# ESTIMATION OF THIN PLATE SPLINE WARP PARAMETERS FROM PROTEIN SPOT POSITIONS IN 2D ELECTROPHORESIS GELS

*Lars Pedersen*

Informatics and Mathematical Modelling
Richard Petersens Plads, Building 321, Technical University of Denmark, DK-2800 Kgs. Lyngby
Denmark
e-mail: lap@imm.dtu.dk

**ABSTRACT**

The purpose of this study is to compare two types of image transformations, a first order polynomial (bilinear) transformation and a Thin Plate Spline transformation in the registration of 2D gel electrophoresis images. In the registrations protein spot positions serve as landmarks and the Normalised Mutual Information is used to evaluate the registration quality. The two types of transformations are compared using a varying number of landmarks in the transformation parameter estimation.

## 1. DATA, METHOD AND NOTATION

Consider two images, $I^p$ and $I^q$ with $N$ protein spots each. For each image we have a list of protein spots with 3 attributes for each spot: spot centre coordinates (x, y) and spot % integrated optical density (%IOD). We denote the spot centres (in homogeneous coordinates) from image $I^p$, $P^p = \{p_i\},\ i = 1, 2, \ldots, N$ where $p_i = (1, p_{ix}, p_{iy})$ and equally for image $I^q$, $P^q = \{q_i\},\ i = 1, 2, \ldots, N$ where $q_i = (1, q_{ix}, q_{iy})$. The %IOD's are denoted $D^p$ and $D^q$, respectively. The true correspondence between spots from $I^p$ and $I^q$ is known.

Initially the full-size images (approx. $2000 \times 2000$ pixels) and point sets have been aligned to eliminate a *global* first order bilinear transformation using 20 manually depicted spots as landmarks.

To ease the computational burden we examine sub-regions (approx. $450 \times 450$ pixels) of the images ($I_r^p$ and $I_r^q$) in order to estimate the transformations accounting for the remaining disparity in the regions. Ideally we would omit the global registration and instead estimate the transformations (TSP and first order polynomial) on the entire dataset (the full-size images). It is however quite time consuming to warp the full-size images, especially in the TPS case.

From the sub-regions $I_r^p$ and $I_r^q$ (Fig. 1), covering approximately the same area in the two gels, $L$ landmarks are selected. The landmarks $L_r^p$ are chosen from $P_r^p$ according to %IOD cardinality, i.e., the $L$ most intensive spots are used as landmarks. The corresponding spots in $P_r^q$ are selected using the known spot correspondence to form the set of landmarks in the sub region $I_r^q$, $L_r^q$.

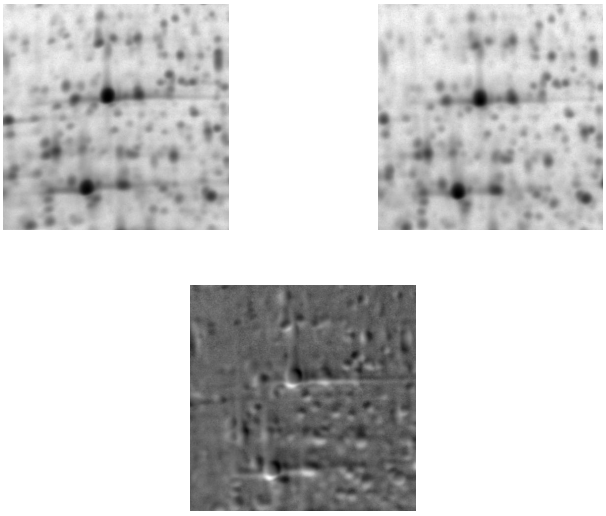

**Fig. 1:** Sub-regions $I_r^p$ (top left) and $I_r^q$ (top right). Bottom: difference image $I_r^p - I_r^q$.

These landmarks are used to estimate both a TPS transformation and a first order polynomial transformation to align the sub-regions, and the idea is to compare the performance of the two transformation methods.

