Preprints of the 9th IFAC Symposium on Biological and Medical Systems
The International Federation of Automatic Control
Berlin, Germany, Aug. 31 - Sept. 2, 2015

Improvement of Statistical Shape Models for Soft Tissues Using Modified-Coherent Point Drift

Mahdi Delavari*. Amir Hossein Foruzan*, **. Yen-Wei Chen**

**Department of Biomedical Engineering, Shahed University, Tehran, Iran (md.delavari@shahed.ac.ir, a.foruzan@shahed.ac.ir).*

***Intelligent Image Processing Lab, College of Information Science and Engineering, Ritsumeikan University, Japan (chen@ritsumei.ac.jp).*

Abstract: In this paper, we present a method to build statistical shape models for soft tissues. The efficiency of the model is close to the models built by the MDL (Minimum Description Length) algorithm; however, our method is much faster. The core of our method is based on finding corresponding points by a modified Coherent Point Drift method. The conventional CPD algorithm is modified to prepare a robust methods to find corresponding points of soft tissues. Our method achieved the Compactness, Generality, and Specificity of 18, 3.64(±0.85) mm, and 0.21(±0.06) respectively. Our results are close to the MDL results. The run-time of our method is 68 seconds which is faster than the MDL (3600 seconds) and TPS-RPM (390 seconds) methods.

Keywords: Corresponding points, Statistical Shape Models, Modified Coherent Point Drift, Medical Image Registration.

1. INTRODUCTION

Statistical Shape Models (SSMs) (Cootes et al., 1995) use statistical information of shapes to interpret input images. Construction of an SSM for a non-rigid tissue, such as human's liver, is very difficult since the variation of the shape is large among cases. We aim at constructing a model to describe variations of the shape. To perform this, we need a set of training images including legal variations of the object's shape. The strength of an SSM depends on the size and diversity of the training set. The first step to build the model is to represent a 2D/3D shape by some landmarks located on the border/surface of the object's shape. Assuming the correspondence of the landmarks is known, the mean shape is built by averaging the landmarks of the aligned shapes. The modes of the variations are obtained by calculating the landmarks covariance matrix and are then reduced by the PCA (Principal Component Analysis) analysis (Heimann and Meinzer, 2009).

One of the main challenges in building a shape model is to establish correct correspondences between landmarks. In cases of 2D shapes, it can be achieved manually by an expert. However, in cases of 3D shapes, finding the corresponding landmarks is a time-consuming and difficult task. Thus, establishing accurate correspondences between two shapes is considered as one of the fundamental challenges of shape model building (Heimann and Meinzer, 2009).

The correspondence problem is solved by three types of approaches (Sotiras et al., 2013). In the first class, one shape is registered over the other shape and the nearest neighbors are considered as corresponding points. In the second approach, the points are classified by higher level structures such as lines, curves and surfaces. These structures are used as features to find corresponding points. The third method solves registration and correspondence problems simultaneously.

For a point on the first shape, its corresponding landmark on the second shape is its nearest neighbor. Registration methods are generally divided into rigid and non-rigid approaches. Rigid methods only consider global deformations of the shape. Therefore they can be applied to register rigid objects such as bones. Non-rigid algorithms describe both global and local variations. The accuracy and precision of registration results are affected by the selection of the non-rigid transformation kernel (Sotiras et al., 2013).

In the second approach, the correspondence problem is solved without calculating transformation functions. To obtain acceptable results, the features should be selected accurately, a large number of points are needed to have sufficiently smoothed curves, and deformations of the two shapes should be large. To search the whole points for finding corresponding landmarks is very time-consuming and sometimes impossible. The feature selection suffers from noise and outliers. Some researchers have used weighted graphs to restrict the search region. For an organ, the graphs should be designed manually and cannot be generalized to other organs (Sotiras et al., 2013).

