A new distance measure for non-rigid image matching

Anand Rangarajan[1], Haili Chui[1], and Eric Mjolsness[2]

[1] Departments of Diagnostic Radiology and Electrical Engineering,
Yale University,
New Haven, CT, USA
[2] Jet Propulsion Laboratory,
California Institute of Technology,
Pasadena, CA, USA

Abstract. We construct probabilistic generative models for the non-rigid matching of point-sets. Our formulation is explicitly Platonist. Beginning with a Platonist super point-set, we derive real-world point-sets through the application of four operations: i) spline-based warping, ii) addition of noise, iii) point removal and iii) amnesia regarding the point-to-point correspondences between the real-world point-sets and the Platonist source. Given this generative model, we are able to derive new non-quadratic distance measures w.r.t. the "forgotten" correspondences by a) eliminating the spline parameters from the generative model and by b) integrating out the Platonist super point-set. The result is a new non-quadratic distance measure which has the interpretation of weighted graph matching. The graphs are related in a straightfoward manner to the spline kernel used for non-rigid warping. Experimentally, we show that the new distance measure outperforms the conventional quadratic assignment distance measure when both distances use the same weighted graphs derived from the spline kernel.

1 Introduction

The need for non-rigid image matching arises in many domains within the field of computer vision. Some form of non-rigid matching is required to match objects that have undergone complex deformations. The extent of the deformation required to achieve non-rigid matching is an important quantity as it provides a convenient measure of distance between the two objects.

Non-rigid matching methods can be broadly divided into two categories; intensity-based and feature-based. Intensity-based methods attempt to calculate the optical flow between the two images. Usually, these methods require fairly strong brightness constancy assumptions between the two images. Feature-based methods attempt to match two sets of sparse features that have been extracted from the underlying image intensities. Usually, these methods require both object and deformation models in order to constrain the set of allowed matches and deformations.

Object models are typically constructed using a hierarchy of features: points, lines, curves, surfaces, etc. If matching is performed using generic, unlabeled point features, then the correspondence problem is acute. On the other hand, if high-level feature representations are used, the correspondence problem is alleviated but the matching is not likely to be robust against missing features. In addition, the constraints on the deformation model become more complex when high-level features are used.

In this paper, we are mainly concerned with deriving new distance measures for non-rigid matching of unlabeled point features. The new distance measure is a function of the unknown point-to-point correspondences and can handle outliers as well. Since we mostly focus on the new distance measure, at this point we are not presenting an algorithm to minimize this distance.

The formulation of our problem is explicitly Platonist. We begin by assuming a Platonist super point-set of unlabeled features. By using a probabilistic thin-plate spline warping model, we are able to generate real-world point feature sets. Outliers are explicitly modeled by forcing each real-world warped point-set to be a strict subset of the Platonist super point-set. The final step in this generative model is the loss of information of the point-to-point correspondences between the real-world point-set and the Platonist super point-set.

After exploiting a Platonist analogy in formulating this model, we then derive a new distance measure between the real-world point-sets. First, we eliminate the thin-plate spline warpings from the model by setting these parameters to their maximum *a posteriori* estimates. Then, in typical Bayesian fashion, we integrate out the hidden Platonist super point-set. The result is a new non-quadratic distance measure between all of the real-world point-sets defined solely in terms of the unknown correspondences.

Having derived the new non-quadratic distance measure, we present comparisons with the more traditional quadratic assignment distance measures. As a by-product of our derivation, we are able to show that the new distance measure is closely related to a weighted graph matching distance measure with the "graphs" determined by the thin-plate spline kernel. Both distance measures (quadratic and non-quadratic) use the same graphs derived from the spline kernels. Finally, we show that our new distance measure significantly outperforms the quadratic distance measure indicating a payoff resulting from our principled derivation.

2 Review

The various approaches to non-rigid image matching can be broadly grouped into two categories—intensity-based and feature-based. Intensity-based methods begin by assuming some form of brightness constancy reminiscent of optical flow methods [2]. Most methods in this class attempt to minimize an energy function that consists of two terms. The first term simply sums over the square of the intensity differences between the two images at each pixel. The second term is an elastic matching term which is typically derived from considerations of smoothness of the displacement field [7]. A free parameter is used to tradeoff between these two terms. The principal difficulty with this entire class of methods is that the brightness constancy assumption is frequently violated. Recently, there has been considerable interest in using entropy and mutual information-based intensity distance measures [30, 16] to overcome these limitations. A second problem with these methods is related to the lack of explicit object modeling. Since no attempt is made to construct object models, these methods cannot enforce *correspondence* constraints on structures that are *a priori* known to match. Recently [5], there has been some effort to overcome this limitation by including region segmentation information into the computation of optical flow. However, it is fair to say that at the present time, intensity-based image matching methods have yet to fully solve the aforementioned (two) problems by incorporating segmentation information into mutual information-based estimation of displacement fields (flow).

Feature-based image matching methods form the second class of methods. In contrast to the optical flow-based intensity matching methods, feature-based methods are more varied. One way of dividing the space of feature-based methods is along the lines of sparse versus dense features. Labeled landmark points are the most popular kind of sparse features since non-rigid matching of landmarks does not require a solution to the point-to-point correspondence problem For instance in [6], thin-plate splines (TPS) [31] are used to characterize the deformation of landmarks. Basically, the non-rigid matching problem is solved by minimizing the bending energy of a thin-plate spline while forcing corresponding landmarks extracted from the two images to perfectly match. Landmark positioning "jitter" can be accounted for in this model by allowing a tradeoff between the landmark position *least-squares* matching energy term and the spline bending energy term. This is analogous to the "vanilla" optical flow image matching method mentioned above. The major drawback of this method is the over-reliance on a few landmarks. Extracting labeled and corresponding landmarks from the two images is a difficult problem. Moreover, the method is quite sensitive to the number and choice of landmarks.

