Automatic Mutual Nonrigid Registration of Dense Surface Models by Graphical Model based Inference

Xiao Dong and Guoyan Zheng
ARTORG Center - ISTB, University of Bern Switzerland
Email: guoyan.zheng@ieee.org

1. Introduction

Fully automatic nonrigid registration of shapes is highly desired for computer vision and medical image processing. For a nonrigid registration between two shapes \( \text{Shape}_A \) and \( \text{Shape}_B \), the ultimate goal should be to find the correspondence between subparts of the two shapes instead of to achieve a perfect match by applying a forced deformation without constraints on the objects. Such a correspondence assignment is usually defined and achieved as an optimization problem with a cost function abstracted as

\[
H(\text{Shape}_A(\mathbf{Q}_A), \text{Shape}_B(\mathbf{Q}_B), \mathbf{T}_{A \rightarrow B}, \mathbf{C}_{A \rightarrow B})
\]

where \( \text{Shape}_A(\mathbf{Q}_A) \) and \( \text{Shape}_B(\mathbf{Q}_B) \) are the shape descriptors of \( \text{Shape}_A \) and \( \text{Shape}_B \) with shape parameters \( \mathbf{Q}_A \) and \( \mathbf{Q}_B \) in their own representation spaces, \( \mathbf{T}_{A \rightarrow B} \) is a nonrigid transformation between the spatial coordinate spaces of the two shapes, and \( \mathbf{C}_{A \rightarrow B} \) is the correspondence assignment between the subparts of the two shapes.

There are two ways to define the cost function \( H(\text{Shape}_A(\mathbf{Q}_A), \text{Shape}_B(\mathbf{Q}_B), \mathbf{T}_{A \rightarrow B}, \mathbf{C}_{A \rightarrow B}) \). In the robust point matching (RPM), active shape model (ASM), active appearance model (AAM) (Cootes and Taylor, 2004), and nonrigid shape matching algorithms using the level set based contour and surface representation (Xu et al., 2001), shape signature harmonic embedding (Lee et al., 2006) and Gaussian mixture based shape models (Roy et al., 2007) (Jiang et al., 2007), the cost function is defined as

\[
H(\text{Shape}_A(\mathbf{Q}_A), \text{Shape}_B(\mathbf{Q}_B), \mathbf{T}_{A \rightarrow B}, \mathbf{C}_{A \rightarrow B}) = E_{\text{diff}}(\text{Shape}_A(\mathbf{Q}_A), \text{Shape}_B(\mathbf{Q}_B), \mathbf{T}_{A \rightarrow B}, \mathbf{C}_{A \rightarrow B}) + \gamma E_{\text{deform}}(\mathbf{T}_{A \rightarrow B})
\]

