

Point-Based Non-Rigid Surface Registration with Accuracy Estimation

Hidekata Hontani and Wataru Watanabe
Nagoya Institute of Technology
Gokiso, Showa-ku, Nagoya, Aichi, 466-8555 JAPAN

Abstract

This article presents a new method for non-rigid surface registration between a surface model and a surface of an internal organ in a given 3D medical image. The surface is represented with a set of feature points, of which locations are represented by a graphical model. For constructing the representation, a set of corresponding points is distributed on each of training surfaces based on an entropy-based particle system. From these corresponding points, we estimate probability densities of the location of each feature point, the conditional probability distribution of the local image pattern around each feature point, and the probability distributions of relative positions between two neighboring feature points. When a new image is given, these densities are used for estimating the location of each feature point by means of a non-parametric belief propagation. The proposed method can estimate not only the locations of the feature points but also their conditional marginal distributions in a given image. Some experimental results obtained from real X-CT images are presented to show its performance.

1. Introduction

It is an important step for automated image analysis to register a surface model of a target organ with a given 3D medical image[1][9]. Registering the surface, we can segment the organ and measure its volume, shape, and other features.

Accuracy of the registration changes depending on the image contrasts along the surface and on its shape. For example, a high-contrast flat part of a surface will be registered accurately, and a low-contrast complicated part, on the other hand, will be registered more inaccurate. Estimation of the accuracy crucially increases availability of the resultant surface obtained by the registration[14]. The contribution of this article is to propose a new point-based surface registration method that can estimate the accuracy of the registration for each location on the surface.

Surface registration can be roughly partitioned into three

stages[1]: specification of transformation, elaboration of surface representation and similarity criterion, and matching and optimization. The transformation can be divided into two classes: rigid one[2] and non-rigid one[3][4]. For registering organ surfaces, we must employ non-rigid transformation. Here, in order to obtain appropriate results, we need a transformation model that is specific to a target surface of organ[9]. The proposed method will use a statistical model of the transformation that will be constructed from a set of training samples.

There are several methods for surface representation: for example, implicit ones[6], parametric ones[3], and primitive-based ones[8][11]. Our method is primitive-based and represents a surface by using a set of points, which is distributed on it by using the entropy-based particle system[5]. We register the surface by estimating the location of each of the points, and estimate its accuracy by computing the probability density distribution of the estimated location.

Many non-rigid registration methods register a surface by minimizing a cost function $J = J(J_d, J_t)$, where J_d is a cost denoting the distance between a surface model and surface candidate points detected from a given image, and J_t is a cost of the transformation. For example, the active contour method[3] firstly detects edges as the surface candidate points and then registers a contour model with the edges by minimizing $J = J_d + \alpha J_t$, where J_d is defined as a sum of squared normal distances between the model and the edges and J_t is the Mahalanobis distance between the current model shape and the average one. Here, it should be noted that the edge detector is designed independently of the contour representation and that it is not easy to estimate the accuracy of the location of each detected edge point. In addition, it is difficult to theoretically determine the value of the positive coefficient α . Both the accuracy of the edge detector and the value of α are required to be properly determined for the estimation of the registration.

In many registration methods, the three stages are designed separately and this causes the difficulty of the estimation of the registration accuracy. In this article, a statistical framework that unifies those three stages is proposed:

A statistical model of the transformation and a detector of each of the points that represent a target surface are both constructed simultaneously from a set of training samples, and the non-parametric belief propagation[13] is used for the registration and for the estimation of its accuracy.

2. Construction of Surface Model

2.1. Outline

Given a set of training samples of X-CT images, we construct a surface model for some target internal organ. Before constructing the model, we normalize the shape, size, and location of bodies in the images. This can be done by automatically extracting skeletons[10] or other organs[7] as landmarks. There exist some organs that can be detected more robustly than the target internal organs. This normalization should be identical to a method that will be used for registering the constructed surface model to a new medical image. The error of this normalization will diffuse the probability distributions of surface models.

Let $\mathcal{I} = \{I^i | i = 1, 2, \dots, M\}$ denote the set of the normalized images. From each image I^i , the surface is extracted manually by an expert. Let S^i denote the surface extracted from I^i . Here, we assume that S^i is smooth and closed. In addition, we assume that, after the image normalization, the shape, the size, and the location of each surface S^i are similar to each other enough for determining the distributions of corresponding points.