Dense feature-based matching methods run the gamut of matching points, lines, curves, surfaces and even volumes [4]. These methods usually begin with an object parameterization. Then, the allowable ways in which the object can deform is specified [18, 23]. The methods that fall into this class differ in object parameterizations and in the specification of the kinds of allowed deformations. In most cases, curves and/or surfaces are first fitted to features extracted from the images and then matched [18, 28, 27, 10]. These methods work well when the surfaces (and curves) to be matched are reasonably smooth. Also, the surface fitting step that precedes matching is predicated on good feature extraction. These methods have not been widely accepted in domains such as brain matching due to the extreme variability of cortical surfaces.

One of the principal reasons for the emphasis on object modeling in nonrigid matching is that it allows us to circumvent the correspondence problem. For example, once a smooth curve is fitted to a set of feature points, the matching can be taken up at the curve level rather than at the point level. Curve correspondence is easier than point correspondence [28] due to the strong constraint imposed by the smooth curve on the space of possible point-to-point correspondences. While the surface case is more complicated, surfaces can be approximately matched when they are smooth and the allowed deformations are not very complex [18, 27]. The downside is the lack of robustness. Sensor noise sometimes makes it difficult to fit smooth curves and surfaces to an underlying set of feature points. In such cases, while point feature locations may still be trustworthy, the fitting of surface normals and other higher-order features becomes problematic. Consequently, these higher-order features cannot be used.

In this paper, we begin with an integrated pose and correspondence formulation using point features. Essentially, we modify the pose parameters to include non-rigid deformations. We now turn to a review of recent approaches that attempt to integrate the search for correspondence in non-rigid matching. While the correspondence problem has a long history in rigid, affine and projective point matching [14], there is relatively a dearth of literature on non-rigid point matching. Recently, there has been some interest in using point-based correspondence strategies in non-rigid matching [23, 10, 29, 32, 21]. The modal matching approach in [23] relies on the point correspondence approach pioneered in [24] and further developed in [25]. The basic idea here is to use a pairing matrix that is built up from the Gaussian of the distances between any point feature in one set and the other. The modes of this matrix are used to obtain the correspondence. In [23], following [8], the deformation modes of the point-sets are obtained from the principal components of the covariance matrix of a pre-specified training set of shapes. The main drawback of this approach is that it does not use the *spatial relationships* between the points in each set to constrain the search for the correspondences and the mapping. In [9], after pointing out this drawback, the inter-relationships between the points is taken into account by building a graph representation from Delaunay triangulations. The search for correspondence is accomplished using inexact graph matching [26]. However, the spatial mappings are restricted to be affine or projective. In [1], decomposable graphs are hand-designed for deformable template matching and minimized with dynamic programming. However, the graphs are not automatically generated and there is no direct relationship between the deformable model and the graphs that are used. In [17], a maximum clique approach [20, 12] is used to match relational sulcal graphs. Again, the graphs are hand designed and not related to spatial deformations.

3 Deriving the distance measure

We first present background material on the thin-plate spline—our choice for the non-rigid spatial mapping. Then, we bring in the unknown correspondences and proceed with the derivation of the new distance measures.

3.1 Thin-plate splines

Our main reason for choosing the thin-plate spline is due its well understood behavior in landmark matching [6]. Essentially, the thin-plate spline produces a smoothly interpolated spatial mapping (with adherence to landmarks handled by the data term). The thin-plate spline (TPS) formulation required here is for 2D and 3D point matching. In both cases, we'll consider smoothness terms comprising of second-order derivatives of the interpolating function. Lack of space does not permit us to present the thin-plate spline in great detail. Instead, we merely present a "bare-bones" derivation. The interested reader is referred to [31] for the general formulation and to [6] for the application to landmark matching. In Figure 1, we depict an example thin-plate spline warping. Note the decomposition into the affine and warping components—a special characteristic of thin-plate spline mappings.

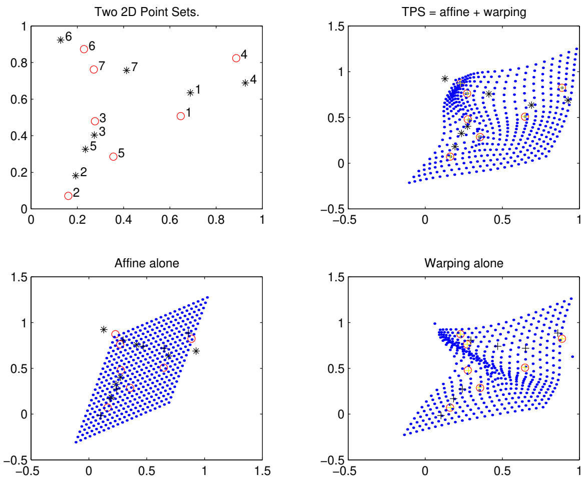

Fig. 1. Top Left: Original and warped point-sets. Top Right: Visualization of the thin-plate mapping. Bottom Left: Affine component of the mapping. Bottom right. Warping component of the mapping.

