A Bayesian Network Framework for Relational Shape Matching

Anand Rangarajan
University of Florida
Gainesville, FL US 32611
anand@cise.ufl.edu

James Coughlan
Smith-Kettlewell Institute
San Francisco, CA 94115
coughlan@ski.org

Alan L. Yuille
UCLA
Los Angeles, CA 90095
yuille@stat.ucla.edu

Abstract

A Bayesian network formulation for relational shape matching is presented. The main advantage of the relational shape matching approach is the obviation of the non-rigid spatial mappings used by recent non-rigid matching approaches. The basic variables that need to be estimated in the relational shape matching objective function are the global rotation and scale and the local displacements and correspondences. The new Bethe free energy *approach is used to estimate the pairwise correspondences between links of the template graphs and the data. The resulting framework is useful in both registration and recognition contexts. Results are shown on hand-drawn templates and on 2D transverse T1-weighted MR images.*

1. Introduction

Shape matching in its broadest context is a very difficult problem since it involves establishing distances between shapes that can be different due to intrinsic reasons—differing topologies, non-rigid deformations and the like—and due to extrinsic reasons—illumination changes, unknown correspondence, segmentation errors, outliers, missing and occluded information etc. However, if shape equivalences and distances in shape space can be constructed and more importantly efficiently computed, the payoff is huge since a host of applications ranging from non-rigid registration in biomedical and other imaging areas [27], object recognition [9], indexing in image databases and statistical shape analysis [5] awaits. In this work, we formulate a relational shape matching objective which at its core, casts shape matching as the warping of a template point-set with a fixed topology onto an unstructured data point-set. By leveraging recent developments in Bayesian networks—notably the Bethe free energy approach [30]—we show that efficient relational shape matching algorithms can be designed.

2 Previous Work

Taxonomically, the principal difference in both registration and recognition applications is between intensity- and feature-based methods. Since we are focused on feature-based methods, our review of intensity-based methods is cursory.

The main recent advance in intensity-based methods is the use of mutual information (MI) [28] as a distance measure for non-rigid registration [19]. Somewhat surprisingly, MI-based methods have generated much less interest in recognition applications. We speculate that as MI-based non-rigid matching methods improve in speed, they will become more viable for recognition. Or in more general terms, any efficient method for optical flow which avoids the brightness constancy assumption is simultaneously of interest to both registration and recognition.

There exist a plethora of approaches and methods for feature-based matching as applied to non-rigid registration and recognition. For a comprehensive review of feature matching approaches (with an emphasis on registration applications) please see [4]. Also the work in [10] is a particularly good review of feature-based matching in general.

Most of the feature-based recognition approaches begin with the fundamental assumption that the template and/or the data consists of a single closed curve or surface ([15, 22, 25, 26, 13, 12]). While these methods can be extended to some extent to the multiple curve/surface case, we were motivated by this previous work to propose a shape matching strategy that did not begin with this fundamental assumption. The *spin image* representation in [11] does not require segmentation of the data into surfaces despite the template being a single surface. The spin image is a novel shape attribute at each point on the template surface as is *shape context* [2]. In the latter, a joint histogram of inter-point angles and distances is used to develop an attribute vector at each point. The attribute vector is the shape context and is the quantity that is compared. The shape context approach solves for point correspondences using a weighted bipartite matching algorithm. And, in a move that is strikingly remi-

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

niscent of our previous TPS-RPM approach to non-rigid point matching [3], the template point-set is repeatedly warped using a thin-plate spline. The main drawback in the shape context approach and in all of our own previous work is the fundamental disconnect between the spline used for warping and the representation of the template point-set [17]. Unlike much earlier approaches such as [1, 20], in the work of [2, 3], there is no *relational* object model used for warping. While the shape context is a relational attribute, it plays *no role* in warping. In both cases, the thin-plate spline forces the template points to become correlated during warping [17], but this correlation is *completely independent of any underlying relational object representation*. In contrast, the work in [20] which is based on work in [21, 23] was the first to use correlated ways of movement of different features to assign correspondence. This paper is a spiritual cousin of [20] with the principal differences being the correspondence model and the minimization algorithms. Finally, very recent work in stereo matching [24] and object recognition [6] using Bayesian belief propagation [14] share a lot in common with this work.

After surveying the previous work in registration and recognition, it should be clear that we are focused on i) developing a common framework for both registration and recognition problems which ii) utilizes a relational template representation that is not restricted to merely a single curve or surface, iii) solves for correspondences and deformations using the underlying relational representation iv) resulting in a shape distance which can be used for non-rigid registration and recognition.