The third method is the best among the three approaches which solves both the correspondence and transformation problems concurrently. It is an iterative method which finds corresponding point in one step and employs the landmarks to find transformation function (Sotiras et al., 2013).

2. RELATED WORKS

The well-kwon Iterative Closest Point (ICP) (Besl and Mckay, 1992) is a very fast method and its convergence to a local

minimum is guaranteed. This method utilizes rigid transformation and iteratively finds binary correspondences between landmarks using nearest-neighbor relationships in each step. It requires that the initial positions of the shapes to be sufficiently close to each other which is not always possible. The use of binary correspondences may cause the algorithm to trap in a local minimum and may lead to wrong correspondences. Some modifications are applied to the classic ICP to improve the capability (EM-ICP (Hermans et al., 2011), Modified-ICP (Niculescu et al., 2007), Affine-ICP (Zheng et al., 2008)) but it still fails to find correct correspondences for non-rigid tissues and in the presence of noise and outliers.

The TPS-RPM (Thin Plate Spline – Robust Point Matching) algorithm described by Chui et al. (2003) uses Gaussian mixtures (GMM) and the non-rigid TPS transformation function and solves the correspondence problem by EM (Expectation Maximization) algorithm (Moon, 1996). The EM algorithm is a dual-update process which in the E-step the correspondences are estimated and in the M-step the transformation is updated. This algorithm is guaranteed to converge from almost any initial condition. Contrary to the ICP method, this method utilizes fuzzy correspondence assignment. The variances of the GMM are reduced through Deterministic Annealing process which results in the conversion of the fuzzy correspondence into binary correspondence. To overcome outliers, a uniform/Gaussian distribution is included into the GMM and its variance is fixed during Deterministic Annealing procedure. At the end of the TPS-RPM, the points are not completely aligned to each other. Yang et al. (2011) introduced a dynamic algorithm to reject outliers during the process and matches the corresponding points more accurately by a local refinement technique. Spatial deformation of the moving shape is lower in the Yang's method compared with the Chui's. However, the computational cost of the algorithm is increased and it is implemented for 2D data. Mourning et al. (2010) used GPU to accelerate TPS-RPM. They had not a noticeable improvement (in run-time) for shapes with less than 2000 points.

Mcneill et al. (2006) divided a shape into its natural subdivisions and constrained each part of the moving shape to be matched to its analogous part of the reference shape. The natural subdivisions are attained by the Natural Point Distribution method. For example, the 2D contour of a human body can be divided into 5 parts: two hands, two feet, a head and a trunk. This method requires an expert to define subdivisions and the accuracy of the algorithm largely depends on the accuracy of the divisions.

Coherent Point Drift (CPD) algorithm described by Myronenko et al. (2010) uses a non-rigid approach to shape registration and point matching using GMMs. Contrary to the TPS-RPM method which assumes a TPS kernel for transformation of the points, the Myronenko's method does not assume any transformation function explicitly. They implicitly add a motion constraint on the velocity field of Gaussian centroids. Using Motion Coherence Theory (Yuille, 1988), inclusion of higher order derivatives of velocity field results in smoother transition compared to the TPS-RPM method. However, the method is sensitive to the selection of the parameters (especially the variance of the Gaussian mixtures).

MDL introduced by Davies et al. (2008) for both 2D and 3D shapes. They mapped each shape in the training set onto a sphere. A point on a sphere has two degree of freedom in polar space (θ,φ) which are polar angles. Displacement of the points leads to new correspondences. An objective function was defined to minimize the description length of the model. Using gradient descent strategy, the points are moved to get the most compact model. Then, the points were mapped back to the real world and the model was made. The MDL method has a large computational cost and its convergence is not guaranteed. Meanwhile, a re-parameterized shape has not an even distribution of landmarks over its surface.