### 1.1. Bilinear transformation

Consider the two 2D point sets of corresponding points (landmarks) $X = (x_i, y_i)$ and $Y = (u_i, v_i)$, $i = 1...L$, where $L$ is the number of points in the sets. We estimate the parameters in the first order polynomial transformation (or bilinear transformation) [3]:

$$u_i = a_{00} + a_{10}x_i + a_{01}y_i + a_{11}x_iy_i \quad (1)$$

$$v_i = b_{00} + b_{10}x_i + b_{01}y_I + b_{11}x_iy_i \quad (2)$$

The parameters are estimated using a least squares method and requires at least 4 sets of matching pairs to estimate

the eight parameters. Having found the $a$ and $b$ parameters we can now warp all pixels in an image according to this transformation. The bilinear warped version of an image $I_r^q$ is denoted $I_{r,B}^q$

### 1.2. TPS transformation

The Thin-Plate Spline (TPS) transformation is useful for mapping landmarks [1], and [2] proposes an elegant transition from the standard TPS functional to an energy function, that can be minimised.

Given two point sets $P$, and $Q$ in homogeneous coordinates, (i.e., $p_i = (1, p_{ix}, p_{iy})$ and $q_j = (1, q_{jx}, q_{jy})$) with each $L$ corresponding (landmark) points we wish to find a TPS transformation $f$ that minimises the TPS energy function

$$E_{TPS} = \sum_{i=1}^{L} \|p_i - f(q_i)\|^2 \quad + \\ \lambda \iint \left( (\frac{\partial f^2}{\partial^2 x})^2 + (\frac{\partial f^2}{\partial x \partial y})^2 + (\frac{\partial f^2}{\partial^2 y})^2 \right) dxdy, \quad (3)$$

where $f(q_i)$ is the TPS transformation (or warp) of the points in set $Q$. By minimising the first term the point set $Q$ is mapped as closely as possible to $P$. Including the second term in the minimisation imposes a limit on the second partial derivatives of $f$, i.e., a smoothness constraint on the bending energy of the spline. $f$ can be decomposed into an affine part and a non-affine part (e.g., [2]). For any point $s$ in $\mathbb{R}^2$ the TPS mapping of $s$ can be written as

$$f(s) = s \cdot d + \phi(s) \cdot w, \quad (4)$$

where $d$ is a $3 \times 3$ matrix of affine transformation parameters and $w$ is a $L \times 3$ matrix of non-affine warping coefficients. $\phi(s)$ is a $1 \times L$ vector for any point $s$,

$$\phi(s) = \begin{bmatrix} \phi_1(s) \\ \phi_2(s) \\ \vdots \\ \phi_L(s) \end{bmatrix} \quad (5)$$

where the i'th element is $\phi_i(s) = c\|s - q_i\|^2 \log \|s - q_i\|,\ i = 1, 2, \ldots, L$.

Following the TPS energy function formulation in [2] we minimise:

$$E_{TPS}(d, w) = \|P - Qd - \Phi w\|^2 \quad + \\ \lambda_1 trace(w^T \Phi w) + \lambda_2 trace[d - I]^T [d - I] \quad (6)$$

where $P$ and $Q$ are the corresponding point sets (landmarks), $d$ is the parameter matrix accounting for the affine part of the TPS transformation and $w$ is the non-affine part parameters. $\lambda_1$ and $\lambda_2$ are regularization parameters for the non-affine and the affine terms respectively. $\Phi$ is a $L \times L$ matrix formed from the $\phi$'s. Please refer to [2] for further details.

In the experiments, the regularization parameters for the TPS warp $\lambda_1 = 1$ and $\lambda_2 = 0.01$, are chosen relatively small to allow almost free behaviour of the TPS. $\lambda_2$ is set smaller than $\lambda_1$ to allow more freedom for the affine part of the transformation. If $\lambda_1 = \lambda_2 = 0$ we allow for maximum bending of the spline and the error at the landmarks will be small, but then the spline will not be smooth.