3 A Relational Correspondence and Deformation Approach to Non-rigid Shape Matching

In this section, we formulate the relational correspondence and deformation cost function for deforming a template graph onto an unlabeled data point-set. After formulating the cost function, we turn our attention to the prior in Section 3.2. Our notation is summarized in Table 1.

3.1 The relational shape matching (RSM) cost function

As mentioned in Section 2, one of the main motivations in this work is to move away from using non-rigid spatial mappings *that are essentially unrelated* to the object (or objects) that are being deformed. To this end, we assume that the template object is represented by an unlabeled point-set X (in $\mathcal{R}^2$ or in $\mathcal{R}^3$) and an undirected graph $G(V, E)$ (which characterizes the interconnection topology). In this setup, a closed digital curve would have a topology consisting of two neighbors for each point—one neighbor being the next point along the curve and the other neighbor being the previous point. Obviously, our use of a general graph G allows us to go beyond curves and surfaces as object representations. With this representation, the deformation of the object onto an unstructured data point-set can be posed as a *joint correspondence and deformation* problem. The relational matching cost function is written as

$$E_{\text{relational}}(M, U, s, R) = \sum_{ijab} M_{ia} M_{jb} G_{ij} D_{ijab} - \gamma \sum_{ijab} M_{ia} M_{jb} G_{ij}, \quad (1)$$

where $D_{ijab} = ||\sqrt{s}R[(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)] - \frac{1}{\sqrt{s}}(\mathbf{y}_a - \mathbf{y}_b)||^2$ with $M_{ia} \in \{0, 1\}$, and $\sum_i M_{ia} \leq 1$ and $\sum_a M_{ia} \leq 1$ being the constraints. In (1), we assume the existence of missing/extra points on both sides (template and data) and consequently, the need for outlier rejection arises. The parameter $\gamma > 0$ is a robustness parameter. The unknown variables in (1) are the correspondence M, the rotation R and scale s parameters and the deformation U. We use lowercase bold characters to denote vectors (such as $\mathbf{x}_i$) and uppercase characters to denote matrices [such as $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_1})$]. The relational shape matching cost function (1) attempts to deform a template (X, G) onto data (Y) using a deformation (U) and pose (s, R) in the presence of unknown correspondence (M). Please note that (1) is invariant under translations of X and Y. Typically, the graph G is quite sparse and this (as we shall see) plays a major role in the design of efficient minimization algorithms. We sometimes refer to the relational shape matching cost function as the likelihood since the deformation "prior" is quite important in this approach.

We can easily solve for the rotation and scale parameters in (1) by directly minimizing the cost function w.r.t. R and s respectively. This is one of the reasons why we have elected to not formulate (1) in an *intrinsic* [22] manner. For instance, we could have attempted to match distances $||(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)||$ and $||\mathbf{y}_a - \mathbf{y}_b||$ w.r.t. unknown scale (and no rotation), deformation and correspondence variables. However, since distances do not have the orientation information possessed by a vector [of the form $(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)$], we elected not to do so. Nonetheless, when we eliminate the rotation and scale parameters from (1) by minimizing w.r.t. them, the resulting equivalent cost function is invariant under *similarity* transformations of both the template and data. The solution for the global scale parameter is

$$\hat{s} = \sqrt{\frac{\sum_{ijab} G_{ij} M_{ia} M_{jb} ||\mathbf{y}_a - \mathbf{y}_b||^2}{\sum_{ijab} G_{ij} M_{ia} M_{jb} ||(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)||^2}}. \quad (2)$$

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

Table 1: Matrix and index notation for variables and constants

Variables and Constants	Index notation	Matrix notation	Dimension
Template	$\mathbf{x}_i,\ \forall i \in \{1,\ldots,N_1\}$	X	$D \times N_1$
Data	$\mathbf{y}_a,\ \forall a \in \{1,\ldots,N_2\}$	Y	$D \times N_2$
Deformation	$\mathbf{u}_i,\ \forall i \in \{1,\ldots,N_1\}$	U	$D \times N_1$
Correspondence	$M_{ia},\ i \in \{1,\ldots,N_1\},\ a \in \{1,\ldots,N_2\}$	M	$N_1 \times N_2$
Template graph	$G_{ij},\ i \in \{1,\ldots,N_1\},\ j \in \{1,\ldots,N_1\}$	G	$N_1 \times N_1$
Rotation	Euler angles or quaternions	R	$D \times D$
Regularization		μ	scalar
Robustness		γ	scalar
Scale		s	scalar