Assume for the moment that we have N pairs of corresponding points in either 2D or in 3D. Denote the two point-sets by Z and X respectively. The representations of the point-set Z is shown for the cases of 2D and 3D below:

$$\text{In 2D } Z = \begin{bmatrix} 1 & z_1^1 & z_1^2 \\ 1 & z_2^1 & z_2^2 \\ 1 & z_3^1 & z_3^2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & z_N^1 & z_N^2 \end{bmatrix}, \text{ and in 3D } Z = \begin{bmatrix} 1 & z_1^1 & z_1^2 & z_1^3 \\ 1 & z_2^1 & z_2^2 & z_2^3 \\ 1 & z_3^1 & z_3^2 & z_3^3 \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & z_N^1 & z_N^2 & z_N^3 \end{bmatrix} \quad (1)$$

A similar representation holds for the point-set X as well. The representations in (1) are the so called *homogeneous* coordinates.

We now set up a thin-plate spline mapping from X to Z. Thin-plate splines are asymmetric in the sense that a mapping from X to Z cannot be easily inverted to yield a mapping from Z to X. Minimizing the following energy function gives us a smooth spline interpolant capable of warping points in X arbitrarily close to points in Z. A regularization parameter λ determines the closeness of the fit. In 2D,

$$\begin{aligned} E_{\text{tps}}(f) &= \sum_{i=1}^{N} \|Z_i - f(X_i)\|^2 \\ &+\lambda \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\left(\frac{\partial^2 f}{\partial (x^1)^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x^1 \partial x^2} \right)^2 + \left(\frac{\partial^2 f}{\partial (x^2)^2} \right)^2 \right] dx^1 dx^2. \end{aligned} \tag{2}$$

A similar expression holds in 3D.

Define $t = (x^1, x^2, \ldots, x^D)$ where $D = 2$ for 2D and $D = 3$ for 3D. Then $t_i \stackrel{\text{def}}{=} (x_i^1, x_i^2, \ldots, x_i^D)$. Also, in 2D, $[\phi_1(t), \phi_2(t), \phi_3(t)] \stackrel{\text{def}}{=} [1, x^1, x^2]$ with a straightforward extension to 3D. For the thin-plate spline energy function given in (2), it is possible to show that there exists a unique minimizer f_λ given by

$$f_\lambda(t) = \sum_{k=1}^{D+1} d_k \phi_k(t) + \sum_{i=1}^{N} c_i E(t - t_i), \tag{3}$$

where $E(t - t_i)$ is the Green's function for the thin-plate spline: $E(\tau) = \tau^2 \log \tau$ in 2D and $-|\tau|$ in 3D. Here $|t - t_i| = \sqrt{\sum_{k=1}^{D} (x^k - x_i^k)^2}$. The minimizer f of the thin-plate spline energy function given in (3), is specified in terms of two unknowns c and d. Using (3), it is possible to eliminate f from the thin-plate spline energy function. When this is done, we get

$$E_{\text{tps2}}(c, d) = \|Z - Xd - Kc\|^2 + \lambda \text{ trace } c^T K c. \tag{4}$$

In (4), Z and X are the $N \times (D+1)$ point-sets, d is a $(D+1) \times (D+1)$ *affine* transformation consisting of translation, rotation and shear components, c is a $N \times (D+1)$ matrix of warping parameters (with all entries in the first column set to zero), and K is a $N \times N$ matrix corresponding to the Green's function (which is different for 2D and 3D). The principal difference between $E(t - t_i)$ and K is that the latter is defined only at the landmark points: the matrix entry $K_{ij} = E(t_i - t_j)$.

As it stands, finding least-squares solutions for the pair (c, d) by directly minimizing (4) is awkward. Instead, a QR decomposition is used to separate the affine and warping spaces. For more details, please see [31]:

$$X = [Q_1, Q_2] \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{5}$$

where Q_1 and Q_2 are $N \times (D+1)$ and $N \times (N - D - 1)$ orthonormal matrices, respectively. R is upper triangular. With this transformation in place, (4) becomes

$$E_{\text{tps-final}}(\gamma, d) = \|Q_2^T Z - Q_2^T K Q_2 \gamma\|^2 + \|Q_1^T Z - Rd - Q_1^T K Q_2 \gamma\|^2 + \lambda \gamma^T Q_2^T K Q_2 \gamma, \qquad (6)$$

where $c = Q_2\gamma$ and γ is a $(N-D-1) \times (D+1)$ matrix. Given this definition of c, $X^T c = 0$. The least-squares energy function in (6) can be first minimized w.r.t. γ and then w.r.t. the affine transformation d. The final result after minimization is

$$\hat{\gamma} = (Q_2^T K Q_2 + \lambda I_{(N-D-1)})^{-1} Q_2^T Z, \text{ and } \hat{d} = R^{-1}(Q_1^T X - K Q_2 \gamma). \qquad (7)$$

The bending energy of the thin-plate spline after eliminating (c, d) is

$$E_{\text{bending}}(Z) = \text{trace}\left[Z^T Q_2 (Q_2^T K Q_2 + \lambda I_{N-D-1})^{-1} Q_2^T Z \right]. \qquad (8)$$

3.2 A Platonist formulation

Having described the thin-plate spline spatial mapping in its two conventional (integral and matrix kernel) forms, we turn to the integrated pose and correspondence formulation.

