{x,j-1}^{(i)} &= \alpha_{x,j}^{(i)}s_{ij} + \gamma_{x,j}^{(i)}, \\ \beta_{x,j-1}^{(i)} &= \beta_{x,j}^{(i)}, \\ \delta_{x,j-1}^{(i)}s_{ij} + \psi_{x,j-1}^{(i)} &= \delta_{x,j}^{(i)}s_{ij} + \psi_{x,j}^{(i)}, \\ \varphi_{x,j-1}^{(i)} &= \varphi_{x,j}^{(i)}, \end{aligned} \quad (5)$$

for  $i = 1, \dots, M, j = 2, \dots, N$ . Our objective function is obtained by adding to the objective function from (3) a term of the form

$$S = \sum_{i=1}^M \sum_{j=2}^N c_{ij} \left( (\alpha_{x,j-1}^{(i)} - \alpha_{x,j}^{(i)})^2 + (\delta_{x,j-1}^{(i)} - \delta_{x,j}^{(i)})^2 \right), \quad (6)$$



**FIG. 2.** Segmentation of the gel sample according to the  $x$  coordinates of the landmarks, which are shown as crosses.

where  $c_{ij} = \omega_s / (s_{i,j-1} - s_{i,j+1})^2$ ,  $\omega_s$  is a constant weight, and  $\mathcal{S}$  represents a measure of the curvature of the global mapping. We thus arrive to the following QP:

$$\min_{A_{x,j}^{(i)}, b_{x,j}^{(i)}, \check{l}_j} \mathcal{S} + \sum_{i=1}^M \sum_{j=1}^N \left( \|A_{x,j}^{(i)} - I\|_F^2 + \|b_{x,j}^{(i)}\|_F^2 \right) + \omega \sum_{j=1}^N \|\check{l}_j - \bar{l}_j\|_2^2 \quad (7)$$

s.t. the continuity constraints (5) hold

$$\|T_{x,j}^{(i)}(l_{ij}) - \check{l}_j\|_\infty \leq \epsilon_{ij}, \quad i = 1, \dots, M$$

$$\|\check{l}_j - \bar{l}_j\|_\infty \leq E_j, \quad j = 1, \dots, N$$

where  $\check{l}_j$  is the  $j$ th ideal landmark for this model. There are  $6MN + 2N$  unknowns in (7). When  $\epsilon_{ij} = E_j = 0$ , (7) has  $2N + 2MN + 4M(N - 1) = 6MN + 2N - 4M$  equality constraints. Moreover there are at most three constraints for each set of  $(\alpha_{x,j}^{(i)}, \beta_{x,j}^{(i)}, \gamma_{x,j}^{(i)})$  or  $(\delta_{x,j}^{(i)}, \varphi_{x,j}^{(i)}, \psi_{x,j}^{(i)})$ . We thus claim that (7) is always feasible.

We next repeat the above procedure by segmenting along the  $y$  direction. We now assume  $y_{i1} \leq y_{i2} \dots \leq y_{iN}$ . For each sample  $i$ , we compute  $t_{ij} = (y_{i,j-1} + y_{ij})/2$  for  $j = 2, \dots, N$ , and we define  $t_{i1} = \check{\tau}_i^y$  and  $t_{i,N+1} = \hat{\tau}_i^y$ . Let us consider the segment

$$\Sigma_j^{(i)} = \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, t_{ij} \leq y \leq t_{i,j+1} \right\}, \quad i = 1, \dots, M, \quad j = 1, \dots, N,$$

and define the corresponding local affine transformation  $T_{y,j}^{(i)} : \Sigma_j^{(i)} \rightarrow \mathbb{R}^2$  as

$$T_{y,j}^{(i)}(z) = \begin{pmatrix} \alpha_{y,j}^{(i)}x + \beta_{y,j}^{(i)}y + \gamma_{y,j}^{(i)} \\ \delta_{y,j}^{(i)}x + \varphi_{y,j}^{(i)}y + \psi_{y,j}^{(i)} \end{pmatrix}.$$

In a similar manner, we construct a QP based on the partition along the  $y$ -direction. The ideal landmarks  $\hat{l}_j$  in this model, along with the transformations  $T_{y,j}^{(i)}$ , are obtained as the solution of this QP. Figure 3 shows a sample gel image with both  $x$ - and  $y$ -direction partitions. Finally, we can combine the segmentation in the  $x$ -direction with the segmentation in the  $y$ -direction in order to obtain a more balanced transformation. Given a set of landmarks  $l_{ij}, i = 1, \dots, M, j = 1, \dots, N$ , we compute  $s_{ij}$  and  $t_{ij}$  for  $i = 1, \dots, M, j = 1, \dots, N + 1$  as described previously. We then divide each gel into  $N^2$  segments: for  $i = 1, \dots, M, j = 1, \dots, N, k = 1, \dots, N$ , we define  $\Omega_{jk}^{(i)}$  as

$$\Omega_{jk}^{(i)} = \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \in R^2, s_{ij} \leq x \leq s_{i,j+1}, t_{ik} \leq y \leq t_{i,k+1} \right\}$$

and the corresponding local affine transformation  $T_{jk}^{(i)} : \Omega_{jk}^{(i)} \rightarrow R^2$  as a convex combination of the two corresponding transformations, i.e.,

$$T_{jk}^{(i)}(z) = \lambda T_{x,j}^{(i)}(z) + (1 - \lambda) T_{y,k}^{(i)}(z), \quad i = 1, \dots, M, \quad j = 1, \dots, N, \quad k = 1, \dots, N,$$