Since the relational shape matching cost function is convex w.r.t. s, this is the global solution. Please note that the use of a regularizing prior which imposes a penalty on large or small values of scale would alter this solution. Solving for the rotation matrix (in 2D) results in

$$\tan(\hat{\theta}) = \frac{\sum_{ij} G_{ij}[Z_{ij}^{(1)}W_{ij}^{(2)} - Z_{ij}^{(2)}W_{ij}^{(1)}]}{\sum_{ij} G_{ij}[Z_{ij}^{(1)}W_{ij}^{(1)} + Z_{ij}^{(2)}W_{ij}^{(2)}]} \quad (3)$$

where $\mathbf{Z}_{ij} \stackrel{\text{def}}{=} \mathbf{z}_i - \mathbf{z}_j$, $\mathbf{W}_{ij} \stackrel{\text{def}}{=} \mathbf{w}_i - \mathbf{w}_j$, $\mathbf{z}_i \stackrel{\text{def}}{=} \mathbf{x}_i + \mathbf{u}_i$ and $\mathbf{w}_i \stackrel{\text{def}}{=} (M\mathbf{y})_i = \sum_a M_{ia}\mathbf{y}_a,\ \forall i$. The superscripts in (3) indicate 2D components of the vectors $\mathbf{Z}_{ij}$, $\mathbf{W}_{ij}$, $\mathbf{z}_i$, $\mathbf{z}_j$ and $\mathbf{w}_i$, $\mathbf{w}_j$. Substituting the solutions for s and R back in (1), we obtain the following similarity transformation-invariant cost function

$$E_{\text{effective}}(M,U) \quad = \quad 2\sqrt{E_A} - 2\sqrt{E_B^2 + E_C^2} \quad (4)$$

where

$$E_A = \sum_{ijab} G_{ij}M_{ia}M_{jb}||\mathbf{Z}_{ij}||^2 \sum_{ijab} G_{ij}M_{ia}M_{jb}||\mathbf{y}_a - \mathbf{y}_b||^2,$$

$$E_B = \sum_{ij} G_{ij}[Z_{ij}^{(1)}W_{ij}^{(1)} + Z_{ij}^{(2)}W_{ij}^{(2)}],$$

$$\text{and } E_C = \sum_{ij} G_{ij}[Z_{ij}^{(1)}W_{ij}^{(2)} - Z_{ij}^{(2)}W_{ij}^{(1)}].$$

Note that the scale parameter s is not unilaterally applied to the template X but is bilaterally applied to both X and Y. (This is why we were able to eliminate both rotation and scale parameters—they were decoupled in the cost function). It is somewhat difficult to see why (4) is invariant under rotation.

With perfect hindsight, it would have been possible to design such an intrinsic shape matching cost function at the very outset. However, we prefer to use the non-intrinsic relational shape matching cost function in (1) for the simple reason that it is a much more tractable quadratic cost function w.r.t. U and as we will show, it is a much more tractable cost function w.r.t. the correspondences M as well.

3.2 A deformation prior

The relational shape matching cost function in (1) attempts to find the best *pairwise* least-squares fit between template and data point-sets. However, since the number of deformation variables ($u_i,\ \forall i \in \{1,\ldots,N_1\}$) *scales* with the cardinality of the template, there are numerous solutions. To see this, first restrict the correspondences such that there are no outliers in point-sets X and Y. Then, by inspection, we see that there is a solution for U for each of the $N_1! = N_2! = N!$ correspondences. One such solution is

$$\hat{\mathbf{u}}_i = \frac{1}{s}\sum_a M_{ia}R^T\mathbf{y}_a - \mathbf{x}_i. \quad (5)$$

This "solution" also sends the original cost to zero if it is assumed that we have the correct solution for s and R. Clearly, we need to regularize the cost function by penalizing excessively large deformations. In most of our previous work, we used a thin-plate spline (or a Gaussian radial basis function spline) to regularize the spatial mapping. The spline essentially imposed a penalty on large deformations of the plane (in 2D). The choice of prior reflects our emphasis on an object model as opposed to a spatial mapping.

Our prior is a pragmatic choice. Since the relational shape matching cost function is based on pairwise matching, the prior should reflect that. Instead of merely penalizing large deformations, we choose to penalize large relative deformations. This is done via a prior cost function of the form

$$E_{\text{prior}}(U) = \sum_{ij} G_{ij}||\mathbf{u}_i - \mathbf{u}_j||^2 = 2\,\text{trace}(U^T L U) \quad (6)$$

where $L_{ij} = n_i\delta_{ij} - G_{ij}$ is the $(i,j)^{\text{th}}$ entry of the *Laplacian* of the graph G where $n_i \stackrel{\text{def}}{=} \sum_j G_{ij}$. The

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

Laplacian of a graph is a non-negative definite matrix. The first zero eigenvalue along with the eigenvector $e = \frac{1}{N_1}[1, 1, \ldots, 1]^T$ corresponds to the translation invariance of U. If a second eigenvalue of the Laplacian is also zero, then the graph is bipartite [7]. The prior in (6) is invariant under translation and rotation of U. With this prior in place, we may write the overall RSM cost function as

$$E_{\mathrm{RSM}}(M, U, s, R) = E_{\mathrm{relational}}(M, U, s, R) + \mu E_{\mathrm{prior}}(U). \quad (7)$$

In (7), the regularization parameter $\mu > 0$ determines the tradeoff between the relational shape matching cost and the degree of deformation.