Having found $d$ and $w$ we can now warp all pixels in an image according to this transformation. The TPS warped version of an image $I_r^q$ is denoted $I_{r,TPS}^q$

### 1.3. Normalised Mutual Information

We use the Normalised Mutual Information (NMI) to measure how well two images are registered. The NMI of two images $M$ and $N$ is defined as [4]:

$$NMI(M, N) = \frac{1}{2} \frac{H(M) + H(N)}{H(M, N)} \quad (7)$$

where $H(M)$ and $H(N)$ are the marginal entropies of the two images $M$ and $N$ respectively and $H(M, N)$ is the joint entropy between the two images. The NMI is a real number between 0 and 1. The better $M$ and $N$ are registered the closer $NMI(M, N)$ is to 1. To measure the results of the warpings we compute the normalised mutual information before warping: $NMI(I_r^p, I_r^q)$. After bilinear warping: $NMI(I_r^p, I_{r,B}^q)$, and after TPS warping: $NMI(I_r^p, I_{r,TPS}^q)$.

Histograms are necessary in the computation of the entropies and ideally we should have a class for each image grey level, i.e., $2^N$ classes for a $N$-bit image. Since the joint histogram is a $2^N \times 2^N$ matrix this computation is infeasible for 16-bit images. The joint histogram will be a very large $65536 \times 65536$ matrix. Therefore,we are forced to use a lower number of classes in the computation of the histograms, i.e., to reduce the bit-depth if the images. For a given registration the NMI is rather sensitive to the number of bins used in the histogramming, but for a fixed number of classes it is still possible to compare NMI values from different registrations. In the experiments we have used 200 classes. This number is chosen as a tradeoff between large NMI values and low execution times.

## 2. EXPERIMENTS AND COMMENTS

For a different (increasing) number of landmarks we have computed the optimal parameters for both the bilinear and the TPS transformation and subsequently warped the sub images accordingly to the parameters.

Figs. 2 and 3 show example results using Bilinear and TPS warping to align the regions $I_r^p$ and $I_r^q$. The normalised mutual information (NMI) the regions before alignment is 0.5868 and after warping of the image the NMI is 0.6084 and 0.6411 for bilinear and TPS warp respectively (see also Table 1).

In general one would expect the NMI to increase after warping and also the NMI is expected to increase as the