where \( E_{\text{diff}}(\text{Shape}_A(\mathbf{Q}_A), \text{Shape}_B(\mathbf{Q}_B), \mathbf{T}_{A \rightarrow B}, \mathbf{C}_{A \rightarrow B}) \) is a measurement of the shape difference between the aligned shapes with the nonrigid transformation \( \mathbf{T}_{A \rightarrow B} \) and the correspondence assignment \( \mathbf{C}_{A \rightarrow B} \), \( E_{\text{deform}}(\mathbf{T}_{A \rightarrow B}) \) defines the bending energy of the nonrigid deformation \( \mathbf{T}_{A \rightarrow B} \) and \( \gamma \) is a factor to balance the two terms. The advantage of this approach is that both the shape difference and the bending energy are explicitly expressed and we have a wide freedom to select their definitions with different underlying physical pictures. The disadvantage is that the explicit representations of \( \mathbf{T}_{A \rightarrow B} \) and \( \mathbf{C}_{A \rightarrow B} \) are iteratively defined on each other, i.e., a \( \mathbf{T}_{A \rightarrow B} \) can only be defined on a given \( \mathbf{C}_{A \rightarrow B} \) and vice versa. This usually results in that \( \mathbf{T}_{A \rightarrow B} \) and \( \mathbf{C}_{A \rightarrow B} \) can only be updated iteratively during the optimization. Such an optimization strategy can only converge to local optima and it usually asks for a proper initialization.
In the shape context (Belongie et al., 2002), the spectral decomposition (Jain and Zhang, 2006) and graphical model based registration algorithms (Coughlan and Ferreira, 2002)(Caetano et al., 2004)(Rangarajan et al., 2003), both the shape difference and the bending energy of the nonrigid deformation are defined implicitly, where both of them are encoded in the relative positions between correspondent landmarks of the shapes. In such an approach, the optimization procedure aims to directly find the optimal correspondence assignment \( C_{A \rightarrow B} \) and the nonrigid transformation \( T_{A \rightarrow B} \) never appears during the optimization. The advantage of this approach is that it can lead to a fully automatic registration since the optimization does no rely on an explicit transformation \( T_{A \rightarrow B} \) between the two shape. The disadvantage is that usually this approach is computational expensive. In (Coughlan and Ferreira, 2002)(Caetano et al., 2004)(Zhang and Seitz, 2005)(Sun et al., 2003), graphical model based Bayesian inference was used for the purpose of point set matching, image segmentation and stereo matching. In (Xiao et al., 2007), this idea was also explored for the purpose of surface matching. One limitation of the work of (Xiao et al., 2007) is that it only matches shapes with the same number of landmarks and assumed that a correspondence assignment between the landmarks of these two shapes can always be achieved. The optimization procedure to find the optimal correspondence assignment was achieved by Gibbs sampling, which may not be efficient to explore the whole configuration space (Gibbs, 2000) and therefore apt to local optima.

This paper will explore the idea of graphical model based approach for the nonrigid surface matching of two dense surfaces. The main contributions of this paper are: (a) A cost function, which is immune to rigid transformations and robust to nonrigid shape deformations, is used to measure the difference between shapes; (b) Belief propagation is used in the optimization procedure to find the optimal correspondence assignment, which can converge efficiently for our registration purpose; (c) A hierarchical matching strategy is proposed so that it can on one hand reduce the computational complexity and on the other hand, during the matching we can always assign a sparse surface to a denser surface so that the existence of correspondent landmarks can be ensured; and (d) A parallel graphical model is constructed so that the two surfaces will mutually register to each other and also exchange local assignment information during the registration, which can accelerate the convergence of the optimization procedure and overcome local optima.

2. Graphical model based nonrigid surface matching

2.1 The cost function for nonrigid surface matching

The inputs of our system are two dense triangulated surfaces \( \text{Surf}_A, \text{Surf}_B \) with \( N_A \) and \( N_B \) \((N_A \approx N_B)\) triangle faces respectively. The surfaces can be described as \( S_A = [s_A^0, s_A^1, \ldots, s_A^{N_A-1}] \) and \( S_B = [s_B^0, s_B^1, \ldots, s_B^{N_B-1}] \) respectively, where \( s_A^i = [x_A^i, n_A^i] \) \((s_B^i = [x_B^i, n_B^i])\) stands for the 3d position of the center and the normal direction of the \( i^{th} \) face of \( \text{Surf}_A(\text{Surf}_B) \).

For each landmark \( s_A^i \) on the surface \( \text{Surf}_A \), we can compute a shape signature \( d_A^i \) of the landmark at the \( i^{th} \) face of \( \text{Surf}_A \) as

\[
\begin{align*}
  d_A^i &= [d_A^{i,0}, \ldots, d_A^{i,N}, \ldots, d_A^{i,N_A-1}]_{i \neq \cdot}, \\
  & = [D_A^{ij}, A_A^{ij}, AL_A^{ij}],
\end{align*}
\]