In this paper, we propose a method to find corresponding points approaching the MDL algorithm with respect to a model evaluation metrics while resolving time-cost and conversion issues. The innovations of our method include establishing correspondences for human liver shapes by means of Coherent Point Drift (CPD) method, optimization and conscious selection of non-rigid registration parameters, decreasing computational cost, and developing a robust registration algorithm. The remaining of the paper is organized as follows. In the third section the proposed method is described. The results of our method are presented in section four. Section five and six are devoted to discussion and conclusion respectively.

3. PROPOSED METHOD

The flowchart of our method is shown Fig. 1. Our method begins with reading input mask data. After smoothing the liver masks by a 3x3x3 mean filter, surface meshes are built using marching cube algorithm (Lorensen et al., 1987). A typical mesh contains about 240,000 points. Finding correspondence between such two sets requires a large computational cost. Thus, the number of the points are reduced to 1000 landmarks. Then, the Modified-CPD method is employed to establish a one-to-one correspondence between landmarks. Finally, the meshes are aligned by the Generalized Procrustes Analysis (Gower, 1975) and the shape model is built.

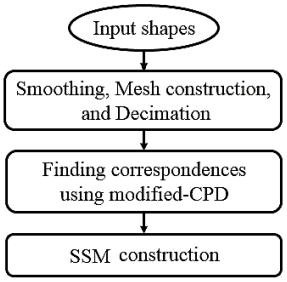

Fig. 1. The flowchart of the proposed method.

3.1 Mesh construction

A generalized statistical shape model requires a relatively large amount of input data which comprises all valid variations of the shapes. Typical liver shapes employed in the construction of the shape model are shown in Fig. 2.

We employ the marching cube algorithm to construct meshes and reduce the resulting mesh points to 1000. Instead of

building meshes in the voxel coordinates, they are constructed in the world coordinate by considering the element sizes.

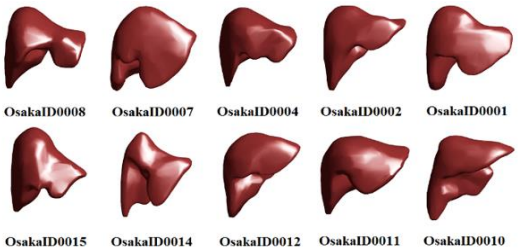

Fig. 2. Typical liver shapes in training set.

3.2 The Modified-CPD method

CPD is a point-based registration method using the kernel density estimation technique. To register a moving shape on a fixed shape, it adds a constraint on the velocity field to insure the movement coherency of nearby points. The method solves both non-rigid transformation and point correspondences problems simultaneously using EM strategy. Motion coherency imposes smoothness on the transformation function and forces the neighboring points to move coherently. Thus the topology of the moving shape is preserved during registration.

Assuming two liver meshes are represented by $X = (x_1, \dots, x_N)^T$ and $Y = (y_1, \dots, y_M)^T$ where X belongs to the fixed shape and Y denotes the moving shape. The distribution of the points x_i are modeled by a Gaussian mixture. The center, covariance matrix and weight of a component are y_m, $\sigma^2 I_D$ and *1/M* respectively (1).

$$p(x) = \sum_{m=1}^{M} \frac{1}{M} N(y_m, \sigma^2 I_D) \tag{1}$$

The points of *Y* are considered as the center of the equally weighted (*1/M*) Gaussian kernels with identical covariance matrices ($\sigma^2 I_D$). The new locations of the centers are then found using their previous coordinates by using (2).

$$Y = v(Y_0) + Y_0 \tag{2}$$

v is a continuous velocity field of the centroid movements and Y_0 is the initial locations of them. The prior knowledge of the centroids is defined as (3).

$$p(Y \mid \lambda) \propto \exp(-\frac{\lambda}{2}\phi(Y)) \tag{3}$$

λ is a weighting constant and $\phi(Y)$ is a function that regularizes the motion smoothness. We aim to estimate the probability density of *X* by means of *Y*. thus, the Bayes theorem is employed to calculate the posterior probability ($p(Y|X)$) or the likelihood probability ($p(X|Y)$) by minimizing the energy function in (4).

$$E(Y) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\|\frac{x_n - y_m}{\sigma}\|^2} + \frac{\lambda}{2}\phi(Y) \tag{4}$$

x_n and y_m represent each point in X and Y respectively and σ is the standard deviation of each component in GMM.