We construct a model surface that consists of N feature points $\{P_j | j = 1, 2, \dots, N\}$ as follows. We firstly distribute a set of N points $\mathcal{P}^i = \{P_j^i | j = 1, 2, \dots, N\}$ on each surface S^i . We obtain M sets of points $\{\mathcal{P}^i | i = 1, 2, \dots, M\}$. Let $\mathcal{P}_j \equiv \{P_j^i | i = 1, 2, \dots, M\}$ denote a set of M corresponding points distributed on M surfaces. Let x_j^i denote the 3D coordinates of P_j^i and let I_j^i denote a local image around P_j^i in I^i .

We consider the M points of \mathcal{P}_j are drawn from $p(x_j)$, where $p(x_j)$ is a prior distribution of a feature point P_j . We estimate $p(x_j)$ from \mathcal{P}_j . In addition, we estimate the conditional probability density $p(I_j | x_j)$ for each P_j based on a set $\{I_j^i | i = 1, 2, \dots, M\}$. This density will be used for computing a likelihood distribution in a newly given image to estimate the location of P_j .

In addition, we construct a graphical model G with N nodes $\{v_j\}$, each of which corresponds to the location of P_j . Two node v_j and v_k are linked by an edge e_{jk} if P_j^i and P_k^i are close enough in average. If v_j and v_k are linked by an edge, then we estimate the probability density $p(x_k - x_j)$. This density represents the variety of local deformations of the surface in a probabilistic framework.

2.2. Distribution of Corresponding Points

For arranging the corresponding points, we employ an entropy-based method proposed by Cates[5]. In this subsection, we briefly explain the entropy-based method.

This method distributes a set of corresponding points on given surfaces based on their geometry: No image pattern information is referred to for this distribution. The method quantifies two geometric features and obtains a distribution that satisfies those two features: the *uniformity* of the point distribution on each surface and the *compactness* of the inter-surface distribution of the corresponding points. The former feature is required for the representation in order to avoid sparsely distributed regions. The latter feature will make the prior distribution of each feature point P_j compact.

For the quantification of the uniformity, the points on each surface S^i are considered random variable X_j^i drawn from a probability density function (PDF) $p^i(X_j^i)$. For simplicity, let $p^i(x) = p^i(X_j^i = x)$. Then, the differential entropy $H[X^i]$ of the PDF $p^i(x)$ is defined as

$$H^i[X^i] = - \int_{X^i} p(x) \log p(x) dx. \quad (1)$$

Applying Parzen window method, we approximate the PDF $p^i(x_j^i)$ on S^i as follows:

$$p^i(x_j^i) \simeq \frac{1}{N(N-1)} \sum_{k=1, k \neq j} \mathcal{N}(x_j^i; x_k^i, \sigma^2), \quad (2)$$

where $\mathcal{N}(x; \mu, \sigma^2)$ is a 3D isotropic normal distribution with mean μ and variance σ^2 . From (1) and (2), we obtain

$$H^i[X^i] \simeq \sum_j \frac{1}{N(N-1)} \sum_{k \neq j} \mathcal{N}(x_j; x_k, \sigma^2). \quad (3)$$

The distribution of points will be computed by means of the steepest descent method. The gradient of H^i is as follows:

$$\frac{\partial H^i}{\partial x_j^i} = \frac{1}{\sigma^2} \sum_{k=1, k \neq j}^N (x_j - x_k) w_{jk}, \quad (4)$$

where $\sum_k w_{jk} = 1$. This means that a repulsive force is generated between two points on S^i when H^i is maximized.

For the quantification of the compactness, the a set of N points on each surface is considered random variable Z drawn from a normal distribution. Let z denote a realization of Z and let $z^i = [x_1^{i\top} x_2^{i\top} \dots x_N^{i\top}]^\top$ denote a $3N$ -vector that represents the set of N points $\{P_j^i\}$ ($j = 1, 2, \dots, N$) distributed on S^i . Assuming each set z^i is drawn from a normal distribution, we can quantify the compactness or similarity of the distribution among the surfaces using the

covariance Σ of the normal distribution. The entropy is given by

$$H[Z] \simeq \frac{1}{2} \log |\Sigma| = \frac{1}{2} \sum_{l=1}^{3N} \log \lambda_l, \quad (5)$$

where λ_l is the eigenvector of Σ . Let $Y = [z^1 - \bar{z}z^2 - \bar{z} \cdots z^M - \bar{z}]$ denote a $3N \times M$ matrix of deviations, where \bar{z} is the average of $\{z^i\}$. Then we can estimate the entropy as follows:

$$H[Z] \simeq \frac{1}{2} \left| \frac{1}{M-1} Y^\top Y \right|. \quad (6)$$

Denoting an identity matrix I , we obtain the gradient of $H[Z]$ as

$$\frac{\partial H}{\partial z} = -Y(Y^\top Y + \alpha I)^{-1}, \quad (7)$$

where α is a regularization on the inverse of $Y^\top Y$. The negative of the gradient (7) updates the entire system to improve the compactness.