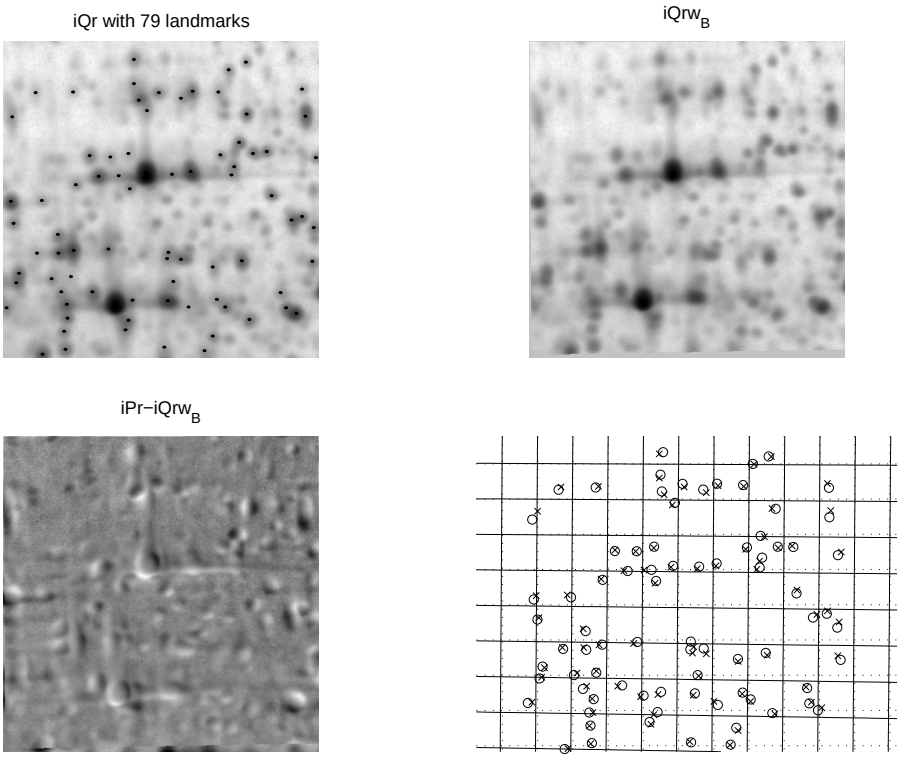


**Fig. 2:** Bilinear warping results using $\tau = 0.05$ (79 landmarks). Top left: image $I_r^q$ with landmarks. Top right: image $I_{r,B}^q$. Bottom left: difference image $I_r^p - I_{r,B}^q$. Bottom right: Bilinear warp of grid with original (x) and warped (o) landmarks.

number of landmarks is increased, i.e., when the landmarks used spread out and "cover" more and more of the entire image area.

In Fig. 4 is the difference images from Figs. 1, 2 and 3 is shown, scaled to the same gray level interval. It seems clear, that the TPS warp is superior to the bilinear registration.

| | | NMI | | |
|---|---|---|---|---|
| $\tau$ | # landmarks, $L$ | Before | After Bilinear | After TPS |
| 0.15 | 12 | .5868 | 0.6004 | 0.6209 |
| 0.10 | 30 | do | 0.6076 | 0.6241 |
| 0.05 | 79 | do | 0.6084 | 0.6411 |
| 0.01 | 240 | do | 0.6088 | 0.6527 |
| 0.005 | 280 | do | 0.6087 | 0.6531 |
| 0.001 | 309 | do | 0.6089 | 0.6551 |

**Table 1:** Selected experimental results.

Fig. 5 shows the number of landmarks and the NMI after bilinear and TPS warp of the images. Note that the NMI after TPS warp is always higher than the NMI after bilinear warp. Furthermore the NMI increases slightly as the number of landmarks increases. For the bilinear transformation the NMI is almost constant as function of the number of landmarks. In the computation of the bilinear transformation the problem is over determined for more than 4 cor-

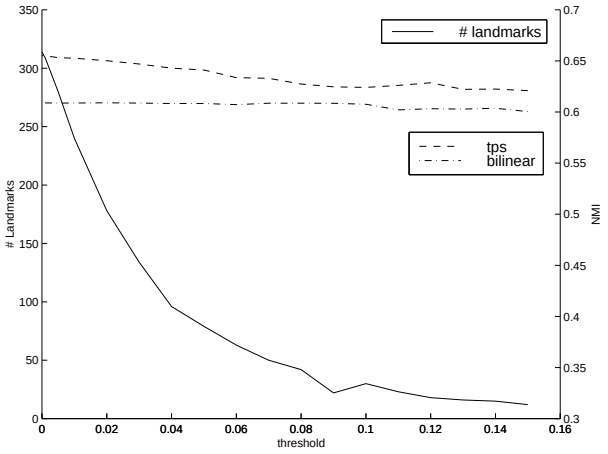


**Fig. 5:** Number of landmarks and normalised mutual information after TPS and bilinear warps as functions of %IOD threshold.