How does this prior constrain the set of possible solutions for U? Recall from (5) that the relational shape matching term allowed arbitrary deformations with no penalty. To get a better understanding of this prior in the context of the relational shape matching cost, we again switch to the no outlier scenario: $\sum_a M_{ia} = 1$ and $\sum_i M_{ia} = 1$. When this is done, the relational shape matching cost is drastically simplified and we may write (after some straightforward algebraic manipulations)

$$\begin{aligned} E_{\mathrm{equiv}}(U) &= \mathrm{trace}(sU^T LU + 2sU^T LX \\ &\quad -2U^T LMYR^T + \mu U^T LU). \end{aligned} \quad (8)$$

Here $E_{\mathrm{equiv}}(U)$ is a rewrite of $E_{\mathrm{RSM}}(M, U, s, R)$ but focusing solely on those terms that depend on U. We are doing this in order to better understand the nature of the prior *vis-a-vis* the relational shape matching cost. In (8), the first and fourth terms are convex w.r.t. U and the second and third terms are linear w.r.t. U. Consequently, we can evaluate the set of fixed points for U by differentiating (8) w.r.t. U and setting the result to zero. When this is done, we get

$$L\hat{U} = \frac{1}{s + \mu} L(MYR^T - sX). \quad (9)$$

Since L is non-negative definite, we cannot invert the above solution for U. However, we can clearly see the influence of the regularization parameter μ in the solution; larger the value of μ, smaller the resulting deformation and vice versa. The role played by the Laplacian matrix L is very interesting. The Laplacian imposes a correlation structure on U which is driven by the topology of G. This is to be expected since both the likelihood and the prior induce correlations in U via G.

4 An optimization algorithm based on the Bethe Approach

In this section, we focus on algorithms for minimizing the objective function in (7). A motivation for this entire approach is the excellent fit between the relational shape matching cost function and new and exciting work in the interface between Bayesian networks and statistical physics [30, 29, 31]. In fact, we were emboldened to design the relational shape matching cost function mainly due to a fundamental advance that has recently been made in algorithms for Bayesian networks. In this section, we describe this advance and relate it to our new cost functions.

4.1 Bayesian Belief Propagation and the Bethe free energy

In our previous work [4], we have designed efficient algorithms for point matching by relaxing the notion of binary correspondence (as a permutation matrix and outliers) to that of a doubly substochastic matrix. We have empirically shown that this relaxation allows us to efficiently solve for correspondences and non-rigid deformations. The principal drawback in this previous work is the inability to take relationships between the point features into account during matching. In this work, we have designed new cost functions which incorporate a set of relationships (G) between the template features. This representation matches up well with recent work in Bayesian networks [29].

Assume for the moment that the template graph is a tree T. Also, assume that the correspondences are now no longer binary but are "soft" in the sense of taking values in the interval $[0, 1]$. The soft correspondences will be denoted by $\{p_{ia}\}$. Also, assume that the constraints on the correspondences $\{p_{ia}\}$ are relaxed in the following way; $p_{ia} > 0$, $\sum_a p_{ia} = 1$. For the sake of exposition, assume that there are no outliers and that $N_1 = N_2 = N$. The new constraints on $\{p_{ia}\}$ have relaxed the original binary correspondence to that of a "fuzzy" correspondence. Also, the earlier two way constraints ($\sum_i M_{ia} = 1$, $\sum_a M_{ia} = 1$) have given way to a one way constraint ($\sum_a p_{ia} = 1$). With this change in representation, we can conceive of the correspondence matrix as the *marginal* probability of node "i" in a Bayesian network with each node having $a = \{1, 2, \ldots, N\}$ possibilities. Since the structure of the Bayesian network is a tree T, we can use the "vanilla" Bayesian belief propagation algorithm [14] to find the optimum solution for the correspondence matrix M now reconceived as a marginal probability $\{p_{ia}\}$ (for fixed settings of the deformation U). In addition to the marginal distribution $\{p_{ia}\}$ we also have the pairwise joint distribution $\{p_{ijab}\}$ which can be interpreted as the set of pairwise relational "soft" correspondences indicating the degree or extent to which point "i" matches to point "a" *AND* point "j" matches to point "b."