where \( D_A^{ij} = \| x_A^i - x_A^j \| \) is the spatial distance of landmark \( s_A^i \) and \( s_A^j, A_A^{ij} \) is the angle between the normal directions \( n_A^i, n_A^j \) of the two faces and \( AL_A^{ij} \) is the angle between the normal direction \( n_A^i \) and the vector connecting \( x_A^j \) and \( x_A^i \) as shown in Fig. 1(a). This shape descriptor measures the distances and angles between landmarks and vectors, therefore a spatial
shape \( \text{Surf}_A(\text{Surf}_B) \) can be fully determined by the collection of all these shape signatures \( \{\mathbf{d}_{A}^i\}|\{\mathbf{d}_{B}^i\}\). This shape descriptor is immune to rigid transformations. In the case of a non-rigid deformation, the landmarks and normal vectors can be regarded as connected by elastic strings and the deviation of these values can then be understood as the bending energy of the deformation. Therefore this descriptor can also handle nonrigid shape deformations.

Without losing any generality, we consider the case to register \( S_A \) to \( S_B \). Given the shape descriptors \( \{\mathbf{d}_{A}^i\} \) and \( \{\mathbf{d}_{B}^i\} \), the difference between two surfaces with a given correspondence assignment between the landmarks of the surfaces \( C_{A\rightarrow B}(i) = j, i = 0,1,\ldots,N_A, j = 0,1,\ldots,N_B \) can then be written as

\[
H(S_A, S_B, C_{A\rightarrow B}) = \sum_i (\mathbf{d}_{A}^i - \mathbf{d}_{B}^i)\Lambda(\mathbf{d}_{A}^i - \mathbf{d}_{B}^i)^T
\]

where \( \mathbf{d}_{B}^i = [\mathbf{d}_{B}^{C_{A\rightarrow B}(0)}(i), \mathbf{d}_{B}^{C_{A\rightarrow B}(1)}(i), \ldots, \mathbf{d}_{B}^{C_{A\rightarrow B}(N_B-1)}(i)] \), \( \Lambda = \text{Diag}(\sigma_{D}^{2}, \sigma_{A}^{2}, \sigma_{N}^{2})^{-1} \) is the covariance matrix to balance the strengths of the elastic strings that connect the landmarks. The cost function measures the difference between the shape descriptor of \( S_A \) and the shape descriptor of \( S_B \) at the correspondent landmarks.

### 2.2 Graphical model based correspondence assignment optimization

For the purpose of defining a 3-dimensional surface, the cost function in (3) is over-determined since its underlying graph \( G \) is a complete graph. In (Caetano et al., 2004), the authors indicated that to determine a rigid shape in \( \mathbb{R}^d \), the minimal underlying graph \( G \) is \( k \)-tree in \( \mathbb{R}^d \). A denser graph topology will improve the robustness of the representation to keep the shape with the expense of a higher computational cost. Define a graph \( G = (V, E) \) with \( N_A \) vertices, \( V = \{v_i\} \) with each vertex correspondent to one landmark in \( S_A \) and \( E = \{e_{ij}\} \) with \( e_{ij} = 1 \) if there is an edge connecting \( v_i \) and \( v_j \). Then a generalized representation of (3) can be given as

\[
H_G(S_A, S_B, C_{A\rightarrow B}) = \sum_{i,j|e_{ij}} E^i_{A}(C_{A\rightarrow B}) = \sum_{i,j|e_{ij}} (\mathbf{d}_{A}^i - \mathbf{d}_{B}^i)\Lambda^i_{A}(\mathbf{d}_{A}^i - \mathbf{d}_{B}^i)^T
\]

where \( \mathbf{d}_{B}^i = \mathbf{d}_{B}^{C_{A\rightarrow B}(j)} \), \( \Lambda^i_{A} = \text{Diag}(\sigma_{D}^{2}, \sigma_{A}^{2}, \sigma_{N}^{2})^{-1} \).