where  $\lambda \in [0, 1]$  is a constant and  $T_{x,j}^{(i)}, T_{y,k}^{(i)}$  are the solutions of the QP models based on  $x$ - and  $y$ -direction partitions, respectively. Finally we define a global transformation  $T^{(i)}$  for gel sample  $i$  as

$$T^{(i)} = T_{jk}^{(i)} \quad \forall z \in \Omega_{jk}^{(i)}, \tag{8}$$

and the ideal landmarks as

$$l_j = \lambda \check{l}_j + (1 - \lambda) \hat{l}_j, \quad j = 1, \dots, N.$$

It is easy to check that  $T^{(i)}$  satisfies constraints similar to those in the QP models based on the  $x$ - and  $y$ -direction partitions. The transformation  $T^{(i)}$  is thus a global piecewise affine transformation which aligns the landmarks with some error tolerance and is globally continuous on the whole gel sample, even beyond the convex hull of the given landmarks. Moreover, a weighted mean of its curvature and distance to the identity is optimized. Figure 4 shows the image in Fig. 3 under such a transformation.

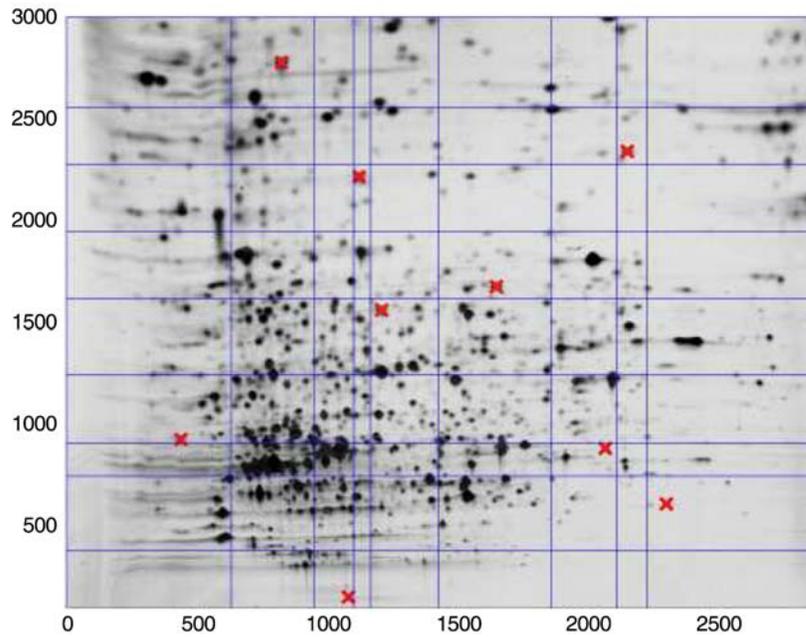


FIG. 3. 2D segmentation of the gel sample according to both  $x$  and  $y$  coordinates of the landmarks, which are shown as crosses.

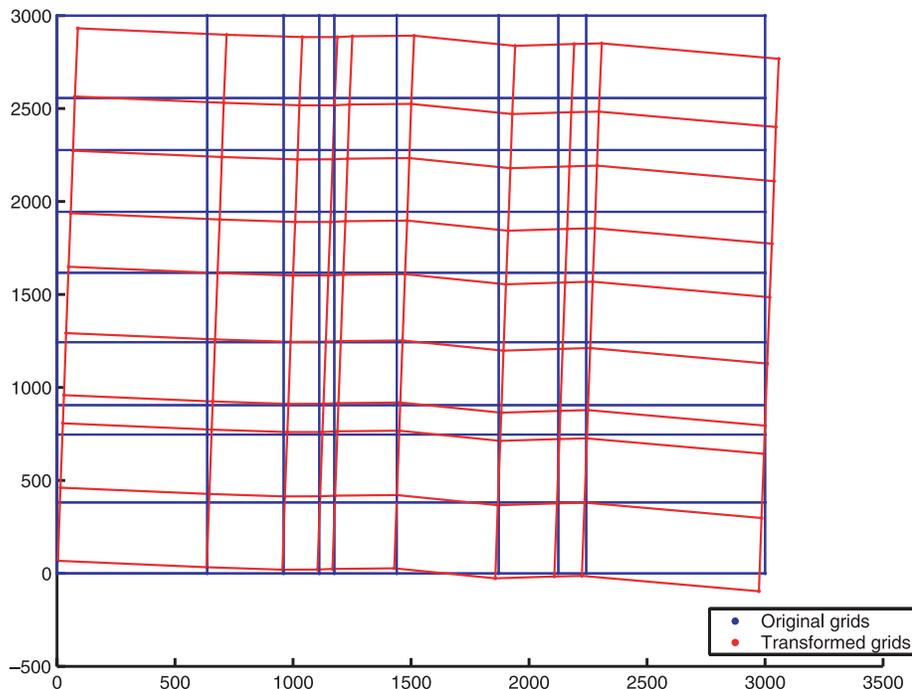


FIG. 4. The same grid before and after transformation.