The entropy-based method minimizes the following cost function Q to distribute corresponding points on all given surfaces.

$$Q(x_j^i) = H[Z] - \sum_{i=1}^M H^i[X^i]. \quad (8)$$

We minimize Q by means of the steepest descent method as

$$x_j^i \leftarrow x_j^i - \gamma \left(\frac{\partial H}{\partial x_j^i} - \frac{\partial H^i}{\partial x_j^i} \right), \quad (9)$$

where γ is a time step. As shown in (4), in general, the gradient $g_j^i = \partial Q / \partial x_j^i$ is not included in the tangent space of S^i . As the result, $x_j^i - \gamma g_j^i$ becomes out of the surface S^i . To compensate this, we project the point $x_j^i - \gamma g_j^i$ to the nearest point on S^i . For computing this projection effectively, the surface S^i is represented by a level set of its distance function.

2.3. Estimation of Probabilistic Distribution

We construct a graph with N nodes $\{v_j | j = 1, 2, \dots, N\}$, each of which corresponds to P_j . Let d_{jk}^i denote the distance between P_j^i and P_k^i . For computing efficiency, we employ the Euclidean distance for d_{jk}^i . Let \bar{d}_{jk}^i denote the average of $\{d_{jk}^i\}$ for $i = 1, 2, \dots, M$. Connecting v_j and v_k by an edge e_{jk} if $\bar{d}_{jk}^i < d_0$, we obtain a graph G that represents neighbors for each feature point. Let \mathcal{E} denote a set of edges d_{jk} that satisfy $\bar{d}_{ij} < d_0$. Here, the threshold d_0 is experimentally determined. Based on M graphs $\{G^i\}$ and M images I^i , we estimate following three probabilistic distributions.

1. The prior probability distribution $p(x_j)$ of the location x_j of i -th feature point P_j .
2. The conditional probability distribution $p(I_j | x_j)$ of the local image pattern I_j under the condition that the location x_j is given.
3. The probability distribution $p(x_j - x_k)$ of the relative position between P_j and P_k .

We assume that $p(x_j)$ is a normal distribution and $\{P_j^1, P_j^2, \dots, P_j^M\}$ are randomly drawn from $p(x_j)$. We estimate the mean and the variance of the normal distribution based on $\{x_j^1, x_j^2, \dots, x_j^M\}$. The requirement of the compactness described in previous subsection decreases this variance.

The probability $p(I_j | x_j)$ is used for detecting each feature point P_j from a given image. Here, I_j denotes a local small image whose size is $L \times L \times L$ and its center is at x_j . We assume that $p(I_j | x_j)$ obeys a normal distribution. We estimate its mean \bar{I}_j and covariance Σ_j based on $\{P_j^i\}$ ($i = 1, 2, \dots, M$).

$$p(I_j | x_j) = \mathcal{N}(I_j; \bar{I}_j, \Sigma_j), \quad (10)$$

where $\mathcal{N}(I; \mu, \Sigma)$ denotes a L^3 dimensional normal distribution with the average μ and the covariance Σ .

We estimate $p(x_j - x_k)$ only when the feature points P_j and P_k are connected with an edge in the graph. We assume that $p(x_j - x_k)$ obeys a normal distribution whose mean is \bar{x}_{kj} and its covariance is Σ_{kj} .

$$p(x_j - x_k) = \mathcal{N}(x_j; \bar{x}_{kj}, \Sigma_{kj}) \quad (11)$$

The mean and the covariance are estimated based on M samples $x_j^i - x_k^i$ ($i = 1, 2, \dots, M$) where $e_{jk} \in \mathcal{E}$. This probability density $p(x_j - x_k)$ denotes the variety of the local shape change around P_k .