The velocity field of the moving centroids must be smooth. Smoothness of a signal can be evaluated by measuring the amount of energy in high frequencies. The high frequency content of a signal can be obtained by multiplying it to a Gaussian high pass filter (HPF) or similarly by dividing it to a Gaussian low pass filter (LPF) in frequency space. Thus, minimizing the high frequencies of the velocity filed may leads to a smooth movement. The high frequency content of the velocity filed can be obtained by (5).

$$\phi(v) = \int \frac{|\tilde{v}(s)|^2}{\tilde{G}(s)} ds \tag{5}$$

$\tilde{v}$ is the Fourier transform of the velocity field and $\tilde{G}$ is a Gaussian LPF. Substitution of (5) in (4) may leads to (6).

$$E(\tilde{v}) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\|\frac{x_n - y_m}{\sigma}\|^2} + \frac{\lambda}{2} \int_{R^d} \frac{|\tilde{v}(s)|^2}{\tilde{G}(s)} ds \tag{6}$$

First, we substitute the Fourier transform of velocity field in (5) and calculate the derivation of (6) with respect to *v*. then, it can be shown that the velocity filed minimizes the energy function (6) is in the form of (7) (similar to Motion Coherence Theory (Yuille et al., 1988)).

$$v(z) = \sum_{m=1}^{M} w_m G(z - y_{0m}) \tag{7}$$

G is a Gaussian kernel which is not related to the GMM of *Y* and w_m are the weights of *G*. Equation (7) is equivalent to a weighted sum of all orders of derivations of velocity field and is in the form of (8), where *D* is the derivation operator that acts as (9).

$$\frac{\lambda}{2}\int \frac{|\tilde{v}(s)|^2}{\tilde{G}} ds \rightarrow \frac{\lambda}{2}\int \sum_{m=1}^{\infty} \frac{\beta^{2m}}{m! 2^m} (D^m v)^2 \tag{8}$$

$$\begin{cases} D^{2m} v = \nabla^{2m} v \\ D^{2m+1} v = \nabla(\nabla^{2m} v) \end{cases} \tag{9}$$

By substituting (2) and (8) in (6), the (6) changes into (10).

$$E(W) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M} e^{-\frac{1}{2}\left\| \frac{x_n - y_{0m} - \sum_{k=1}^{M} w_k G(y_{0k} - y_{0m})}{\sigma} \right\|^2} + \frac{\lambda}{2} tr(W^T G W) \tag{10}$$

$W_{M \times D}$ is the weight matrix and $G_{M \times M}$ is a symmetric square matrix with the elements $g_{ij} = e^{-\frac{1}{2}\left\|\frac{y_{0i} - y_{0j}}{\beta}\right\|^2}$.

An upper bound (11) can be found for the energy function (10). Thus, minimizing *Q* may cause minimizing (10).

$$Q(W) = \sum_{n=1}^{N}\sum_{m=1}^{M} P^{old}(m \mid x_n) \frac{\|x_n - y_{0m} - G(m,.)W\|^2}{2\sigma^2} + \frac{\lambda}{2} tr(W^T G W) \tag{11}$$

P^{old} is the posterior probability matrix which is calculated from current locations of the points of *Y*. *G(m,.)* denotes the m^{th} row of *G*. By computing the derivation of (11) with respect to *W* and equating it to zero, *W* can be computed from (12).

$$((diag(P1))G + \lambda\sigma^2 I)W = PX - diag(P1)Y_0 \tag{12}$$

The diag(.) notation denotes a diagonal matrix and 1 is a column vector of all ones. P is the posterior probability matrix with the elements calculating by (13).