## 5. HIERARCHICAL PIECEWISE AFFINE TRANSFORMATIONS USING UNIFORM GRID

In the previous section, a globally continuous piecewise affine transformation was considered based on a segmentation of the gel that depends on the distribution of the landmarks. It could happen that some segments are much larger than others if the coordinates of the landmarks are not uniformly distributed. Moreover, the resulting partition of the gel may not be fine enough to capture all possible local deformations present in the gel. In the above framework it is difficult to further refine the partition. In order to address this limitation, we propose a hierarchical approach based on constructing for each gel  $i$  a sequence of increasingly finer partitions  $\mathcal{P}_1^{(i)}, \mathcal{P}_2^{(i)}, \dots$ , that do not depend on the distribution of the landmarks. Partition  $\mathcal{P}_\kappa^{(i)}$  consists in dividing gel  $i$  into  $p_{i,\kappa}q_{i,\kappa}$  rectangles where  $p_{i,1}, p_{i,2}, \dots, p_{i,\kappa}, \dots$  and  $q_{i,1}, q_{i,2}, \dots, q_{i,\kappa}, \dots$  are two increasing sequences of positive integers. In our application, we take  $p_{i,\kappa+1} = 2p_{i,\kappa}$ ,  $q_{i,\kappa+1} = 2q_{i,\kappa}$ , so that the number of rectangles increases by a factor of 4 at each grid refinement. Similar hierarchical partitions have been considered in the context of pairwise alignment via a least-square optimization approach by Salmi *et al.* (2002).

Let us describe a typical partition. For simplicity, we drop the index  $\kappa$ , and we assume that gel sample  $i$  is included in the rectangle  $\Omega^{(i)} = [\check{\tau}_i^x, \hat{\tau}_i^x] \times [\check{\tau}_i^y, \hat{\tau}_i^y]$ . This rectangle is divided into  $p_i \times q_i$  equal rectangles

$$\Omega_{jk}^{(i)} = \left\{ z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, s_{ij} \leq x \leq s_{i,j+1}, t_{ik} \leq y \leq t_{i,k+1} \right\}, j = 1, \dots, p_i, k = 1, \dots, q_i,$$

with vertices

$$s_{ij} = \check{\tau}_i^x + (j-1) \frac{\hat{\tau}_i^x - \check{\tau}_i^x}{p_i}, j = 1, \dots, p_i + 1, \quad (9)$$

$$t_{ik} = \check{\tau}_i^y + (k-1) \frac{\hat{\tau}_i^y - \check{\tau}_i^y}{q_i}, k = 1, \dots, q_i + 1. \quad (10)$$

On each subrectangle, we define an affine transformation  $T_{jk}^{(i)} : \Omega_{jk}^{(i)} \rightarrow R^2$  as

$$T_{jk}^{(i)}(l) = A_{jk}^{(i)}l + b_{jk}^{(i)} = \begin{pmatrix} \alpha_{jk}^{(i)}x + \beta_{jk}^{(i)}y + \gamma_{jk}^{(i)} \\ \delta_{jk}^{(i)}x + \varphi_{jk}^{(i)}y + \psi_{jk}^{(i)} \end{pmatrix}, \quad (11)$$

where

$$A_{jk}^{(i)} = \begin{pmatrix} \alpha_{jk}^{(i)} & \beta_{jk}^{(i)} \\ \delta_{jk}^{(i)} & \varphi_{jk}^{(i)} \end{pmatrix}, \quad b_{jk}^{(i)} = \begin{pmatrix} \gamma_{jk}^{(i)} \\ \psi_{jk}^{(i)} \end{pmatrix}.$$

In order to enforce continuity of the piecewise affine map, one could proceed as in the previous section. However, since affine transformations map lines into lines, it is sufficient to impose the following constraints involving the vertices  $v_{jk}^{(i)} = (s_{ij}, t_{ik})$ :

$$\begin{cases} T_{j-1,k-1}^{(i)}(v_{jk}^{(i)}) = T_{j-1,k}^{(i)}(v_{jk}^{(i)}) \\ T_{j-1,k}^{(i)}(v_{jk}^{(i)}) = T_{jk}^{(i)}(v_{jk}^{(i)}) \\ T_{jk}^{(i)}(v_{jk}^{(i)}) = T_{j,k-1}^{(i)}(v_{jk}^{(i)}) \end{cases}, \quad j = 2, \dots, p_i, k = 2, \dots, q_i, \quad (12)$$

$$\begin{cases} T_{1,k-1}^{(i)}(v_{1k}^{(i)}) = T_{1,k}^{(i)}(v_{1k}^{(i)}), & T_{p_i,k-1}^{(i)}(v_{p_i+1,k}^{(i)}) = T_{p_i,k}^{(i)}(v_{p_i+1,k}^{(i)}), & k = 2, \dots, q_i \\ T_{j-1,1}^{(i)}(v_{j1}^{(i)}) = T_{j,1}^{(i)}(v_{j1}^{(i)}), & T_{j-1,q_i}^{(i)}(v_{j,q_i+1}^{(i)}) = T_{j,q_i}^{(i)}(v_{j,q_i+1}^{(i)}), & j = 2, \dots, p_i \end{cases} \quad (13)$$

The constraints (12) involve all interior vertices, while the constraints (13) involve all boundary vertices that are common to different rectangles.