3. Registration of Surface Model

Given a new normalized medical image I , we register the surface model to the image by estimating the conditional marginal distribution $p(x_i | I)$ by means of a non-parametric belief propagation (NBP). Based on the surface model, we can quantify a joint probability distribution $p(\{x_i\}, \{I_i\})$ in terms of potential functions as follows:

$$p(\{x_i\}, \{I_i\}) \propto \prod_i \psi_i(x_i, I_i) \prod_{e_{ij} \in \mathcal{E}} \psi_{i,j}(x_i, x_j), \quad (12)$$

where $\psi_i(\cdot, \cdot)$ and $\psi_{i,j}(\cdot, \cdot)$ are single-node potentials and pairwise ones, respectively. Here, these potentials are defined as follows:

$$\psi_i(x_i, I_i) = p(x_i)p(I_i | x_i), \quad \psi_{i,j}(x_i, x_j) = p(x_i - x_j). \quad (13)$$

The NBP algorithm iteratively updates the inference of the marginal distribution $p(x_i|I)$ at each node v_i . Let $\hat{p}^n(x_i|I)$ denote the inference at iteration n . Each node v_i computes \hat{p}^n by multiplying its single-node potential with messages m_{ki}^n sent from its neighbor nodes v_k as follows:

$$\hat{p}^n(x_i|I) \propto \psi_i(x_i, I_i) \prod_{e_{ki} \in \mathcal{E}} m_{ki}^n(x_i), \quad (14)$$

where

$$m_{ki}^n(x_i) \propto \int_{x_k} \psi_{i,k}(x_i, x_k) \frac{\hat{p}^{n-1}(x_k)}{m_{ik}^{n-1}(x_k)} dx_k. \quad (15)$$

Given a new image I , we can compute the likelihood $p(I_i|x_i)$ at each voxel based on (10). The distributions of the likelihood is in general hard to represent parametrically. This is the reason why we employ non-parametric method for the estimation. As proposed by Sudderth[13], we represent the messages $m_{ki}(x_i)$ and the single-node potentials $\psi_i(x_i)$ by Gaussian mixture models. For example, the messages $m_{ki}(x_i)$ are defined by W components of Gaussian distributions as follows:

$$m_{ki} = \sum_{\alpha=1}^W w_{ki}^{(\alpha)} \mathcal{N}(x_i; \mu_{ki}^{(\alpha)}, \Sigma_i), \quad (16)$$

where α denotes the number of the Gaussian component, $w_{ki}^{(\alpha)}$ denotes its weight, and $\mu_{ki}^{(\alpha)}$ denotes its mean. Each of these Gaussian components is represented by a particle defined by $\{w_{ki}^{(\alpha)}, \mu_{ki}^{(\alpha)}, \Sigma_i\}$.

When a node v_i receives a set of messages m_{ki}^n from its neighbor nodes v_k ($e_{ki} \in \mathcal{E}$), the node v_i computes the right hand side of (14) and estimates $\hat{p}^n(x_i|I)$. For the computation of the product of several Gaussian mixtures, a Gibbs sampler is employed.

Let $\{w_i^{(\alpha)}, \mu_i^{(\alpha)}, \Sigma_i\}$ denote particles ($\alpha = 1, 2, \dots, W$) that represent $\hat{p}^n(x_i|I)$. In other words, assume that $\hat{p}^n(x_i|I)$ is represented as follows:

$$\hat{p}^n(x_i|I) = \sum_{\alpha=1}^W w_i^{(\alpha)} \mathcal{N}(x_i; \mu_i^{(\alpha)}, \Sigma_i). \quad (17)$$

Then, we can compute messages m_{ij} based on the particles $\{w_i^{(\alpha)}, \mu_i^{(\alpha)}, \Sigma_i\}$ and $p(x_j - x_i)$ in (11). We first move each particle as

$$\mu_{ij}^{(\alpha)} = \mu_i^{(\alpha)} + \nu^{(\alpha)}, \quad (18)$$

where $\nu^{(\alpha)}$ is a random noise that obeys the normal distribution $\mathcal{N}(\cdot; \bar{x}_{ij}, \Sigma_{ij})$ in (11). Each of the particles is then weighted by the reminder (15): $w_{ij}^{(\alpha)} = w_i^{(\alpha)} / m_{ji}(\mu_i^{(\alpha)})$.

The registration algorithm is summarized as follows.

Feature Point Detection The likelihood distribution for each feature point P_i is computed based on (10).

1. Normalize a given image based on the shape of the body and obtain I .
2. Compute the likelihood distribution $p(I_i|x_i)$ for each feature point P_i using (10).
3. Compute the joint probability distribution $p(x_i, I_i)$ for each P_i by computing $p(x_i)p(I_i|x_i)$.