$$p_{mn} = e^{-\frac{1}{2}\left\| \frac{y_m^{old} - x_n}{\sigma} \right\|^2} \Big/ \sum_{m=1}^{M} e^{-\frac{1}{2}\left\| \frac{y_m^{old} - x_n}{\sigma} \right\|^2} \tag{13}$$

After finding W, the centroids are moved to new locations (14).

$$Y = Y_0 + GW \tag{14}$$

Solving (12) for W is the M-step and calculating P is the E-step of the EM algorithm. The convergence of the EM to a local minimum is guaranteed from almost any starting point.

3.2 Free parameters

σ, λ, and β are the free parameters in the CPD algorithm. The parameter λ in (11) compromises between the point registration and the motion smoothness and is set to 2 according to (Yuille, 1988). The parameter β in G regularizes the interactions among the points in Y and determines how many neighboring points in Y are constraint to move together coherently. Small values of β leads to a local transformation, while big values lead to a rigid transformation. Parameter β is set to 1.87 so that all of the elements of G get above 0.36. Because if we consider that the effective range of a Gaussian is its mean ± standard deviation, the standard deviation is where the Gaussian reaches to 0.36 of its final value. This cause the farthest points (Gaussian centroids) in Y to have at least a small effect on each other and thus the topology of the moving shape is preserved during registration procedure.

The parameter σ in (11) is the standard deviation of each Gaussian in the GMM. This parameter is decreased by a constant factor (between 0.93 and 0.97) during the registration process using the Deterministic Annealing strategy. Due to normalization of the shapes before registration, the initial value of the parameter is set to 1. Thus, almost every point in X and Y may correspond to each other at the start of the process. The annealing factor is set to 0.95.

Due to complexity of liver shapes, we found that if the parameter σ is decreased less than a threshold, some points in X may fall far from the Gaussian contours and gross errors happen. We quantitatively found the threshold value for sigma to be 0.028 that is equivalent to 70 iterations of the algorithm. The flowchart of the proposed method is shown in Fig. 3.

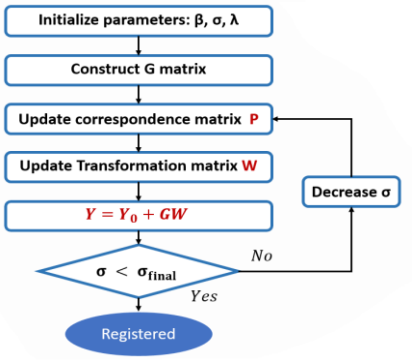

Fig. 3. The Modified-CPD flowchart.

3.3 Model construction

After finding corresponding points, each shape in training set must be converted to a vector in the form of (15) and then each shape vector is normalized such that $|X| = 1$.

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & z_m \end{pmatrix} \rightarrow X = (x_1, y_1, z_1, \ldots, x_m, y_m, z_m)^T \tag{15}$$

Then, the normalized shapes are aligned using Generalized Procrustes Analysis and the mean shape is calculated (16).

$$\bar{X} = \frac{1}{s}\sum_{i=1}^{s} X_i \tag{16}$$

s is the number of shapes in training set. The covariance matrix corresponding to training shapes is then calculated using (17).

$$S = \frac{1}{s-1}\sum_{i=1}^{s}(X_i - \bar{X})(X_i - \bar{X})^T \tag{17}$$

If the eigenvalues of S are sorted in a descending order, any valid shape can be reconstructed using a linear combination of the mean shape and the eigenvalues using (18).

$$X \approx \bar{X} + \Psi b \tag{18}$$

Ψ is the eigenvector matrix corresponding to the sorted eigenvalues and b is a parameter vector.

4. RESULTS

We evaluated the proposed method with both rigid (ICP) and non-rigid (TPS-RPM and MDL) approaches. Comparison was performed using Compactness, Specificity, and Generalization ability.