First, we no longer assume that the correspondence between the point-sets Z and X is known. We introduce a correspondence matrix M which obeys the following constraints.

1. The correspondences are binary: $M_{ai} \in \{0, 1\}$.
2. Every point in X is matched to one point in Z: $\sum_a M_{ai} = 1$.
3. Every point in Z is matched to one point in X or is an *outlier* w.r.t. X: $\sum_i M_{ai} \leq 1$.

In informal terms, M is a matrix with binary entries whose columns sum to one and whose rows may either sum to one or be all zero. The one-to-one correspondence constraint is not sacrosanct and can be modified to a classification (many-to-one) constraint. This is because, in non-rigid mapping, a one-to-one constraint is incorrect when, for example, points are deformed into lines.

The bending energy of the thin-plate spline needs to be modified to take into account the introduction of the correspondence matrix M. Note that M allows us to generalize to the case of unequal point counts between X and Z.

We now write a probabilistic generative model for obtaining Z given X and a set of spline parameters (c, d). Since the correspondence M is unknown, it is included in the generative model as a hidden variable.

$$p(Z, M | X, c, d) = \frac{\exp\left[-E_1(Z, M, c, d)\right]}{Z_{\text{part1}}} \qquad (9)$$

where

$$E_1(Z, M, c, d) = \sum_{ai} M_{ai} \|Z_a - (Xd)_i - (Kc)_i\|^2. \qquad (10)$$

In (10), $(Xd)_i$ and $(Kc)_i$ are the i^{th} elements of the vectors Xd and Kc respectively. The partition function Z_{part1} is a normalization constant. Equation

(10) is reminiscent of a Gaussian mixture model with Z playing the role of the cluster centers and M being the complete data classification matrix [15]. If M is presumed known, the least-squares term in (10) reduces to the thin-plate spline least squares term.

The pure bending energy term is exactly the same as in the thin-plate spline:

$$p(c, d|X, \lambda) = \frac{\exp\left[-E_2(c)\right]}{Z_{\text{part2}}}, \text{ where } E_2(c) = \lambda \text{ trace } c^T K c. \tag{11}$$

The two energy terms in (10) and (11) can be combined into one. The resulting probabilistic generative model for Z can be written (after some algebraic manipulation) as

$$p(Z, M, c, d|X, \lambda) = \frac{\exp\left[-E(Z, M, c, d\right]}{Z_{\text{part}}}, \tag{12}$$

where

$$\begin{aligned} E(Z, M, c, d) = {} & \|MZ - Xd - Kc\|^2 + \lambda \text{ trace } c^T K c \\ & + \text{trace } Z^T[\text{diag}(\sum_i M_{ai}) - M^T M]Z. \end{aligned} \tag{13}$$

The diag operator above takes a vector and rearranges it into a square matrix with the vector entries appearing along the diagonal. The remaining entries are zero. The binary nature of the entries of M makes the last term redundant since it is zero. However, the term should be kept in mind if and when the binary constraint on the entries of M are relaxed; in that event, the last term becomes significant once again.

A few key observations can be made regarding the integrated correspondence-spline energy function in (13). After we define $Z_{\text{perm}} \stackrel{\text{def}}{=} MZ$, note that the form of the energy function is exactly the same as that of the original thin-plate spline bending energy in (4). (The last term does not contain the warping parameters (c, d) and from the perspective of solving for (c, d) the previous statement holds.) Consequently, we can exploit all of the properties of the thin-plate spline that were briefly derived in the previous section to separate out the warping and affine spaces. We can eliminate (c, d) from (13) and this step is quite similar to the work in [33]. The bending energy [after eliminating (c, d)] is

$$E_{\text{corr-bend}}(Z, M) = \text{trace}[Z^T M^T G M Z], \tag{14}$$

where

$$G \stackrel{\text{def}}{=} Q_2(Q_2^T K Q_2 + \lambda I_{N-D-1})^{-1} Q_2^T. \tag{15}$$

In deriving (14), we have dropped the last term in (13).

We are now in a position to extend this formulation to the simultaneous non-rigid matching of several point-sets $X^1, X^2, \ldots, X^K$. We find that the traditional Platonist metaphor suits us admirably. The point-set Z assumes the role of the "light beyond the cave" and each point-set X^k is cast in the role of a "shadow perceived on the cave wall." We model the Platonist super point-set Z as a superset of all the points present in each of the real-world point-sets X^k, $k \in \{1,\ldots,K\}$. We assume the following generative model for obtaining the real-world point-sets from the archetype Z. Each real-world point-set X^k is obtained by (i) warping Z using a thin-plate spline, (ii) removing a subset of points from Z, (iii) adding additive white Gaussian noise (AWGN) to the remaining points and finally, (iv) erasing or forgetting the correspondence information between Z and the newly created point-set X^k.

Since we have already worked out the bending energy expression [in (14)] between Z and a single real-world point-set X, we now extend the formulation to cover the simultaneous matching of Z to *all* of the real-world points-sets X^k, $k \in \{1,\ldots,K\}$. Henceforth, we denote the set comprising all real-world point-sets by $\mathbf{X}$. The sets of all correspondences, warping and affine parameters are denoted by $\mathbf{M}$, $\mathbf{c}$ and $\mathbf{d}$ respectively.