Initialization for NBP Initial values for computing NBP are set.

1. Draw W particles from $\hat{p}_i^0 = p(x_i, I_i)$ to determine $\mu_i^{(\alpha)}$.
2. Initialize $w_i^{(\alpha)} \equiv 1/W$ and $m_{ij} \equiv 1$ for all i and j . Set Σ_i for particles by means of the rule of thumb estimate[12].

Non-parametric belief propagation The marginal distributions $\hat{p}(x_i|I)$ is repeatedly updated.

1. Compute $m_{ij}^{n+1}(x_j)$ from m_{ji}^n and $\hat{p}^n(x_i)$ using (15).
2. Compute $\hat{p}^{n+1}(x_i|I)$ using (14).
3. Iterate until converged.

4. Experimental Results

We applied the proposed method for registering surfaces of aortic arches in given whole-body non-contrasted X-CT images. Automatic registration of aortic arches is required for, e.g., automatically locating para-aortic lymph nodes in given X-CT images. When given images are not contrasted, it is not easy to register aortic arches: if non-contrasted, many false contours are detected near the true ones, and some true contours are missing. Model registration is needed to detect aortic walls accurately. We used a set of $M = 15$ non-contrasted X-CT images, of which resolution is $0.98\text{mm} \times 0.98\text{mm} \times 4.25\text{mm}$.

4.1. Normalization of Shape of Body

For constructing a surface model of aortic arch, we automatically normalized the shape of bodies in the images. For this normalization, we firstly detected all ribs in images and computed their convex hulls. In addition, we detected bronchial bifurcations as landmarks because the aorta runs around the bronchial bifurcation. The spline transformation was used to normalize the shape, size and location of the convex hull and to move the bronchial bifurcation to a pre-determined position. Some examples of this normalization are shown in Fig.1. As shown in Fig.1, the shapes of the chests were normalized.

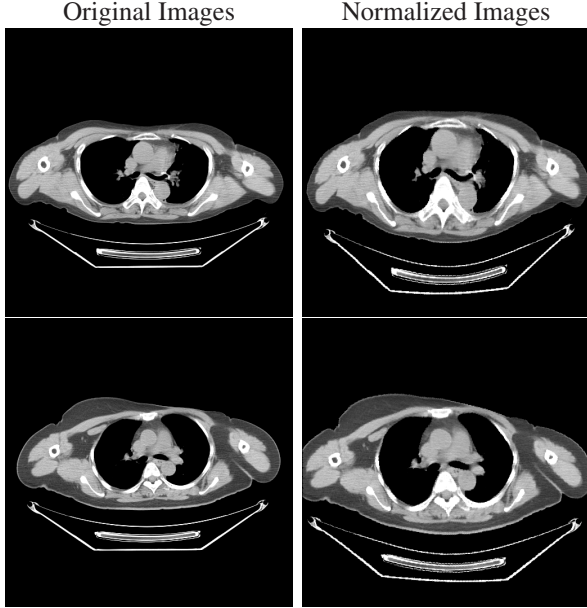


Figure 1. Examples of the Normalization. Left: Given images. Right: Normalized images.

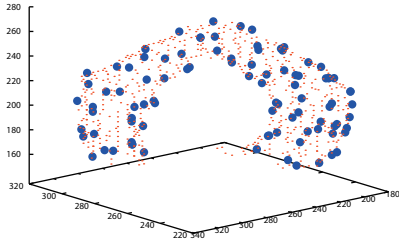


Figure 2. An example of the feature points distribution

4.2. Distribution of Corresponding Points

From the set of these normalized images, an expert manually extracted the surfaces of aortic walls and obtained $\{S_i | i = 1, 2, \dots, 15\}$. Given $M = 15$ surfaces $\{S^i\}$ ($i = 1, 2, \dots, M$), we distributed $N = 100$ points on each surface. Figure 2 shows an example of the feature points distribution on a surface. This distribution was obtained by iteratively minimizing Q in (8).

Some examples of the inter-surface distributions of M points in \mathcal{P}_j are shown in Fig.3(A). Each set \mathcal{P}_j is indicated by its color. As described in subsection 2.3, the prior distribution of a feature point P_j in a given normalized image I is estimated based on \mathcal{P}_j . Based on the resultant MN points $\{P_j^i | i = 1, 2, \dots, M; j = 1, 2, \dots, N\}$, we estimated $p(x_j)$, $p(I_j|x_j)$, and $p(x_k - x_j)$ as described in subsec 2.3. In addition, we construct a graphical model G . Figure 3(B) shows a mesh of feature points, in which any two dependent points are linked by edges.