4.1 Dataset

To evaluate our method, we employed the 2nd phase of 30 CT-scan images of the abdominal region from 16 males and 14 females in the range of 20 to 75 years old. They included 12-bit DICOM images of 512 x 512 x 159 pixels and a size of 0.625 x 0.625 x 1.25 mm^3. Input images were acquired by LightSpeed Ultra GE scanners with eight detectors at Osaka University Hospital. The masks of the data were prepared by a radiology specialist.

The codes are implemented in MATLAB and run in a personal computer with Windows 8.1 OS, AMD-Fx4100 3.6GHz CPU and 8 GB dynamic RAM.

4.2 Evaluation measures

The Compactness of a model is calculated using (19). In (19), λ_i is the i^{th} eigenvalue, $N \times d$ is the total number of the modes, and m is minimum number of the modes so that the parameter C is above 0.98.

$$C = \frac{\sum_{i=1}^{m}\lambda_i}{\sum_{i=1}^{N\times d}\lambda_i} \tag{19}$$

In Fig. 4, the compactness of the model compared to three other methods (TPS-RPM, ICP, and MDL) is shown.

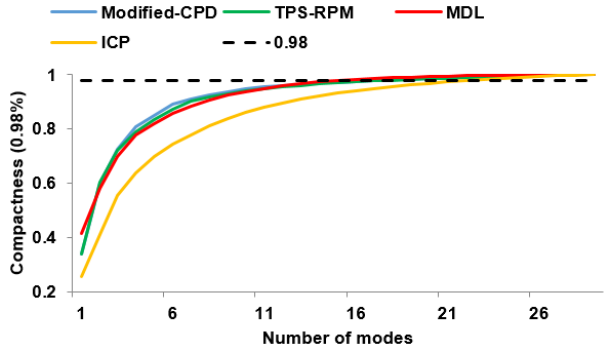

Fig. 4. Compactness of the models versus number of modes.

As can be seen in Fig. 4, the compactness of the ICP, TPS-RPM, MDL, and our method are 23, 19, 18, and 16 respectively.

The Generalization ability of a model is defined as the capability of the model to describe valid shapes which are not seen in the training set (20). The measure is calculated by the leave-one-Out approach.

$$G(m) = \frac{1}{M}\sum_{i=1}^{M}\|x_i - x_i'\| \quad (20)$$

In (20), x is the reconstructed shape by the proposed model, x' is the input shape, and M is the number of training shapes. The Generalization ability is measured in millimeter. In Fig. 5, the Generalization ability for 8 data is shown for the methods.

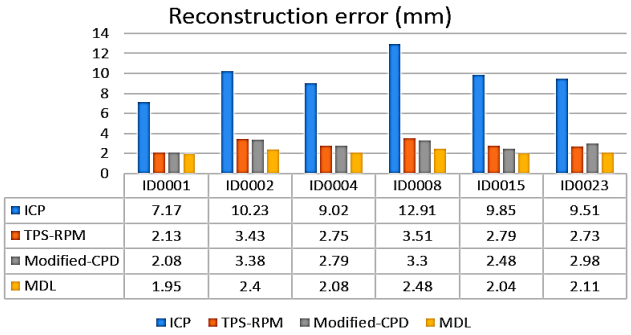

Fig. 5. The Generalization ability for 8 different data for different methods.

The generalization ability for all input dataset was calculated and the average of the results are shown in Table 1.

Table 1. The average results of the Generalization.