While the Bayesian belief propagation algorithm is guaranteed to converge to the correct marginal distribution—which in our case is the set of "soft" correspondences $\{p_{ia}\}$—this result is only true for graphs

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

with no loops or trees [14]. Recently, it was shown that the Bayesian belief propagation algorithm converges to the saddle points of the so-called Bethe free energy *if it converges at all* [30]. The Bethe free energy for the relational shape matching cost function is written as

$$F(\{p_{ijab}, p_{ia}\}, U, s, R) = \frac{1}{\alpha}\sum_{ijab} G_{ij} p_{ijab} \log \frac{p_{ijab}}{p_{ia}p_{jb}} + \frac{1}{\alpha}\sum_{ia} p_{ia} \log p_{ia} - \sum_{ijab} G_{ij} p_{ijab} \log \psi_{ijab}(U, s, R) \tag{10}$$

where

$$\psi_{ijab}(U, s, R) = e^{\gamma - ||\sqrt{s}R[(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)] - \frac{1}{\sqrt{s}}(\mathbf{y}_a - \mathbf{y}_b)||^2} \tag{11}$$

and has the interpretation of minimizing the mutual information on the links while maximizing the marginal entropy (MIME) [18]. The parameter α weighs the energy and entropy terms. Now that there is a cost function (the Bethe free energy) associated with a particular algorithm (belief propagation), alternative minimization strategies can be considered [31, 18]. Rather than describe the development of algorithms for the general Bethe free energy, we immediately specialize to the relational shape matching (RSM) cost function. The full Bethe free energy for the RSM cost function after taking into account the outlier rejection issue as well is

$$\begin{aligned} F(\{p_{ijab}, p_{ia}\}, U, s, R) = {} & \frac{1}{\alpha}\sum_{ijab} G_{ij} p_{ijab} \log \frac{p_{ijab}}{p_{ia}p_{jb}} \\ & + \frac{1}{\alpha}\sum_{ia} p_{ia} \log p_{ia} - \frac{\kappa}{\alpha}\sum_{ia} p_{ia} \log p_{ia} \\ & - \sum_{ijab} G_{ij} p_{ijab} \log \psi_{ijab}(U, s, R) \\ & + \sum_{ij} G_{ij} \sum_{b} \lambda_{ijb}[\sum_{a} p_{ijab} - p_{jb}] \\ & + \sum_{ij} G_{ij} \sum_{a} \lambda_{jia}[\sum_{b} p_{ijab} - p_{ia}] \\ & + \sum_{i} \eta_i^{(1)}(\sum_{a} p_{ia} + r_i^{(1)} - 1) + \sum_{a} \eta_a^{(2)}(\sum_{i} p_{ia} + r_a^{(2)} - 1) \end{aligned} \tag{12}$$

where the constraints

$$\sum_b p_{ijab} = p_{ia}, \sum_a p_{ijab} = p_{jb}, \sum_{i=1}^{N_1} p_{ia} \leq 1, \sum_{a=1}^{N_2} p_{ia} \leq 1 \tag{13}$$

have been implemented using Lagrange parameters λ and η. Note the presence of slack variables (outlier processes) $r^{(1)}$and $r^{(2)}$ in (12). The parameters α and κ determine the "softness" of the correspondences. The original Bethe free energy has κ set to zero since binary-valued probabilities are not a concern in Bayesian networks. As κ is increased, any descent algorithm will convergence to closer to integer solutions. We have considerable experience with such "self-annihilation" cost functions [16]. The following algorithm has none of the drawbacks of the Bayesian belief propagation algorithm. When we restrict ourselves to just the probabilities, it has been shown to converge to a fixed point of (12) [31, 18]. The addition of the rotation, scale and deformation variables does not change this since we are essentially presenting an alternating algorithm which alternates between estimating the rotation, scale and deformation parameters and the correspondences.

In the relational shape matching (RSM) algorithm, the slack variables $r^{(1)}$ and $r^{(2)}$ alter the constraints slightly from the original Bethe energy. The second normalization implements the constraint $\sum_i p_{ia} \leq 1$. This constraint has no analog in Bayesian networks but as shown in our previous work [8], it has a considerable effect on solution quality. The inner loop is executed until

$$\begin{aligned} e_{\text{constraint}} \stackrel{\text{def}}{=} {} & \sum_{ij} G_{ij} \sum_a (\sum_b p_{ijab} - p_{ia})^2 \\ & + \sum_{ij} G_{ij} \sum_b (\sum_a p_{ijab} - p_{jb})^2 \\ & + \sum_i (\sum_a p_{ia} + r_i^{(1)} - 1)^2 \\ & + \sum_a (\sum_i p_{ia} + r_a^{(2)} - 1)^2 < e_{\text{thr}} \end{aligned} \tag{14}$$

Since we are only guaranteed to approach a local minimum of the enhanced Bethe cost function in (12), an important concern is solution quality. The RSM algorithm is shown below.