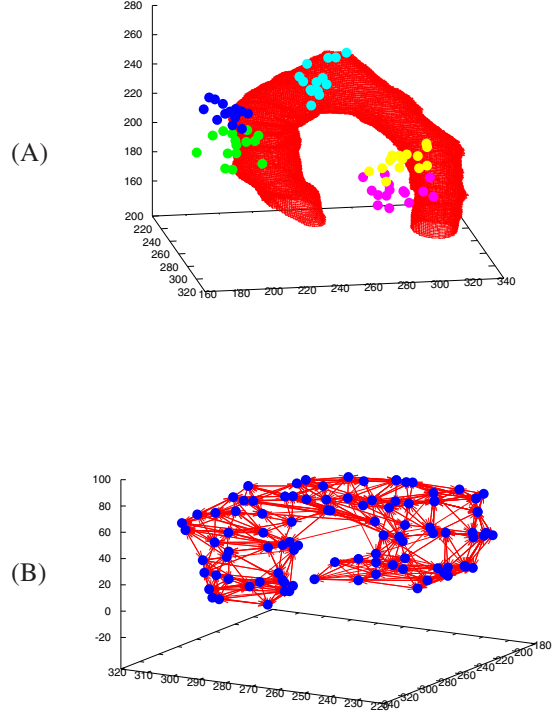


Figure 3. (A):Distributions of \mathcal{P}_j . From these points, $p(x_j)$ and $p(x_j - x_k)$ are estimated. (B):A mesh of feature points. Two feature points are linked by an edge if they are linked in the graphical model.

4.3. Registration of Surface Model

The first step of the registration is to estimate the location of each feature point P_i in a given normalized image. For this estimation we compute the distribution $p(x_i)p(I_i|x_i)$. In this article, in order to compare locations of feature points, we show experimental results obtained from closed dataset. Figure 4 (A) shows the locations of two of the feature points P_j^i , and (B) shows the corresponding prior distributions of $p(x_i)$. The likelihood distribution of each point estimated based on (10) is shown in Fig.4 (C). It should be noted that the distribution of $p(x_i)$ is more compact than that of $p(I_i|x_i)$ because we did not consider the similarities of the local images I_j^i for $i = 1, 2, \dots, M$ in spite that we required the compactness of the corresponding points. The joint (or posterior) distribution is shown in Fig.4 (D). When the local image pattern has strong characteristics, e.g. the corresponding feature point P_j is located on a strong edge at high probability, then the density distribution of

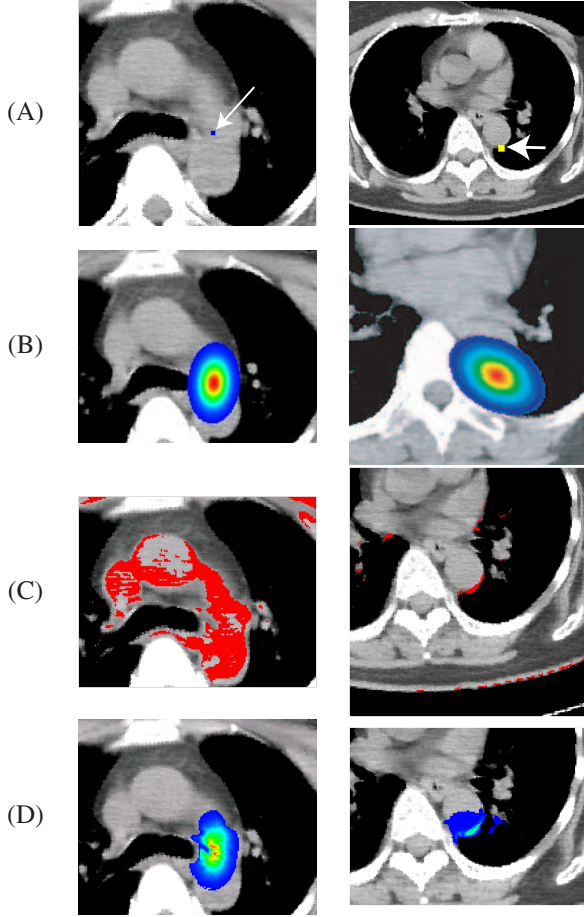


Figure 4. Two experimental results of the feature point detection. The images in each column shows the results obtained from an identical image. (A):Original image and (unknown) true location of feature point. The feature point shown in the right image is located on a strong edge. (B):Prior distribution $p(x_i)$. (C):Likelihood distribution $p(I_i|x_i)$. (D): Joint distribution $p(x_i, I_i)$.