We consider the following measure of the total curvature of our family of piecewise affine transformations  $T^{(1)}, T^{(2)}, \dots, T^{(M)}$ :

$$\begin{aligned} \tilde{S} = & \sum_{i=1}^M \sum_{j=1}^{p_i} \sum_{k=2}^{q_i} \omega_s \left\| \frac{T_{jk}^{(i)}(v_{j,k+1}^{(i)}) + T_{j,k-1}^{(i)}(v_{j,k-1}^{(i)}) - 2T_{jk}^{(i)}(v_{jk}^{(i)})}{\Delta y_i^2} \right\|_2^2 \\ & + \sum_{i=1}^m \sum_{j=2}^{p_i} \sum_{k=1}^{q_i} \omega_s \left\| \frac{T_{jk}^{(i)}(v_{j+1,k}^{(i)}) + T_{j-1,k}^{(i)}(v_{j-1,k}^{(i)}) - 2T_{jk}^{(i)}(v_{jk}^{(i)})}{\Delta x_i^2} \right\|_2^2 \end{aligned} \quad (14)$$

The constraints  $\|T^{(i)}(l_{ij}) - l_j\| \leq \epsilon_{ij}$ , which ensure that all landmarks are mapped into prescribed neighborhoods of the ideal landmarks, become

$$\|T_{(m(i,j),n(i,j))}^{(i)}(l_{ij}) - l_j\|_\infty \leq \epsilon_{ij},$$

where the indices  $(m(i, j), n(i, j))$  are chosen such that

$$l_{ij} \in \Omega_{m(i,j),n(i,j)}^{(i)}.$$

We arrive thus at the following QP:

$$\min_{A_{jk}^{(i)}, b_{jk}^{(i)}, l_j} \tilde{\mathcal{S}} + \sum_{i=1}^M \sum_{j=1}^{p_i} \sum_{k=1}^{q_i} \left( \|A_{jk}^{(i)} - I\|_F^2 + \|b_{jk}^{(i)}\|_F^2 \right) + \omega \sum_{j=1}^N \|\check{l}_j - \bar{l}_j\|_2^2 \quad (15)$$

s.t. the continuity constraints (12) and (13) hold

$$\|T_{(m(i,j), n(i,j))}^{(i)}(l_{ij}) - l_j\|_\infty \leq \epsilon_{ij}, \quad i = 1, \dots, M$$

$$\|l_j - \bar{l}_j\|_\infty \leq E_j, \quad j = 1, \dots, N$$

Let us now analyze the feasibility of the above QP. Altogether we have  $6 \sum_{i=1}^M p_i q_i + 2N$  unknowns and

$$2N + 2MN + 6 \sum_{i=1}^M (p_i - 1)(q_i - 1) + 2 \sum_{i=1}^M (2(p_i - 1) + 2(q_i - 1))$$

$$= 6 \sum_{i=1}^M p_i q_i - 2 \sum_{i=1}^M (p_i + q_i) - 2M + 2N + 2MN$$

equality constraints when  $\epsilon_{ij} = E_j = 0$ . One of the requirements to ensure feasibility is that the total number of degrees of freedom is nonnegative, i.e.,

$$2 \sum_{i=1}^M (p_i + q_i) + 2M - 2MN \geq 0. \quad (16)$$

In our implementation, we take  $p_i = q_i = p$ , so that the above condition is guaranteed to be satisfied if  $p \geq N/2$ , although in practice we have feasible QP's for smaller values of  $p$ . We could have infeasible situations if there are more than three landmarks in one of the subrectangles  $\Omega_{jk}^{(i)}$ . However, this situation cannot happen if  $p$  is large enough.

Our hierarchical approach is described by the following pseudo-code.

#### Algorithm

Determine the size of the initial partition by choosing  $p_i = q_i = \rho$ .

#### repeat

    Compute  $s_{ij}$  and  $t_{ik}$  from (9) and (10);

**if** (15) is infeasible,

        Refine the grid by setting  $p_i = 2p_i, q_i = 2q_i$ ;

**else**

        Obtain  $T^{(i)}, i = 1, \dots, M$  and  $l_j, j = 1, \dots, N$ , from the solution of (15);