4.2 Computational complexity considerations

Since the core of the correspondence algorithm involves updating joint probabilities $\{p_{ijab}\}$ as well as marginal probabilities $\{p_{ia}\}$, the degree of each vertex of G plays an important role in computational complexity considerations. Assume that the degree of each vertex is not fixed and can be $O(N)$. In that case, G is dense and the per iteration complexity is $O(N^4)$ which is too high. However, one of the main attractions of Bayesian networks is working with sparse graphs. If the degree at each vertex is $O(1)$ as in curves and surfaces, then the per iteration

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

- **The Relational Shape Matching (RSM) Algorithm**
- Initialize $\{p_{ijab}, p_{ia}, r_i, s_a\}$. Set $U = 0$, $s = 1$ and $\theta = 0$.
- **Begin A: Outer K Loop:**
  - $s = \sqrt{\frac{\sum_{ijab} G_{ij} M_{ia} M_{jb} ||\mathbf{y}_a - \mathbf{y}_b||^2}{\sum_{ijab} G_{ij} M_{ia} M_{jb} ||(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)|}}$
  - $\tan(\theta) = \frac{\sum_{ij} G_{ij} [(z_i^{(1)} - z_j^{(1)})(w_i^{(2)} - w_j^{(2)}) - (z_i^{(2)} - z_j^{(2)})(w_i^{(1)} - w_j^{(1)})]}{\sum_{ij} G_{ij} [(z_i^{(1)} - z_j^{(1)})(w_i^{(1)} - w_j^{(1)}) + (z_i^{(2)} - z_j^{(2)})(w_i^{(2)} - w_j^{(2)})]}$
  - Conjugate-gradient algorithm for U since cost function is quadratic in U.
  - $p_{ijab} = (\psi_{ijab}(U, s, R))^\alpha = \exp\{-\alpha(||\sqrt{s} R[(\mathbf{x}_i + \mathbf{u}_i) - (\mathbf{x}_j + \mathbf{u}_j)] - \frac{1}{\sqrt{s}}(\mathbf{y}_a - \mathbf{y}_b)||^2 - \gamma)\}$
  - $p_{ia} = p_{ia}^{n_i + \kappa}$
  - **Begin B: Inner L Loop:** Do B until $e_{\text{constraint}} < e_{\text{thr}}$.
    - Simultaneously update $\{p_{ijab}\}$ and $\{p_{ia}\}$ below.
    - $p_{ijab} \leftarrow p_{ijab} \sqrt{\frac{p_{ia}}{\sum_b p_{ijab}}}$
    - $p_{ia} \leftarrow \sqrt{p_{ia} \sum_b p_{ijab}}$
    - Simultaneously update $\{p_{ijab}\}$ and $\{p_{jb}\}$ below.
    - $p_{ijab} \leftarrow p_{ijab} \sqrt{\frac{p_{jb}}{\sum_a p_{ijab}}}$
    - $p_{jb} \leftarrow \sqrt{p_{jb} \sum_a p_{ijab}}$
    - Normalize $\{p_{ia}\}$ across rows and columns
    - $p_{ia} \leftarrow \frac{p_{ia}}{\sum_{a=1}^{N_2} p_{ia} + r_i}$, $r_i \leftarrow \frac{r_i^{(1)}}{\sum_{a=1}^{N_2} p_{ia} + r_i}$
    - $p_{ia} \leftarrow \frac{p_{ia}}{\sum_{i=1}^{N_1} p_{ia} + s_a}$, $s_a \leftarrow \frac{r_a^{(2)}}{\sum_{i=1}^{N_1} p_{ia} + s_a}$
  - **End B**
- **End A**

complexity of our relational shape matching algorithm is $O(N^3)$ which is acceptable for offline applications such as registration and atlas construction in brain mapping and is *unacceptable* for recognition. The main culprit behind the $O(N^3)$ complexity is the update of *all* p_{ijab} where $(a, b) \in \{1, \ldots, N_2\} \times \{1, \ldots, N_2\}$. This can be reduced by restricting the number of possibilities for each edge in $G(V, E)$. We briefly outline this route to a speedup of RSM. For each edge (i, j), if we obtain an $O(1)$ set of pairs (a, b) in Y, then the set of possibilities for the edge (i, j) is drastically reduced. If the set of possibilities for *each* edge $(i, j) \in E$ is $O(1)$, then the per iteration complexity of the correspondence engine in RSM is $O(N^2)$. [From a solution quality standpoint, it makes sense to begin with $O(N)$ edge values and rapidly shrink it to $O(1)$ but we have not yet implemented this feature]. Consequently, for the RSM cost function, the per iteration complexity can be reduced to as low as $O(N^2)$. Our current implementation (discussed in the next section) uses this $O(1)$ restriction.