The likelihood model for the Platonist super point-set Z is

$$p(Z,\mathbf{M},\mathbf{c},\mathbf{d}|\mathbf{X}) = \frac{\exp\left[-\sum_{k=1}^{K} E(Z,M^k,c^k,d^k)\right]}{Z_{\text{partall}}} = \prod_{k=1}^{K} \frac{\exp\left[-E(Z,M^k,c^k,d^k)\right]}{Z_{\text{part}}^k}. \tag{16}$$

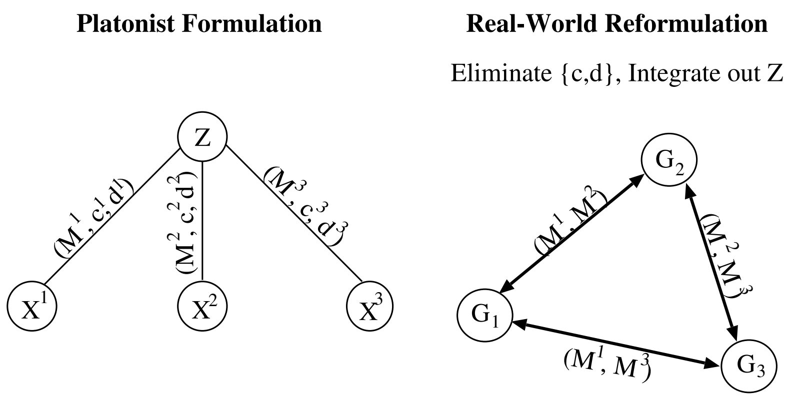

Fig. 2. Left: Platonist Formulation. Right: Real-world reformulation.

An important (and somewhat remarkable) fact about (16) is its separability. The Platonist super point-set Z is the sole bottleneck in the network of connections between the real-world point-sets $\mathbf{X}$. Consequently, with Z fixed, we can easily solve in closed-form for the entire set of thin-plate spline warping parameters $\mathbf{c}$ and $\mathbf{d}$. Note that the set of correspondence matrices $\mathbf{M}$ is also held fixed. This calculation is merely a generalization of the earlier calculation involving Z and X. Here, we have a set of point-sets $\mathbf{X}$ and Z. Our approach schema is depicted in Figure 2. On the left in Figure 2 is the original Platonist formulation with the Platonist super-point set Z acting as a generator for the point-sets X^k. On the right in Figure 2 is the real world reformulation. With $(\mathbf{c}, \mathbf{d})$ eliminated and Z integrated out, we obtain a distance measure between all of the real-world point-sets. Note that the point-sets X^k have been replaced by the corresponding graphs G^k.

3.3 Eliminating the spatial mapping

The spline parameter set $(\mathbf{c}, \mathbf{d})$ is eliminated exactly as before in (14). The only difference is that the elimination is carried out K times—once for each set of parameters $\{c^k, d^k\},\ k \in \{1, \ldots, K\}$. We will not repeat this derivation. The bending energy after eliminating $(\mathbf{c}, \mathbf{d})$ is now a sum over all K bending energies:

$$E_{\mathrm{corr-bend-total}}(Z, \mathbf{M}) = \sum_{k=1}^{K} \mathrm{trace}\ \left[Z^T (M^k)^T G^k M^k Z\right], \tag{17}$$

where

$$G^k \stackrel{\mathrm{def}}{=} Q_2^k \left[(Q_2^k)^T K^k Q_2^k + \lambda I_{N-D-1}\right]^{-1} (Q_2^k)^T.$$

With the above solution for the spatial mapping parameters $(\mathbf{c}, \mathbf{d})$, we may write the likelihood for Z as

$$p(Z, \mathbf{M}, \hat{\mathbf{c}}, \hat{\mathbf{d}} | \mathbf{X}) = \prod_{k=1}^{K} \frac{\exp\left[-E_{\mathrm{corr-bend-total}}(Z, M^k)\right]}{Z_{\mathrm{part}}^k}. \tag{18}$$

3.4 Integrating out the Platonist super point-set

Before integrating out Z, we wish to point out the need for this step. In a standard Bayesian formulation [3], integrating out the latent variables is recommended because the probabilistic structure is preserved by integration.

The distance measure between the real-world point-sets $\mathbf{X}$ is defined as

$$D(\mathbf{M}) \stackrel{\mathrm{def}}{=} -\log \int p(Z, \mathbf{M} | \mathbf{X}) dZ. \tag{19}$$

Note that the distance measure is a function of the unknown correspondences between each real-world point set X^k and the Platonist super point-set Z.

In (19), we have used $p(Z, \mathbf{M}|\mathbf{X})$ as shorthand for $p(Z, \mathbf{M}, \hat{\mathbf{c}}, \hat{\mathbf{d}}|\mathbf{X})$. The Platonist super point-set Z is now integrated out:

$$\begin{aligned} D(\mathbf{M}) &= -\log \int \exp\left[-Z^T \left(\sum_{k=1}^{K} (M^k)^T G^k M^k\right) Z\right] dZ + \text{ terms indep. of } \mathbf{M} \\ &= \frac{1}{2} \log \det \left[\sum_{k=1}^{K} (M^k)^T G^k M^k\right] \end{aligned} \tag{20}$$

This is our non-rigid matching distance measure. It is a function of only the set of correspondences $\mathbf{M}$. The thin-plate spline warping parameters have been eliminated and the Platonist super point-set Z has been integrated out.