**if** some stopping criterion is satisfied or the maximum number of iterations is exceeded,

**Return**;

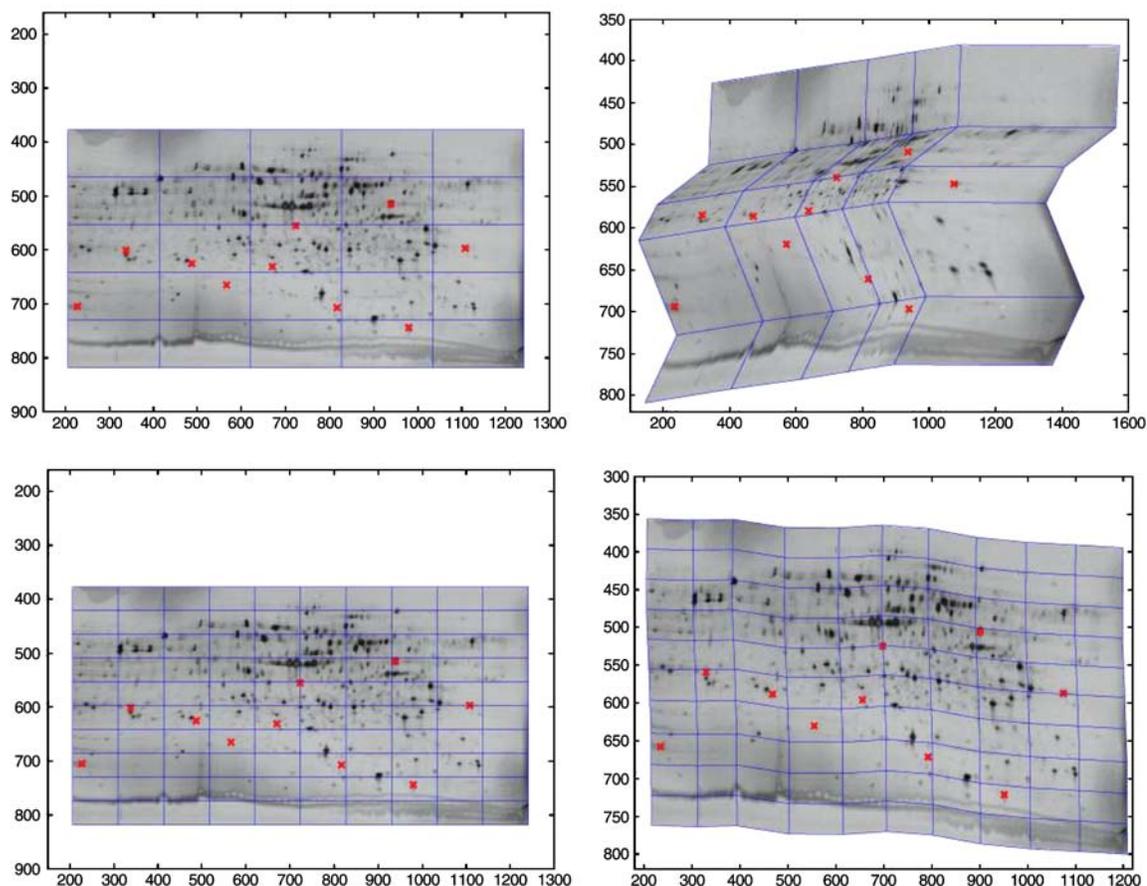
**else**

            Refine the grid by setting  $p_i = 2p_i, q_i = 2q_i$ ;

**endif**

**endif**

In our implementation, we are using an interior-point method for solving the QP that either finds the solution of the QP, or determines that the QP is infeasible in  $O(W^\nu)$  iterations (Ye, 1997), where  $W$  is number of variables of the QP and  $\nu$  is a constant. Each iteration requires  $O(W^\zeta)$  arithmetic operations, where  $2 \leq \zeta \leq 3$  is a constant depending on the sparsity of the problem. Assume that our algorithm terminates after  $k$  refinements, and denote by  $Z$  the number of variables of the QP for the initial grid. Since at each grid refinement, the number of variables increases by a factor of 4, it turns out that the total



**FIG. 5.** The top two figures show an image before and after transformation; the bottom two figures show the transformation of the same images after one refinement.

cost of the above algorithm is

$$CZ^\mu + C(4Z)^\mu + \dots + C(4^k Z)^\mu = CZ^\mu \frac{(4^\mu)^{k+1} - 1}{4^\mu - 1},$$

where  $\mu = \zeta + \nu$ . We conclude that the total cost of our hierarchical approach is only slightly bigger than the cost  $(4^\mu)^k CZ^\mu$  of solving the QP on the finest grid. Thus our hierarchical approach finds the coarsest grid for which the QP is feasible, and the corresponding solution satisfies our accuracy requirements at a cost which is roughly the same as the cost of solving one QP on the (a priori unknown) optimal grid.

Figure 5 shows a typical refinement step in the above algorithm. The top two figures show a gel sample before and after transformation, while the bottom two show the transformation of the same gel after one refinement. The refinement clearly helps in finding a better transformation.

## 6. NUMERICAL EXPERIMENTS

We tested our methods on two sets of protein gel samples.

For the first set of gels, wild type *E. coli* K12 strain W3110 (Coli Genetic Stock Center, Yale University, CT) were grown in a 20L BioFlow IV fermentor (New Brunswick Scientific Co., Inc., Edison, NJ) using glucose/mineral salts minimal medium. Samples were harvested during the exponential growth phase and fed-batch phase, and intracellular protein samples for two-dimensional polyacrylamide gel electrophoresis (2D-SDS PAGE) were prepared according to a protocol adapted from <http://ca.expasy.org/ch2d/protocols/>

*protocols.fn1.html*. Protein samples were focused in the first dimension based on their isoelectric point using 3-10 NL Immobiline pH Gradient (IPG) strips in a Multiphor II IEF apparatus (Amersham Biosciences, Piscataway, NJ), according to the manufacturer's instructions. The focused proteins were reduced and alkylated, by 20 min incubation in SDS equilibration buffer containing DTT and Iodoacetamide, respectively. Then, the strips were run in the second dimension to separate proteins based on their molecular weight using 1.5 mm thick large-format (18.5 cm  $\times$  19 cm) 12%T/2.67%C, continuous Tris-HCl linear gradient gels (10–100 kDa separation) in PROTEAN II XL electrophoresis cell (Bio-Rad laboratories, Hercules, CA), according to the manufacturer's instructions. Gels were stained using a neutral silver stain protocol adapted from Blum *et al.* (1987). Stained gels were digitalized at 400 dpi resolution using a GS-800 imaging densitometer (Bio-Rad Laboratories, Hercules, CA), and the images were exported to 8-bit TIFF format. Nineteen landmarks were picked on each of the 12 gel samples in this dataset. The left figure in Fig. 6 shows one of the gels with its landmarks.

The second set of gels were obtained from an experiment in which 12 rats (6 male, 6 female) were treated either with nicotine (3 male, 3 female) or with buffer control (3 male, 3 female). Subsequently, the animals were sacrificed and their spleen cells cultured (one spleen cell culture for each animal) *in vitro* in the presence of the T-cell mitogen concanavalin A. After stimulation, cells from each individual culture were harvested and proteins solubilized for 2D gel analysis. There are 12 gel samples with 20 landmarks. The right figure in Fig. 6 shows one of these gel images with landmarks.