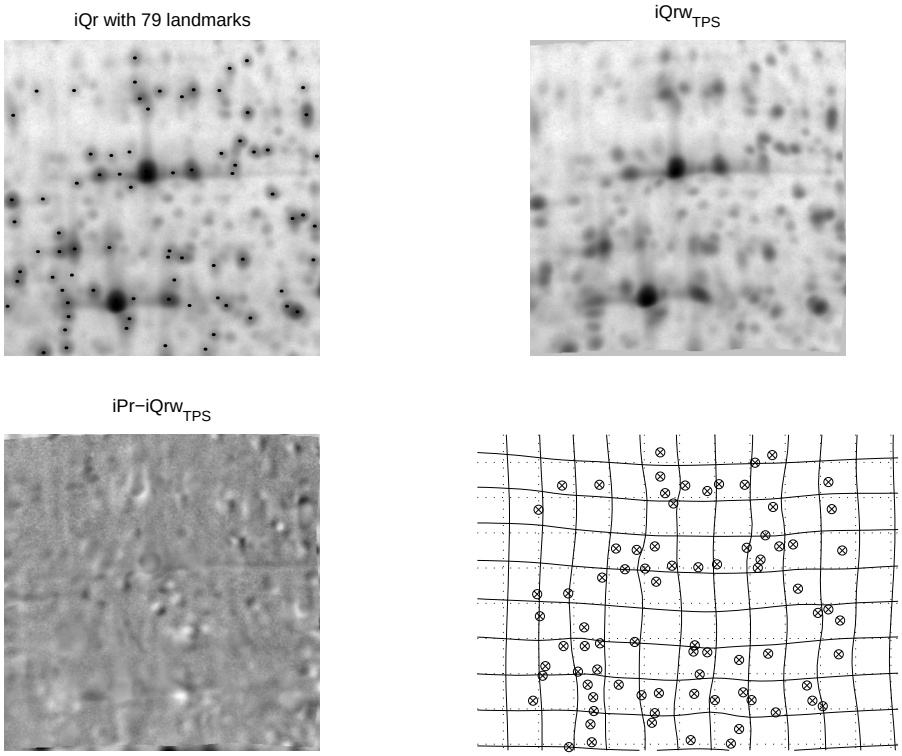


**Fig. 3:** TPS warping results using $\tau = 0.05$ (79 landmarks). Top left: image $I_r^q$ with landmarks. Top right: image $I_{r,TPS}^q$. Bottom left: difference image $I_r^p - I_{r,TPS}^q$. Bottom right: TPS warp of grid with original (x) and warped (o) landmarks.

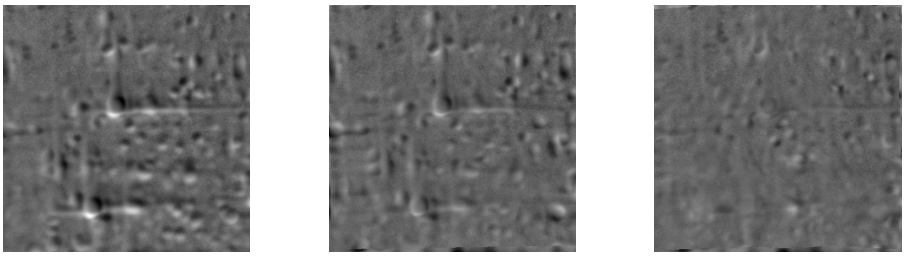

**Fig. 4:** Difference images in same grey level scaling. Left: originals $I_r^p - I_r^q$. Middle: after bilinear warp $I_r^p - I_{r,B}^q$. Right: after TPS warp $I_r^p - I_{r,TPS}^q$

responding landmarks and therefore the registration quality (NMI) is more or less constant as the number of landmarks increase. The TPS, however is capable of capturing all information from all landmarks and the quality of the registration continues to improve as the number of landmarks increase.

In future work it would be necessary to carry out more extensive experiments on more sub-areas of the gel images and on more gels. Also it would be interesting to compare the TPS with other spline functions.

## 3. REFERENCES


[1] F L Bookstein. Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[2] H Chui and A Rangarajan. A new algorithm for non-rigid point matching. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 44–51, 2000.

[3] C A Glasbey and K V Mardia. A review of image warping methods. *Journal of Applied Statistics*, 25:155–171, 1998.

[4] C Studholme, D L G Hill, and D J Hawkes. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition*, 32:71–86, 1999.