Accordingly a distribution \( p(H_G) = \prod_{i,j|e_{ij}} \exp(-E^i_{A}(C_{A\rightarrow B})) \) can be defined and represented by a graphical model as shown in Fig. 1(b). Each vertex \( v_A^j \) represents a landmark \( s_A^j \) of \( S_A \) and its state is a vector \( sv_A^j = [p_A^{j0}, p_A^{j1}, \ldots, p_A^{jN_A-1}] \) representing the probabilities that \( s_A^j \) is assigned as the correspondent landmark of the landmarks of \( S_B \). Vertices \( v_B^j \) stands for the the landmarks in \( S_B \) and its state is a vector \( sv_B^j = [p_B^{j0}, p_B^{j1}, \ldots, p_B^{jN_A-1}] \) representing that \( s_B^j \) should be the correspondence of each landmark in \( S_A \). The vertices \( v_A^j \) are connected according to the topology of \( G \). The potentials between \( v_A^j \) encode probabilities as \( \psi_A(i,j) = \exp(-\beta E^i_{A}(C_{A\rightarrow B})) \), where \( \beta = 1/\tau \) is the inverse of the annealing temperature \( \tau \). The optimization to find the optimal correspondence assignment can then be formalized as a ML estimation to find the peak of \( p(H_G) \), which can be achieved by graphical model based Bayesian inference (Coughlan and Ferreira, 2002)(Caetano et al., 2004)(Zhang and Seitz, 2005)(Sun et al., 2003). In our method, the Bayesian inference is achieved by running the belief propagation (Coughlan and Ferreira, 2002) on the graphical model as given in Fig. 1(b). When
the belief propagation converges, the state $s^{t,i}_{v^i_A}$ is then the marginal distribution of the assignment of vertex $v^i_A$. If the annealing template $\tau$ is low enough, then $\{s^{t,i}_{v^i_A}\}$ will give a unique peak which indicates the global optimum correspondence assignment for the matching.

3. Graphical model based hierarchical mutual surface registration

It can be observed that the above described graphical model based nonrigid surface matching faces three difficulties

- The existence of correspondence assignment: In our algorithm description, we assume that the correspondent landmarks for each landmark of $Surf_A$ can always be (uniquely) found on $Surf_B$. For a general case where $N_A \approx N_B$, this can not always be satisfied.
- The computational complexity: To matching two surfaces with $N_A$ and $N_B$ faces, the computational complexity of each belief propagation iteration is $O(\kappa N_A N^2_B)$, where $\kappa$ is a factor related with the topology of the graphical model. The proposed algorithm is quite computational expensive for practical cases.
- The unbalanced information processing structure: In the graphical model based surface matching, the two surfaces play different roles. The graphical model is built on $Surf_A$ and the shape descriptors of the two surfaces are explored in different ways. Such an unbalanced information processing structure may lead to local optima as in the classic ICP based registration, where the correspondence is achieved by finding closest points from one data set to another. In an ideal surface matching solution, the two surfaces should be dealt with equivalently.

To overcome the above mentioned difficulties, we proposed a hierarchical mutual registration strategy, in which

- A hierarchical shape descriptor is used to describe each surface. In such a hierarchical structure, we can always match a sparse surface to a dense surface so that the existence of correspondence can be ensured.
- The computational cost on the hierarchical shape descriptor can be greatly reduced due to the fact that the matching results at coarse levels can be used as initializations for the finer levels so that the possible configurations of each landmark at finer levels can be greatly reduced.

![Diagram](image)