Method	ICP	TPS-RPM	Modified-CPD	MDL
Compactness (98%)	23	19	18	16
Generality (mean) (mm)	9.8±2.45	3.78 ± 0.96	3.64 ± 0.85	2.64 ± 0.65
Generality (RMS) (mm)	10.8 ± 2.72	4.29 ± 1.18	4.11 ± 0.98	3.11 ± 0.75
Coefficient of variation	0.25	0.25	0.23	0.25

As can be seen, the Generalization ability was computed using the number of the modes which make 98% compactness. The Coefficient of variation is an index to compare variation of different results with different means. It is defined as (21).

$$CV = \frac{standard\ deviaion}{mean} \quad (21)$$

The Specificity of a model is a measure to describe if the model can construct valid shapes or not. If the set covered by a model includes a large space in the shape space, it can include invalid shapes too. To calculate Specificity, M new random shapes are constructed using the model. To construct a random shape, the shape parameter b is randomized and confined between $[-2\sqrt{\lambda_i}, +2\sqrt{\lambda_i}]$. Then, the distance between the new shape and the nearest training shape is used and the average of these distances is employed as the Specificity index (22).

$$S(m) = \frac{1}{M}\sum_{A=1}^{M}\min_i\|y_A - x_i\| \quad (22)$$

y_A is the random shape constructed by the model, and x_i is a shape in the training set, and M is the size of training set. To calculate Specificity, 1000 random shapes were built in the shape space and the average Specificity for ICP, TPS-RPM, MDL, and the proposed method were calculated (Table 2).

Table 2. The average Specificity measures.

Method	ICP	TPS-RPM	Modified-CPD	MDL
Compactness (98%)	23	19	18	16
Specificity (mean)	0.30 ± 0.05	0.21 ± 0.06	0.21 ± 0.06	0.21 ± 0.06
Specificity (RMS)	0.30 ± 0.18	0.22 ± 0.16	0.22 ± 0.17	0.22 ± 0.16
Coefficient of variation	0.17	0.29	0.29	0.29

5. DISCUSSION

For the compactness measure, we used 98% which is a severe metric for SSM models. As is shown in Fig. 4, number of the modes in our method is 18. It reveals that our method can be represented by lower number of modes compared to other methods. However, compared to the MDL algorithm, the MDL produces the most compact model. The compactness of the TPS-RPM and Modified-CPD methods are the same for the first 4 modes. However, the Modified-CPD method reaches faster to 98% threshold compared to the TPS-RPM algorithm. The ICP algorithm has the lowest compactness among other methods, with 23 modes, which can be explained by its rigid registration method.

With respect to the Generalization ability, our method got the 2nd rank (after the MDL). The variation of the measure tells us about the outliers in the results. To compare variations of the results with different means, we employed the Coefficient of Variation (CV). The CV metric of our method is lowest among other method which reveals that the outliers are not very far from the mean compared to other methods. As can be shown in Fig. 5, the Generalization ability of the dataset ID0001 is the lowest compared to other data. This can be described by the similarity of this data to the mean shape of the model. If the number of modes increases, the Generalization ability measure is decreased. Instead of using the same number of modes to compute Generalization ability, we fixed the compactness of all methods to 98% and found the Generalization ability for other methods. The Generalization ability of the Modified-CPD and TPS-RPM methods are similar for up to the 4th mode. However, the error increases in case of the TPS-RPM method which is described by the compactness of the both models.

Regarding the Specificity index, the measure for the Modified-CPD, TPS-RPM, and the MDL methods are similar. For the case of the ICP method, the result has the largest error which can be described by the incorrect correspondences of the method.

An important index of our method is the run-time of the code. For example, the MDL is a very efficient method with respect to the Compactness, Generalization ability, and Specificity. However, it takes a lot of time to run. The long run-time of the method makes it an unpopular algorithm and most researchers prefer to use simple and lower efficient but faster methods (such as the ICP algorithm). The run-time of our method and other methods are shown in Table 3.

Table 3. The average run-time of different methods.

Method	MDL	Modified-CPD	TPS-RPM	ICP
Run-time (second)	3600	68	390	30

The ICP method is the fastest among other methods; however, it gets the lowest indices with respect to SMM model measures. The method is faster compared to the MDL approach. This feature augmented by other measure of our method can be used to conclude that the results of our method are similar to the MDL algorithm; however, it runs much faster than the MDL. Thus, our method is an appropriate substitute for the MDL algorithm.