We now specialize to the case of non-rigid matching of two point-sets X and Y. The distance measure between X and Y is

$$\boxed{D_{\text{log-det}}(M^X, M^Y) = \frac{1}{2} \log \det \left[(M^X)^T G^X M^X + (M^Y)^T G^Y M^Y \right],} \tag{21}$$

where the "graph" G^X is defined as

$$G^X = \lambda Q_2^X \left[(Q_2^X)^T K^X Q_2^X + \lambda I \right]^{-1} (Q_2^X)^T, \tag{22}$$

with a similar expression holding for G^Y. The graph G^X has the nice property that it is *symmetric and non-negative definite*. It can easily be made positive definite which aids in the computation of (21).

3.5 Comparison with traditional quadratic assignment distance measures

The new log-det distance measure in (21) can be directly compared with more traditional quadratic assignment (QAP) distance measures. The QAP distance measure is the obvious foil for comparison since it is the basic quadratic distance measure that is popular and widely used. All the comparisons below are based on the non-rigid matching of two point-sets X and Y. The QAP distance measure is a quadratic distance between the two "graphs" G^X and G^Y. Note that the derivation of the "graphs" from thin-plate spline kernels is a new contribution—one which is quasi-independent of the choice of distance between the two graphs. The QAP distance measure is

$$\boxed{D_{\text{qap}}(M^X, M^Y) = -\text{trace } (M^X)^T G^X M^X (M^Y)^T G^Y M^Y.} \tag{23}$$

Due to the cyclical property of the trace operator, the QAP distance can be simplified as $D(M^{XY}) = -\text{trace } (M^{XY})^T G^X M^{XY} G^Y$ where $M^{XY} \stackrel{\text{def}}{=} M^X (M^Y)^T$.

4 Results

Figures 3 and 4 compare the QAP distance with the new log-det distance. In Figure 3, we've compared the log-det distance measure with the quadratic distance measure. There are no outliers going from the Platonist super point-set to the two real-world point-sets shown at the top left of the figure. All three distance measures perform well with the log-det distance showing the greatest separation. In Figure 4, we've compared the log-det distance measure with the quadratic distance measure. Somewhat surprising is the degree to which the log-det distance measure outperforms the quadratic distance. On the x-axis, we've plotted permutations over a fixed number of points.

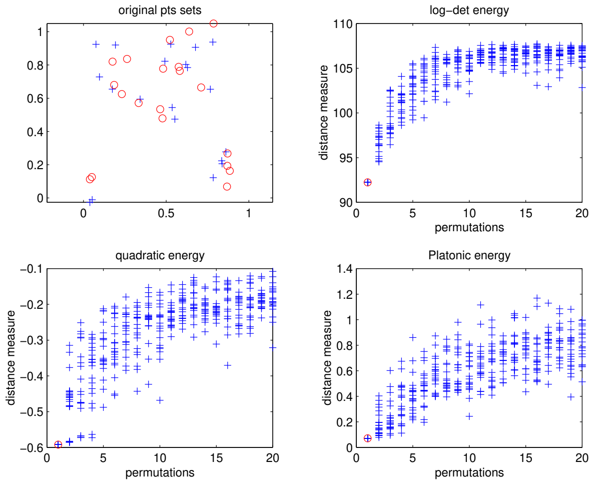

Fig. 3. Top Left: Two warped 2D 20 point-sets originating from a 20 point-set. Top Right: Log-det distance measure. Bottom Left: Quadratic distance measure. Bottom right: Platonic distance measure. The distance measures (ordinate) are plotted against permutations (abscissa). The abscissa value indicates how many points were permuted to obtain the distances. When zero points are permuted, the distance corresponding to the "true" answer is obtained [$\oplus$ is the true distance and $+$ is a distance point for a given permutation] and is plotted at the extreme left on the figure.

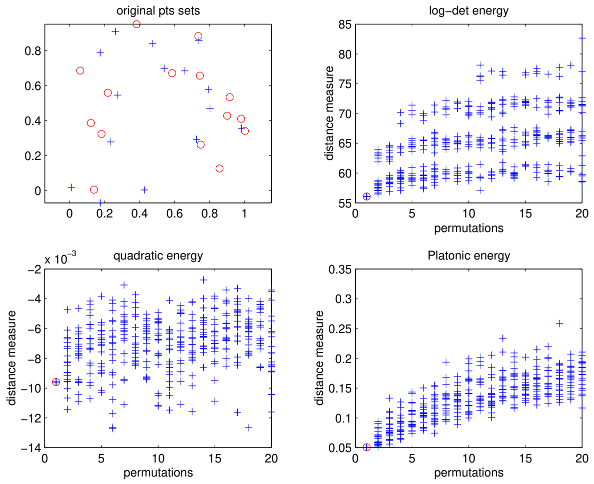

Fig. 4. Top Left: Two warped 2D 15 point-sets originating from a 20 point-set. Top Right: Log-det distance measure. Bottom Left: Quadratic distance measure. Bottom right: Platonic distance measure. The distance measures (ordinate) are plotted against permutations (abscissa). The abscissa value indicates how many points were permuted to obtain the distances. When zero points are permuted, the distance corresponding to the "true" answer is obtained [$\oplus$ is the true distance and $+$ is a distance point for a given permutation] and is plotted at the extreme left on the figure.