(a) The definition of the shape signature at each face of the surface

(b) The graphical models for belief propagation based nonrigid surface matching

Fig. 1. The graphical models for the registration of two surfaces taking $Surf_A$ as the template
3.1 Hierarchical shape descriptors of triangulated surface meshes
To reduce the computational cost, a hierarchical coarse-to-fine strategy is used to represent each surface in $L$ levels as $\{S_{A,L-1}, \ldots, S_{A,1}, S_{A,0}\}$ and $\{S_{B,L-1}, \ldots, S_{B,1}, S_{B,0}\}$. In our system a simple k-mean clustering method is used to accomplish this task. $Surf_A$ and $Surf_B$ are first segmented into $N_{A,L-1}$ and $N_{B,L-1}$ ($N_{A,L-1} \approx N_{B,L-1}$) clusters. Each cluster of $S_{A,L-1}$ ($S_{B,L-1}$) is represented as $s_{i,A,L-1} = [x_{i,A,L-1}, n_{i,A,L-1}]$ ($s_{i,B,L-1} = [x_{i,B,L-1}, n_{i,B,L-1}]$), where $x_{i,A,L-1}$ ($x_{i,B,L-1}$) is the center of the central triangle of this cluster and $n_{i,A,L-1}$ ($n_{i,B,L-1}$) is the mean normal direction of the faces in this cluster. Lower level representations are then obtained by further subdivide each cluster in higher level representations by k-mean clustering and each subdivided cluster can still be described by its center and normal direction. The finest level descriptors $S_{A,0}$ ($S_{B,0}$) are then $S_A$ ($S_B$). To facilitate the correspondence assignment on each landmark $s_{i,A,0}$ ($s_{i,B,0}$) at the finest level, we build an extra representation level $S_{A,-1} = \{s_{i,A,-1}, i = 0,1,\ldots, KN_{A,0} - 1\}$ ($S_{B,-1} = \{s_{i,B,-1}, i = 0,1,\ldots, KN_{B,0} - 1\}$) by uniformly sampling $K$ samples from each face in $S_{A,0}$ ($S_{B,0}$).

With such a hierarchical structure of each surface, we can construct hierarchical shape descriptors at each representation level as described in the previous part. The graphical model based surface matching can then be carried out to match a surface at a coarse level $l$ to another surface at a finer level $l-1$. Due to the fact that $N_A \approx N_B$ and $N_{A,l} << N_{B,l-1}$ ($N_{B,l} << N_{A,l-1}$), the existence of correspondence can then be ensured.

3.2 Graphical model based parallel mutual registration
In the graphical model based strategy to match $S_A$ with $S_B$, only the shape information of $S_A$ encoded in $\{\psi_A(i,j)\}$ is explicitly explored. This may lead to a slow convergence and it is apt to local optima.

The idea of parallel mutual registration is stimulated from the Turbo code in communication field, which indicates that an error-correction code which satisfies two sets of constraints simultaneously can greatly improve the its performance (McEliece et al., 1998). For our purpose of nonrigid shape matching, simultaneously exploring the different shape information of the
two surfaces, although they are not independent since they are similar, may also improve the robustness and accelerate the convergence of our algorithm.

The graphical model for parallel mutual registration of $Surf_A$ and $Surf_B$ between representation level $l - 1$ and $l$ is a duplication of the graphical model of Fig. 1(b) as shown in Fig. 2 with extra directional edges to connect vertices representing landmarks of the same surface but at different representation levels. These edges only exist if a patch at a finer level is a subset of a patch at a coarser level.

The equivalent cost function of the parallel mutual registration scheme can be written as $HG(SA,l,SB,l-1,CA,l\rightarrow l-1) + HG(SB,l,SA,l-1,CB,l\rightarrow l-1)$ with a consistency requirement on $CA,l\rightarrow l-1$ and $CB,l\rightarrow l-1$ encoded in the message exchange between the two subparts of Fig. 2. In the graphical model for the mutual registration, the correspondence assignment information is transferred in the graphical model as shown in Fig. 3.

### 3.3 Hierarchical mutual registration

The parallel registration can be carried out at each successive level of the hierarchical representation of the surfaces, where the result of belief propagation at coarse levels can be used as the initialization at finer levels. The graphical model based hierarchical mutual surface matching is summarized as in Algorithm 1.

1. $l = L - 1$, initialize all the messages as equal distributions at all configurations.
2. Run the parallel mutual matching between level $l$ and level $l - 1$ with $\beta = \beta_l$ till it converges.
3. Using the correspondence assignment result between level $l$ and $l - 1$, initialize the correspondence assignment between level $l - 1$ and $l - 2$. Increase $\beta_l$ to $\beta_{l-1}$, go to 2 till $l = 0$.