For solving the QP's, we used the state-of-the-art interior-point software *MOSEK* ([www.mosek.com](http://www.mosek.com)).

All three methods (Global, 2D-Segmentation, Hierarchical Grid) were tested on both datasets. We chose 50% of the original landmarks randomly as training data and used the remaining as the test data. The transformation coefficients and the ideal positions of the training landmarks on the master gel can be found as the solution of the QP's described in the previous sections. For the test data, the mean of the transformed  $l_{ij}$  on different images,  $\tilde{l}_j = N^{-1} \sum_{i=1}^N T^{(i)}(l_{ij})$ , is considered to be the ideal position of the test landmark on the master gel. The error tolerance  $\epsilon_{ij}$  for the training landmarks were chosen to be the same:  $\epsilon_{ij} = \epsilon$ . The distances  $\tilde{\epsilon}_{ij} = \|T^{(i)}(l_{ij}) - \tilde{l}_j\|_{\infty}$  are considered to be the errors for the test data. The percentage of the test data that lies in a certain pixel range ([0,10], (10,20], etc.) is considered as a measurement of fit of the estimated transformation. Each experiment was repeated eight times and the mean value is reported in the following tables after cross-validation. The numerical results are reported in terms of  $P_1, P_2, P_3, P_4$ , where

$$P_1 = \text{percentage of the cases with } \tilde{\epsilon}_{ij} \in [0, 10],$$

$$P_2 = \text{percentage of the cases with } \tilde{\epsilon}_{ij} \in (10, 20],$$

$$P_3 = \text{percentage of the cases with } \tilde{\epsilon}_{ij} \in (20, 30],$$

$$P_4 = \text{percentage of the cases with } \tilde{\epsilon}_{ij} > 30.$$

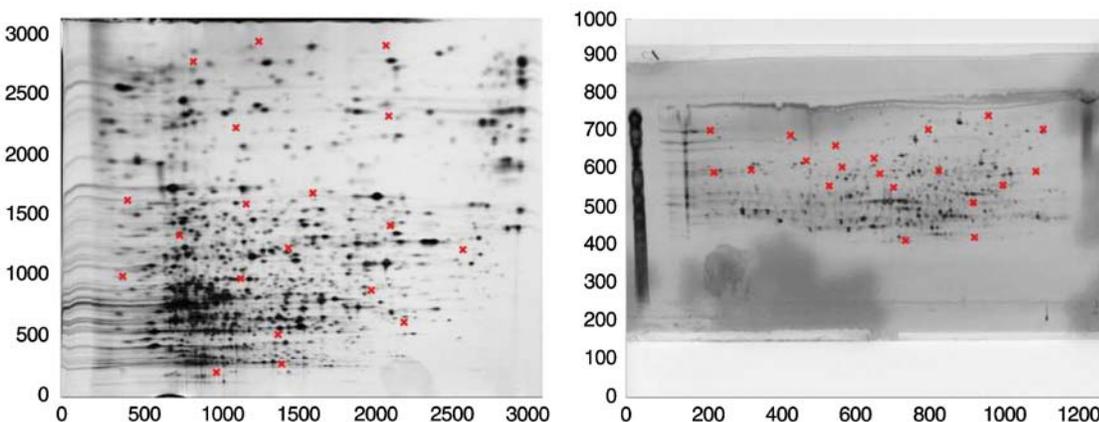


FIG. 6. Gel samples from dataset 1 and 2, where the landmarks are shown as crosses.

TABLE 1. RESULTS FOR DATASET 1<sup>a</sup>

	<i>Global affine</i>				<i>2D segmented affine</i>				<i>Hierarchical affine</i>			
	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$
$\epsilon = 0$		Infeasible			11.9	26.3	20.5	41.2	38.8	45.4	10.6	5.2
$\epsilon = 5$		Infeasible			17.5	35.0	18.3	29.2	57.4	30.3	8.5	3.8
$\epsilon = 10$		Infeasible			19.1	37.2	17.1	26.6	53.8	34.3	8.6	3.6
$\epsilon = 15$		Infeasible			20.1	32.8	15.2	31.9	63.7	28.1	6.8	1.4
$\epsilon = 20$	47.4	39.8	10.0	2.7	21.4	35.0	16.7	27.0	65.2	27.0	6.6	1.1

<sup>a</sup>Comparison of the global affine transformation, 2D segmented affine transformation, and hierarchical piecewise affine transformation, using randomized training data; every experiment was repeated eight times, and the mean is reported.

TABLE 2. RESULTS FOR DATASET 2<sup>a</sup>

	<i>Global affine</i>				<i>2D segmented affine</i>				<i>Hierarchical affine</i>			
	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$
$\epsilon = 0$		Infeasible			76.7	11.4	6.8	5.2	89.4	8.2	1.2	1.1
$\epsilon = 5$		Infeasible			84.2	14.2	1.5	0.1	96.8	3.2	0	0
$\epsilon = 10$	82.2	16.8	0.9	0.1	81.5	14.8	2.5	1.2	92.9	6.7	0.4	0
$\epsilon = 15$	75.4	19.8	3.7	1.0	79.3	14.5	2.9	3.3	95.9	3.9	0.2	0
$\epsilon = 20$	78.9	15.7	3.4	2.0	80.3	12.4	4.5	2.9	96.6	3.2	0.2	0

<sup>a</sup>Comparison of the global affine transformation, 2D segmented affine transformation, and hierarchical piecewise affine transformation, using randomized training data; every experiment was repeated eight times, and the mean is reported.