5 Results

In this section, we apply the RSM algorithm developed in the previous section to three examples: i) a hand-drawn template and hand-drawn data; ii) points sampled from a fish silhouette and a synthetically generated deformation of the same; and iii) two subcortical structures in a T1 MRI image and a synthetically generated deformation of the same. The results are shown in Figures 1-3. We have not yet tested the outlier rejection capabilities of the RSM algorithm. The first two examples are simpler than the third and in these cases, the correspondences found by RSM are nearly exact. We merely thresholded the $\{p_{ia}\}$ matrix and displayed the results. The third example is much more difficult due to a relatively large deformation of the subcortical structures. The correspondences (displayed by thresholding) are correct for the most part with multiple correspondences seen at the bottom of the figure. Please note the small topology difference at the bottom left between the two graphs which was deliberately induced. This is

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

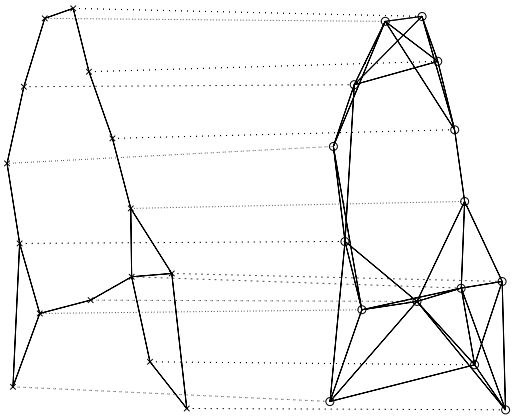

Figure 1: A hand-drawn template is shown on the left with roughly corresponding hand-drawn data shown on the right. The template graph is a symmetrized two nearest neighbor graph while the data graph is a symmetrized three nearest neighbor graph. The correspondences found by the RSM algorithm are shown as dotted lines.

meant to illustrate the relative robustness of using probabilistic correspondences (as in the Bethe free energy) rather than binary correspondences which are more typically used.

6 Discussion

To our knowledge, this is the first time that a relational shape matching framework has been constructed with *pairwise* correspondences and the deformation as the basic variables. The Bayesian network framework is *extraordinarily* well suited to handle the MAP objective functions arising from this formulation. We also stress that we were extremely careful to avoid formulating the objective function as an inexact, weighted graph matching problem. Essentially, the Bayesian network searches for the most probable set of matching deformed template relationships in the data *but not vice versa*. While our experiments are at the proof of concept stage, they serve to demonstrate a concrete application of the relational shape matching Bayesian network framework.

References

[1] Y. Amit and A. Kong. Graphical templates for model recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(4):225–236, 1996.

[2] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. Patt. Anal. Mach. Intell.*, 24(4):509–522, 2002.

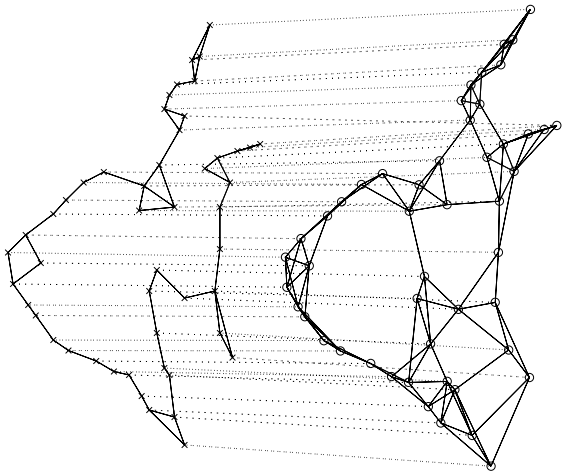

Figure 2: Points sampled from fish silhouette on left and a synthetically deformed version on the right. The template graph is a symmetrized three nearest neighbor graph whereas the data graph is a symmetrized four nearest neighbor graph.