Algorithm 1. Graphical model based hierarchical automatic surface matching
Fig. 4. Randomly generated surfaces from a PCA based statistical model of proximal femurs and their hierarchical representations

Fig. 5. The results of mutual registration at $l = 3$ and $l = 2$, the red lines show the correspondent landmarks on the two surfaces and the green lines show the distance between the assignment results of our algorithm and the correspondence assignment encoded in the PCA model

4. Experiments

We verified the principle of the proposed matching algorithm on a PCA based statistical model of proximal femurs including 4098 vertices and 8192 faces. Two randomly generated instances are obtained from the statistical model and one of them is further scaled by a uniform scaling of 0.9 as shown in Fig. 4(a) and 4(b). Since our matching algorithm is immune to rigid transformation, applying any rigid transformation to these surfaces will not influence the matching result. Hierarchical representations of the surfaces are constructed with $L = 3$, the first two levels ($L = 3$ and $L = 2$) of clustering results of $\text{Surf}_A$ are shown in Fig. 4(c) and 4(d).
The correspondence assignment using the graphical model based mutual registration is then carried out on them. The underlying graph topology $G$ is selected as a random graph, in which the degree of each node is at least 5 (connected with 5 nearest neighbours) and the average degree of one node is 10. $\Lambda' = \text{Diag}(\sigma^2_D, \sigma^2_A, \sigma^2_{AL})^{-1}$ in (4) is selected as $\sigma_D = d_{Ai}/4$, $\sigma_A = \sigma_{AL} = 20^\circ$.

- The results of the mutual registration at the coarsest level $l = 3$ and $l = 2$ are shown in Fig. 5. Obviously the two surfaces are already roughly aligned at the coarsest level and this shows that the proposed algorithm does not ask for any initialization.

- The registration result at coarser level is then used to initialize the correspondence assignment at finer level as shown in Fig. 6. It can be observed that the number of candidates of each landmark at a finer level is greatly reduced due to the initialization.

- The advantage of the parallel mutual registration between the two surfaces can be observed from Fig. 7. The mutual registration can explore the shape information of two surfaces and exchange correspondence assignment information between them. This
strong constraint forces the correspondent landmarks on two surfaces to mutually assign to each other quickly. It’s observed during the experiment the belief propagation at each level can converge in less than 10 iterations.

- As mentioned before, the goal of a nonrigid registration is to find optimal correspondence assignment between shapes. Since both surfaces are generated from the same PCA model, each landmark on a surface carries a point index in the PCA model. We take the landmarks with the same point index on the two surfaces as the ground truth of the correspondence assignment and then compute the registration error as the distance between the correspondent landmark obtained by our registration algorithm and the ground truth position from prior knowledge of the PCA model. The registration error is evaluated on both shapes as $2.7 \pm 2.3$ mm. Of course the prior correspondence knowledge may not be the ground truth but it can be regarded as a proper reference.

5. Conclusions

In this paper we proposed a fully automatic scheme for nonrigid surface matching. The nonrigid surface matching is formalized as a graphical model based Bayesian inference and the belief propagation is used to achieve the optimization to find the optimal correspondence assignment between shapes. To further reduce the computational cost and enhance the robustness to noise and local optima, a hierarchical mutual registration strategy is implemented so that the shape information of the two surfaces can be simultaneously explored. Experiments on randomly generated surfaces from a PCA based statistical model showed the capability of the proposed algorithm to achieve an automatic nonrigid surface registration.

The proposed scheme can also be extended to incorporate other shape descriptors such as the Gaussian curvature as used in (Xiao et al., 2007) and the shape context (Belongie et al., 2002) since they can be easily modeled as local beliefs of each vertex in our graphical model based scheme.

One limitation of the proposed algorithm lies in the way that it handles the nonrigid deformation. Different from the commonly used TPS based deformation energy to set constraints on the shape deformation, our algorithm encodes a nonrigid deformation by the deformation of the subparts of a shape such as the distances and angles between landmarks. It’s difficult to design a metric, which can accurately evaluate the deformation energy. Future work will be carried out to design better cost functions, which can measure the deformation energy more accurately by combining more shape information including local information such as curvature and deformation energies at different representation levels.

6. References


Xu, C., Yezzi, A., Prince, J.: A summary of geometric level-set analogues for a general class of parametric active contour and surface models (2001)