Tables 1 and 2 show the numerical results for both datasets using all three approaches (based on global affine, 2D segmented piecewise affine, and hierarchical piecewise affine transformations, respectively). As we can see, the approach based on the global affine transformation (Section 3) is not feasible when  $\epsilon$  is small. The 2D segmented piecewise affine transformation (Section 4) is less accurate than the hierarchical piecewise affine transformation, but has a much lower computational cost. The finest grid we used in the hierarchical grid approach is  $15 \times 15$  for dataset 1 and  $10 \times 10$  for dataset 2. We expect that the result can be improved further by refining the grid. However, the computational costs increase significantly. The numerical result for dataset 2 is clearly much better than the result for dataset 1. This is due to the fact that the deformations for dataset 2 are mainly translations and rotations, which are well modeled by affine transformations.

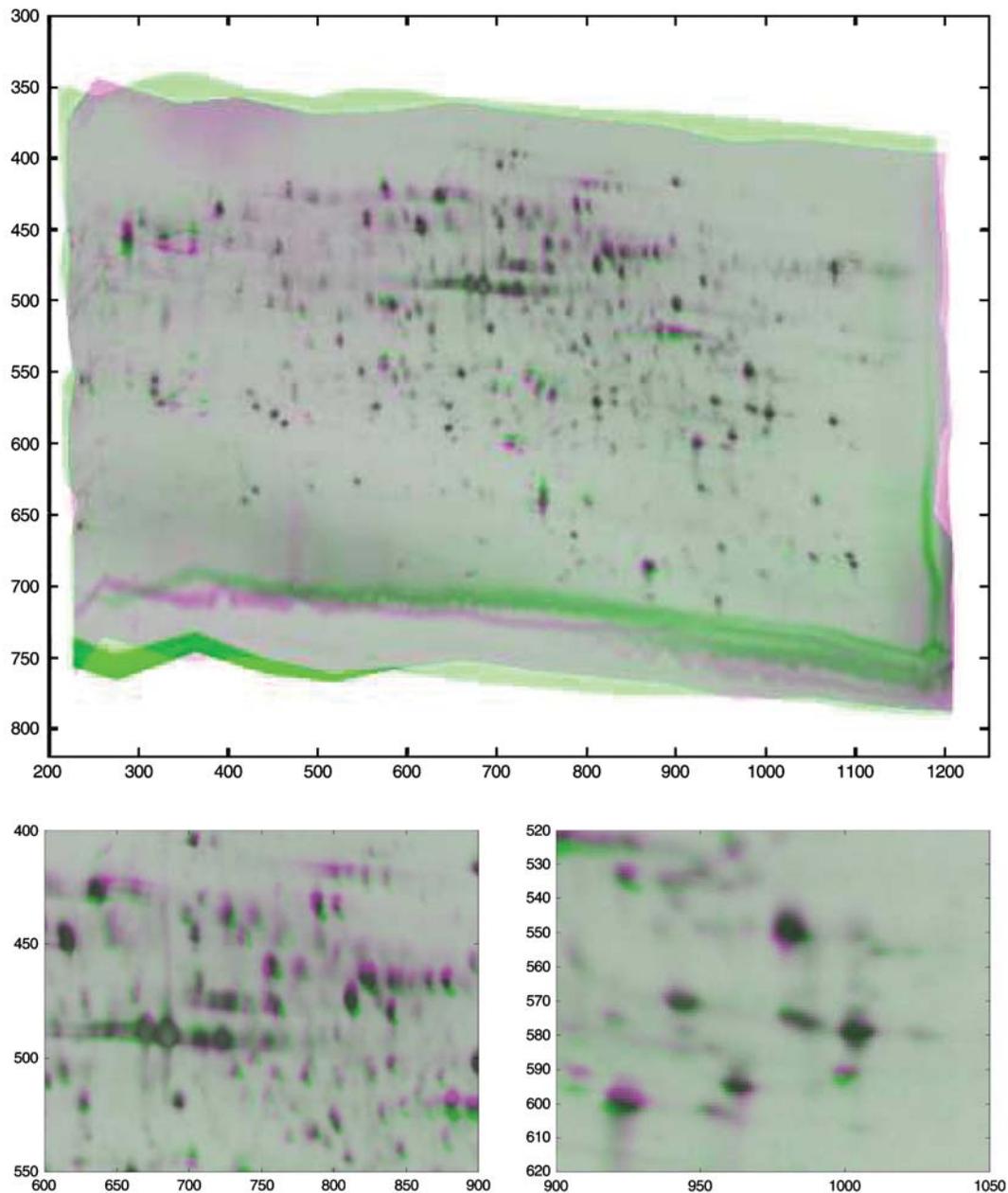
We also repeated our experiments using nonrandomized training data in the following sense. We sorted the landmarks in some order (according to their  $x$ -coordinates, for example) and picked one landmark in two as the training landmarks. By doing so, we obtain a better distribution of the training landmarks over the whole gel. The results are reported in Tables 3 and 4.

As expected, using this set of training landmarks gives much better results overall, especially for the 2D segmented piecewise affine transformations, which can take advantage of the favorable distribution of the training landmarks. This test is more relevant since in practice landmarks are chosen in such a way as to properly cover the gels.