$p(x_i)p(I_i|x_i)$ becomes localized. As shown in Fig.4 (D), it was hard to represent parametrically the joint distributions.

Because we have probabilistic models $p(x_j - x_k)$ that represent the local shape of the surface, we can improve the estimates of the locations of feature points. We drew $W = 1,000$ particles from $p(x_i)p(I_i|x_i)$ and estimated $p(x_i|I)$ by means of the NBP. Figure 5 shows $p(x_i)p(I_i|x_i)$ and $p(x_i|I)$ each of which is represented by the particles. The latter was obtained by the NBP. As shown in the figure, the particles gathered around the correct positions and the variances decreased. We computed the trace of the covariance matrix for the distribution of each feature point. For each surface, we obtained $N = 100$ values of the trace for each surface. Averaging those N values, we obtained

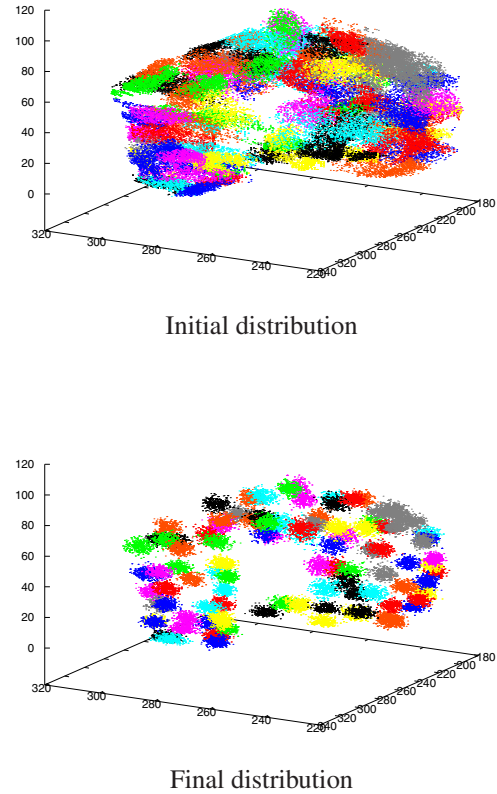


Figure 5. An example of the registration. Left: The initial distributions $p(x_i, I_i)$. Right: Resultant distributions $p(x_i|I)$; Each cloud of particles represents the conditional marginal distribution of each feature point.

the variance for each surface. The graphs in Fig.6 show the change of the variances with respect to the iteration in the NBP. As we can see, all of the $M = 15$ graphs decreased monotonically as the densities were updated.

Figure 7 shows a detailed example of the results of the NBP. The indicated slice images include the true position of the feature point. The images shown in Fig.7 (A) and (B) show the initial distributions of the particles, and (C) and (D) show the final distributions obtained by the NBP from (A) and (B), respectively. It should be noted that the variances of the distributions became smaller. This means that the confidences of the estimates were improved. The variance of the estimated distribution shown in (D) is smaller than that in (C), because the contrast of the edge is low around the feature point in (C). It is one of the advantages of the proposed method the confidence of the location of each feature point can be estimated for each feature point.

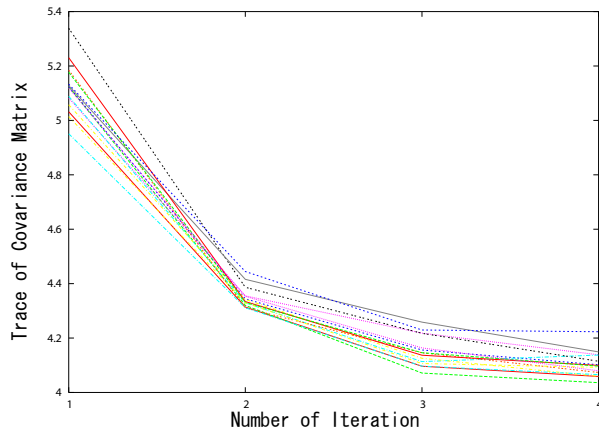


Figure 6. The change of the variance of feature points. The iterative updates in the NBP improved the confidence of the estimates.