As more points are permuted, the distance measure ought to increase. We find that this is the case for the log-det distance measure but not for the quadratic distance. The Platonic distance is the quadratic distance between Z and the real-world point-sets. Since Z contains all the information, this idealized distance performs quite well. Note that there is obviously a question as to what the "true" answer ought to be. However, the bending energy returned by the log-det distance seems to concur with the Platonic bending energy which is reassuring since the latter is the closest you can get to a "gold standard."

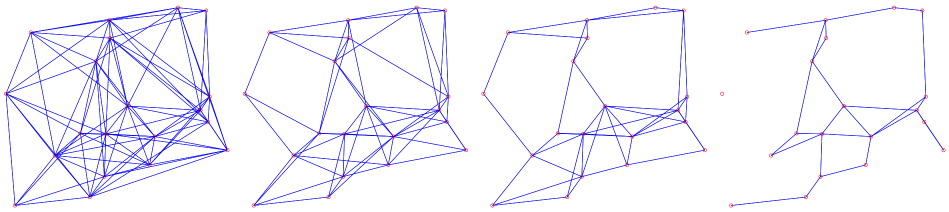

Fig. 5. The topology of the same weighted graph is shown using different thresholds. The node attributes and link weights are absent from this figure. As the threshold is increased (left to right), the topology becomes sparser as expected. The regularization parameter λ is held fixed while the threshold (for displaying a graph weight) is increased. The interplay between the regularization parameter and the graph topology needs to be further investigated.

In Figure 5, we show a point-set with 20 points and associated weighted graphs that have been derived from the spline kernel corresponding to the point-set. We wanted to explore the topology of the graph to see if the spline kernel seemed to use nearest neighbor heuristics in assigning weights. We thresholded the graphs (with increasing thresholds from left to right) after taking teh absolute value of each element. The topology clearly has a nearest neighbor bias which is more evident at larger thresholds. (The threshold for getting a certain number of connections is proportional to the regularization parameter λ.)

In Figure 6, we took a point-set and obtained several point-sets from it by progressively increasing the warping. We depict the topology (using the same threshold for all graphs) for all the warped point-sets. It should be clear from the figure that the topology gets increasingly distorted relative to the origianl topology as you go from smaller to larger warps. However, a family resemblence to the original parent is unmistakeable.

5 Discussion and Conclusion

The two main contributions of this paper are: i) a new definition of weighted graphs based on the thin-plate spline kernel and ii) a new non-quadratic distance measure that significantly outperforms the conventional quadratic assignment

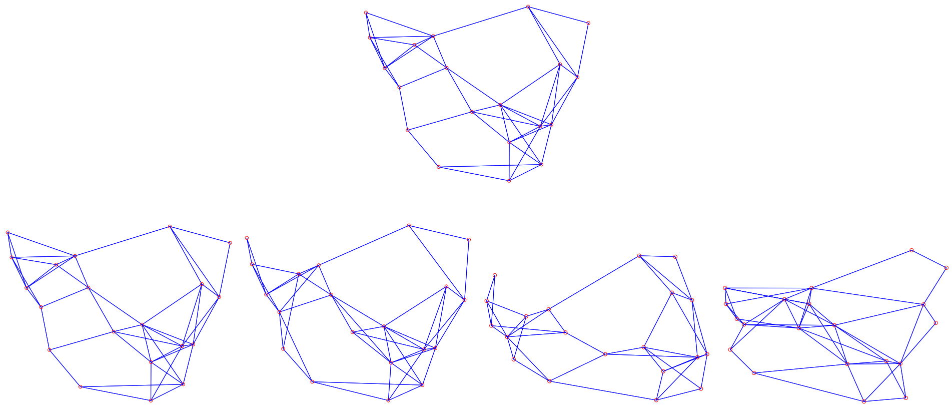

Fig. 6. On the top an original point-set is shown along with its graph depicted for a certain threshold. Below, we show four different thin-plate warped point-sets and their associated graphs. The warping increases as you go from left to right. Note the increased distortion in the topology going from left to right. The same threshold was used for all the graphs.

distance measure. To a certain extent, these two contributions are independent of one another. For instance, it should be possible to take our definition of weighted graphs and use a different distance measure. From the weighted graph standpoint, we have seen that the topology of the thin-plate spline kernel graphs (after thresholding) is somewhat similar to graphs derived from Delaunay triangulations with the important difference being that the spline-based graphs are not planar. The similarity stems from the fact that local connections are favored over more long range ones. We can also use different deformation mappings which should lead to different weighted graph definitions. For example, if we used a radial basis function (RBF) spline for the spatial mapping [33], the weighted graph would have a RBF kernel at its core. From the standpoint of the distance measure, we think that it is very significant that the new log-det distance measure outperforms the quadratic assignment distance. For binary graphs, it has already been shown that non-quadratic distance measures outperform QAP distances [11] and that seems to apply here as well. Enthusiasm must be tempered, however, until fast algorithms can be designed to find good, local minima of the new log-det distance measure.

There are several ways to proceed on the algorithm front. First, it may be possible to extend current deterministic annealing algorithms to the new distance. For instance, we could reduce the difficulty of algorithm design by choosing appropriate Legendre transformations [19] or by using Taylor series approximations of the log-det distance. Another approach would be to take the two topologies (after suitable thresholding) and apply the new maximum clique-based algorithms developed in [22]. After matching the topologies, further refinement using the weights can be performed using the softassign weighted graph matching algorithm [13].