For a graphical illustration of the performance of the hierarchical piecewise affine transformation, Fig. 7 shows the superimposition of the two images from Fig. 1 by this approach.

## 7. SUMMARY

In this paper, we have presented a methodology for aligning families of 2D gels that does not rely on choosing one of the gels as reference, but constructs an ideal gel and the corresponding warping transformations from the solution of a quadratic programming problem. We implemented this methodology



**FIG. 7.** Superimposition of the two images from Fig. 1 after hierarchical piecewise affine transformation.

by using global affine transformations, 2D segmented piecewise affine transformations, and hierarchical piecewise affine transformations. We conclude that the global affine transformation approach works only in case of very simple gel distortions. The 2D segmented affine transformation approach is always feasible and gives relatively good results at low computational costs. The hierarchical piecewise affine transformation approach gives the best results. However, it requires a substantial amount of computation. The piecewise affine warping functions obtained by the last two approaches are globally continuous and have optimized global curvature. Our methodology does not rely on common assumptions made in the literature, such as direction independent warping, and removes the dependence of the alignment methods on arbitrary selection of a reference gel.

TABLE 3. RESULTS FOR DATASET 1<sup>a</sup>

	<i>Global affine</i>				<i>2D segmented affine</i>				<i>Hierarchical affine</i>			
	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$
$\epsilon = 0$		Infeasible			21.0	54.0	20.2	4.8	39.8	38.1	16.9	5.1
$\epsilon = 5$		Infeasible			25.8	55.0	15.0	4.2	61.3	31.5	6.3	0.9
$\epsilon = 10$		Infeasible			27.1	56.8	12.7	3.4	69.1	26.4	3.6	0.9
$\epsilon = 15$	38.2	43.4	11.8	6.6	30.5	54.2	11.9	3.4	70.9	24.5	3.6	0.9
$\epsilon = 20$	34.2	46.8	12.7	6.3	34.5	49.6	11.8	4.2	70.9	24.5	3.6	0.9

<sup>a</sup>Comparison of the global affine transformation, 2D segmented affine transformation, and hierarchical piecewise affine transformation, using nonrandomized training data.

TABLE 4. RESULTS FOR DATASET 2<sup>a</sup>

	<i>Global affine</i>				<i>2D segmented affine</i>				<i>Hierarchical affine</i>			
	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$	$P_1$	$P_2$	$P_3$	$P_4$
$\epsilon = 0$		Infeasible			94.2	5.0	0.8	0	89.2	10.8	0	0
$\epsilon = 5$		Infeasible			95.8	4.2	0	0	98.3	1.7	0	0
$\epsilon = 10$	73.6	20.8	4.2	1.4	91.7	8.3	0	0	97.5	2.5	0	0
$\epsilon = 15$	72.2	19.4	4.2	4.2	95.0	5.0	0	0	98.3	1.7	0	0
$\epsilon = 20$	70.8	19.4	5.6	4.2	95.0	5.0	0	0	97.5	2.5	0	0

<sup>a</sup>Comparison of the global affine transformation, 2D segmented affine transformation, and hierarchical piecewise affine transformation, using nonrandomized data.

ACKNOWLEDGMENT

This research was partially funded by a grant (R01 GM075298) from the joint NSF-NIGMS program in mathematical biology.

REFERENCES

Appel, R.D., Vargas, K.R., Palagi, P.M., Walther, D., and Hochstrasser, D.F. 1997. Melanie II—A third-generation software package for analysis of two-dimensional electrophoresis images: II. Algorithms. *Electrophoresis* 18, 2735–2748.

Blum, H., Beier, H., and Gross, H.J. 1987. Silver stain protocol. *Electrophoresis* 8, 93–99.

Conradsen, K., and Pedersen J. 1992. Analysis of two-dimensional electrophoretic gels. *Biometrics* 48, 1273–1287.

Dowsey, A.W., Dunn, M.J., and Yang, G.-Z. 2003. The role of bioinformatics in two-dimensional gel electrophoresis. *Proteomics* 3, 1567–1596.

Glasbey, C.A., and Mardia, K.V. 1998. A review of image warping methods. *Appl. Statist.* 25, 155–171.

Goshtasby, A. 1986. Piecewise linear mapping functions for image registration. *Pattern Recogn.* 19, 459–466.

Gustafsson, J.S., Blomberg, A., and Rudemo, M. 2002. Warping two-dimensional electrophoresis gel images to correct for geometric distortions of the spot pattern. *Electrophoresis* 23, 1731–1744.

Klose, J. 1975. Protein mapping by combined isoelectric focusing and electrophoresis: A two-dimensional technique. *Humangenetik* 26, 231–234.

O’Farrell, P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* 250, 4007–4021.

Salmi, J., Aittokallio, T., Westerholm, J., Griese, M., Rosengren, A., Numan, T.A., Lahesmaa, R., and Nevalainen, O. 2002. Hierarchical grid transformation for image warping in the analysis of two-dimensional electrophoresis gels. *Proteomics* 2, 1504–1515.

Ye, Y. 1997. *Interior Point Algorithms: Theory and Analysis*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley and Sons, New York, NY.

Address correspondence to:  
*Françoise Seillier-Moiseiwitsch*  
*Dept. of Biostatistics, Bioinformatics and Biomathematics*  
*Georgetown University Medical Center*  
*Suite 180, Building D*  
*4000 Reservoir Road, N.W.*  
*Washington, DC 20057-1484*

*E-mail: seillier@georgetown.edu*