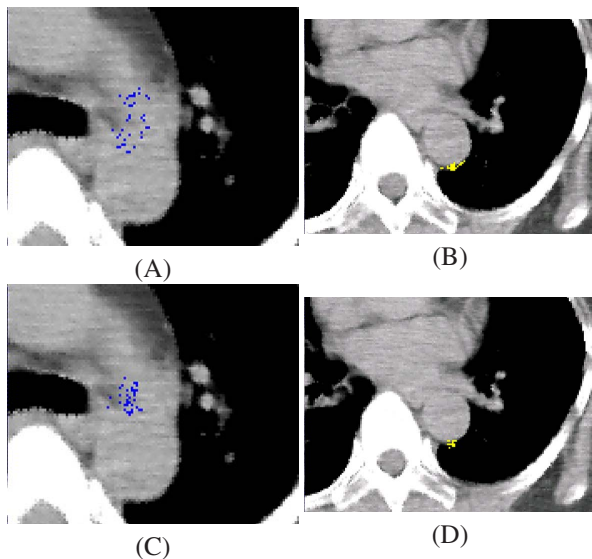


Figure 7. Improvement of the confidence of the estimates. (C) and (D) were obtained by the NBP from (A) and (B), respectively. The particles in (A) and (C) are indicated by blue dots, and those in (B) and (D) by yellow ones.

5. Conclusion

In this article, we proposed a feature-point-based non-rigid surface registration method. We represent a surface model using a graph of feature points $P(x_i)$ and the probability densities $p(x_i)$, $p(I_i|x_i)$, and $p(x_i - x_j)$. For constructing the representation, we distribute N corresponding points on each of the training surfaces based on the entropy-based particle system. We register the surface model with a new given image by estimating the locations of the feature points. For this estimation, we employed the non-parametric belief propagation. One of the advantages of

our method is that we can estimate not only the locations of feature points but also their probability distributions in the image. The future works include to evaluate concordance rate using open data and to analyze the effect of the error of the body shape normalization. It is known that the accuracy of estimates obtained by a belief propagation depends on the structure of the graphical model. The algorithm for inserting edges in the graph G has room for improvement.

References

- [1] M. A. Audette, F. P. Ferrie, and T. M. Peters. An algorithmic overview of surface registration techniques for medical imaging. *Medical Image Analysis*, 4(3):201–217, 2000. 1
- [2] P. Besl and N. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992. 1
- [3] A. Blake and M. Isard. *Chapter 2 of Active Contours*. Springer-Verlag, 1999. 1
- [4] A. Bronstein, M. Bronstein, and R. Kimmel. *Numerical Geometry of Non-Rigid Shapes*. Springer, 2008. 1
- [5] J. E. Cates, P. T. Fletcher, M. A. Styner, M. E. Shenton, and R. T. Whitaker. Shape modeling and analysis with entropy-based particle systems. In *Proceedings of the 20th International Conference on Information Processing in Medical Imaging*, pages 333–345, 2007. 1, 2
- [6] M. Droske and W. Ring. A mumford-shah level-set approach for geometric image registration. *SIAM journal on Applied Mathematics*, 66:2127–2148, 2006. 1
- [7] C. I. Fetita and F. Preteux. Bronchial tree modeling and 3d reconstruction. In *Proceedings SPIE Conference on Mathematical Modeling, Estimation and Imaging*, pages 16–29, 2000. 2
- [8] C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B (Methodological)*, 53(2):285–339, 1991. 1
- [9] D. Hill, P. Batchelor, M. Holden, and D. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1–R45, 2001. 1
- [10] Y. Kang, K. Engelke, and W. Kalender. A new accurate and precise 3-d segmentation method for skeletal structures in volumetric ct data. *IEEE Trans. Medical Imaging*, 22(5):586–598, May 2003. 2
- [11] X. Pennec, N. Ayache, and J.-P. Thirion. Landmark-based registration using features identified through differential geometry. *Handbook of medical imaging, Academic Press*, pages 499–513, 2000. 1
- [12] B. W. Silverman. *Density estimation: for statistics and data analysis*. London, 1986. 4
- [13] E. B. Sudderth, E. T. Ihler, W. T. Freeman, and A. S. Willsky. Nonparametric belief propagation. In *Proceedings of the IEEE Conference of Computer Vision and Pattern Recognition*, pages 605–612, 2003. 2, 4
- [14] M. Taron, N. Paragios, and M.-P. Jolly. Registration with uncertainties and statistical modeling of shapes with variable metric kernels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):99–113, 2009. 1