In summary, we have shown that weighted graphs arise naturally in non-rigid point matching problems. The graphs directly depend on the parameterization of the deformations. In addition, we have found that a principled Bayesian Platonist formulation of the problem naturally leads to a new non-quadratic distance measure that outperforms the traditional quadratic assignment distance measure. It remains to be seen if effective algorithms can be designed that can take advantage of the better properties of the new distance measure.

Acknowledgements

A. R. would like to thank Zoubin Ghahramani for a helpful discussion. A. R. and H. C. are supported by a grant from the Whitaker Foundation.

References

1. Y. Amit. Graphical templates for model recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(4):225–236, 1996.
2. R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1–21, 1989.
3. J. Bernardo and A. Smith. *Bayesian Theory.* John Wiley and Sons, New York, NY, 1994.
4. P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Trans. Patt. Anal. Mach. Intell.*, 14(2):239–256, Feb. 1992.
5. M. Black and A. Jepson. Estimating optical flow in segmented images using variable-order parametric models and local deformations. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(10):972–986, 1996.
6. F. L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(6):567–585, June 1989.
7. G. Christensen, S. Joshi, and M. Miller. Volumetric transformation of brain anatomy. *IEEE Trans. Med. Imag.*, 16(6):864–877, 1997.
8. T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models: Their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.
9. A. D. J. Cross and E. R. Hancock. Graph matching with a dual-step EM algorithm. *IEEE Trans. Patt. Anal. Mach. Intell.*, 20(11):1236–1253, 1998.
10. J. Feldmar and N. Ayache. Rigid, affine and locally affine registration of free-form surfaces. *Intl. J. Computer Vision*, 18(2):99–119, May 1996.
11. A. M. Finch, R. C. Wilson, and E. R. Hancock. An energy function and continuous edit process for graph matching. *Neural Computation*, 10(7):1873–1894, 1998.
12. M. R. Garey and D. S. Johnson. *Computers and intractability: a guide to the theory of NP–completeness.* W. H. Freeman, San Francisco, CA, 1979.
13. S. Gold and A. Rangarajan. A graduated assignment algorithm for graph matching. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(4):377–388, 1996.
14. E. Grimson. *Object Recognition by Computer: The Role of Geometric Constraints.* MIT Press, Cambridge, MA, 1990.

15. T. Hofmann and J. M. Buhmann. Pairwise data clustering by deterministic annealing. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19(1):1–14, Jan. 1997.
16. B. Kim, J. L. Boes, K. A. Frey, and C. R. Meyer. Mutual information for automated unwarping of rat brain autoradiographs. *NeuroImage*, 5:31–40, 1997.
17. G. Lohmann and D. von Cramon. Sulcal basin and sulcal strings as new concepts for describing the human cortical topography. In *Workshop on Biomedical Image Analysis*, pages 41–54. IEEE Press, June 1998.
18. D. Metaxas, E. Koh, and N. I. Badler. Multi-level shape representation using global deformations and locally adaptive finite elements. *Intl. J. Computer Vision*, 25(1):49–61, 1997.
19. E. Mjolsness and C. Garrett. Algebraic transformations of objective functions. *Neural Networks*, 3:651–669, 1990.
20. H. Ogawa. Labeled point pattern matching by Delaunay triangulations and maximal cliques. *Pattern Recognition*, 19:35–40, 1986.
21. S. Pappu, S. Gold, and A. Rangarajan. A framework for non-rigid matching and correspondence. In D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, editors, *Advances in Neural Information Processing Systems 8*, pages 795–801. MIT Press, Cambridge, MA, 1996.
22. M. Pelillo. Replicator equations, maximal cliques and graph isomorphism. *Neural Computation*, 11, 1999. (in press).
23. S. Sclaroff and A. P. Pentland. Modal matching for correspondence and recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 17(6):545–561, Jun. 1995.
24. G. Scott and C. Longuet-Higgins. An algorithm for associating the features of two images. *Proc. Royal Society of London*, B244:21–26, 1991.
25. L. Shapiro and J. Brady. Feature-based correspondence: an eigenvector approach. *Image and Vision Computing*, 10:283–288, 1992.
26. L. G. Shapiro and R. M. Haralick. Structural descriptions and inexact matching. *IEEE Trans. Patt. Anal. Mach. Intell.*, 3(9):504–519, Sept. 1981.
27. R. Szeliski and S. Lavallee. Matching 3D anatomical surfaces with non-rigid deformations using octree splines. *Intl. J. Computer Vision*, 18:171–186, 1996.
28. H. Tagare, D. O'Shea, and A. Rangarajan. A geometric criterion for shape based non-rigid correspondence. In *Fifth Intl. Conf. Computer Vision (ICCV)*, pages 434–439, 1995.
29. S. Umeyama. Parameterized point pattern matching and its application to recognition of object families. *IEEE Trans. Patt. Anal. Mach. Intell.*, 15(1):136–144, Jan. 1993.
30. P. Viola and W. M. Wells III. Alignment by maximization of mutual information. In *Fifth Intl. Conf. Computer Vision (ICCV)*, pages 16–23. IEEE Press, 1995.
31. G. Wahba. *Spline models for observational data*. SIAM, Philadelphia, PA, 1990.
32. T. Wakahara. Shape matching using LAT and its application to handwritten character recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 16(6):618–629, 1994.
33. A. L. Yuille and N. M. Grzywacz. A mathematical analysis of the motion coherence theory. *Intl. J. Computer Vision*, 3(2):155–175, June 1989.

This article was processed using the LaTeX macro package with LLNCS style