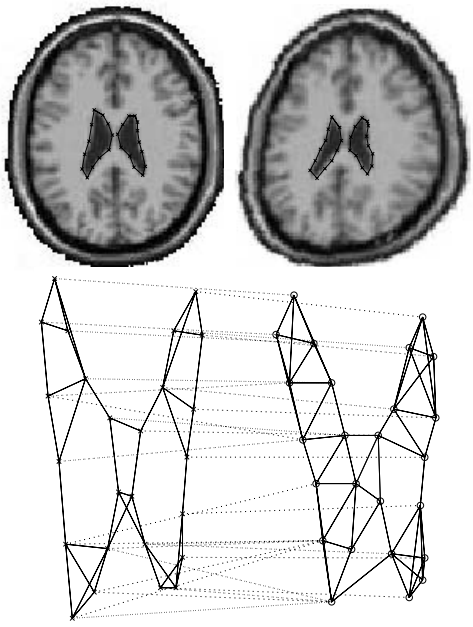

Figure 3: T1-weighted axial MR slice and a synthetically deformed version of the same are shown on top left and right. The template and data graphs were generated as in Figure 2.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 IEEE

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

[4] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 2002. (in press).

[5] 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.

[6] J. M. Coughlan and S. J. Ferreira. Finding deformable shapes using loopy belief propagation. In *European Conf. on Computer Vision (ECCV)*, Lecture Notes in Computer Science, LNCS 2351, pages III:453–468. Springer, 2002.

[7] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czech. Math. Journal*, 25:619–633, 1975.

[8] S. Gold, A. Rangarajan, C. P. Lu, S. Pappu, and E. Mjolsness. New algorithms for 2-D and 3-D point matching: pose estimation and correspondence. *Pattern Recognition*, 31(8):1019–1031, 1998.

[9] E. Grimson. *Object recognition by computer: The role of geometric constraints*. MIT Press, Cambridge, MA, 1990.

[10] M. Hagedoorn. *Pattern matching using similarity measures*. PhD thesis, Univ. Utrecht, Utrecht, Netherlands, 2000.

[11] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered scenes. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21(5):433–449, 1999.

[12] L. J. Latecki, R. Lakamper, and U. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages I:424–429. IEEE Press, 2000.

[13] 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.

[14] J. Pearl. *Probabilistic reasoning in intelligent systems*. Morgan-Kaufmann, 1988.

[15] M. Pelillo, K. Siddiqi, and S. W. Zucker. Matching hierarchical structures using association graphs. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21(11):1105–1120, 1999.

[16] A. Rangarajan. Self annealing and self annihilation: unifying deterministic annealing and relaxation labeling. *Pattern Recognition*, 33:635–649, 2000.

[17] A. Rangarajan, H. Chui, and E. Mjolsness. A relationship between spline-based deformable models and weighted graphs in non-rigid matching. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages I:897–904. IEEE Press, 2001.

[18] A. Rangarajan and A. L. Yuille. MIME: Mutual information minimization and entropy maximization for Bayesian belief propagation. In T. G. Dietterich, S. Becker, and Z. Ghahramani, editors, *Advances in Neural Information Processing Systems 14*, Cambridge, MA, 2002. MIT Press.

[19] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free form deformations: Application to breast MR images. *IEEE Trans. on Medical Imaging*, 18:712–721, 1999.

[20] S. Sclaroff and A. P. Pentland. Modal matching for correspondence and recognition. *IEEE Trans. Patt. Anal. Mach. Intell.*, 17(6):545–561, 1995.

[21] G. Scott and C. Longuet-Higgins. An algorithm for associating the features of two images. *Proc. Royal Society of London*, B244:21–26, 1991.

[22] T. B. Sebastian, P. N. Klein, and B. B. Kimia. Shock-based indexing into large shape databases. In *European Conf. on Computer Vision (ECCV)*, pages III:731–746, 2002.

[23] L. Shapiro and J. Brady. Feature-based correspondence: an eigenvector approach. *Image and Vision Computing*, 10:283–288, 1992.

[24] J. Sun, H.-Y Shum, and N.-N. Zheng. Stereo matching using belief propagation. In *European Conf. on Computer Vision (ECCV)*, Lecture Notes in Computer Science, LNCS 2351, pages II:510–524. Springer, 2002.

[25] 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.

[26] P. Thompson and A. W. Toga. A surface-based technique for warping three-dimensional images of the brain. *IEEE Trans. Med. Imag.*, 5(4):402–417, August 1996.

[27] A. Toga, J. L. Mazziotta, and R. S. J. Frackowiack. *Brain mapping: The trilogy*. Academic Press, New York, 2000.

[28] 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.

[29] Y. Weiss. Correctness of local probability propagation in graphical models with loops. *Neural Computation*, 12:1–41, 2000.

[30] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Bethe free energy, Kikuchi approximations and belief propagation algorithms. In *Advances in Neural Information Processing Systems 13*, Cambridge, MA, 2001. MIT Press.

[31] A. L. Yuille. CCCP algorithms to minimize the Bethe and Kikuchi free energies: Convergent alternatives to belief propagation. *Neural Computation*, 14:1691–1722, 2002.

Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2003) 2-Volume Set
0-7695-1950-4/03 $17.00 © 2003 **IEEE**