6. CONCLUSION AND FUTURE WORKS

We proposed a Modified-CPD method to find corresponding points in a set of input meshes. The results of our method are better than the ICP and TPS-RPM methods. Our results are close to the MDL method results. However, our method is faster compared to the MDL and TPS-RPM. In future, we plan to divide the surface of soft tissues into simpler surfaces based on the natural point distribution (Mcneill et al., 2006). Also, we decide to employ the Couinaud approach to divide the volume of the liver into 8 sub-sections.

ACKNOWLEDGEMENTS

This work is supported in part by the Grant-in Aid for Scientific Research from the Japanese Ministry for Education, Science, Culture and Sports (MEXT) under the Grant No. 15H01130, in part by the MEXT Support Program for the Strategic Research Foundation at Private Universities (2013-2017), and in part by the R-GIRO Research Fund from Ritsumeikan University.

REFERENCES

Besl, P. J., Mckay, N. D. (1992). A method for registration of 3D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(2), pp. 239–256.

Chui, H., Rangarajan, A. (2003). A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2-3), pp. 114–141.

Cootes, Timothy F., Christopher J. Taylor, David H. Cooper, and Jim Graham (1995). Active shape models-their training and application. *Computer vision and image understanding*, 61(1), pp. 38-59.

Davies, R., Twining, C., Taylor, C. (2008). *Statistical Models of Shapes: Optimization and Evaluation*, Springer press, London.

Du, S., Zheng, N., Meng, G., Yuan, Z. (2008). Affine Registration of Point Sets Using ICP and ICA. *IEEE Signal Processing Letters,* 15, pp. 689-692.

Gower, J.C. (1975). Generalized procrustes analysis. *Psychometrika*, 40(1), pp. 33-51.

Heimann, T., et al. (2009). Statistical shape models for 3D medical image segmentation: A review. *Medical image analysis*, 13(4), pp.543-563.

Hermans, J., Smeets, D., Vandermeulen, D., Suetens P. (2011). Robust point set registration using EM-ICP with information-theoretically optimal outlier handling. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2465–2472.

Lorensen,W.E., Cline, H. (1987). Marching Cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21(4), pp. 163-169.

Mcneill, G., Vijayakumar, S. (2006). Part-based Probabilistic Point Matching using Equivalence Constraints. *Advances in Neural Information Processing Systems,* 19, pp. 969-976.

Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal Processing Magazine,* 13(6), pp. 47-60.

Mourning, C., et al. (2010). GPU Acceleration of Robust Point Matching. *Advances in Visual Computing Lecture Notes in Computer Science,* 6455, pp. 417-426.

Myronenko, A., Song, X. (2010). Point-Set Registration: Coherent Point Drift. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32(12), pp. 2262-2275.

Niculescu, G., Forand, D., Nosher, J. (2007). Non-rigid registration of the liver in consecutive CT studies for assessment of tumour response to radiofrequency ablation. *29th Annual International Conference of the IEEE in Engineering in Medicine and Biology Society*, pp. 856-859.

Sotiras, A., Davatzikos, C., Paragios, N. (2013). Deformable Medical Image Registration: A Survey. *IEEE Transaction on Medical Imaging,* 32(7), pp. 1153 – 1190.

Wang, P. et al. (2011). A refined coherent point drift (CPD) algorithm for point set registration. *Science China Information Sciences*, 54(12), pp. 2639-2646.

Yang, J., et al. (2011). Robust hybrid method for nonrigid image registration. *Pattern Recognition*, 44(4), pp. 764–776.

Yuille, A. L., Grzywacz, N.M. (1988). The motion coherence theory. *Int. J. Computer Vision*, 3, pp. 344-353.