Int J CARS
DOI 10.1007/s11548-014-1000-5

ORIGINAL ARTICLE

Capturing large shape variations of liver using population-based statistical shape models

Amir H. Foruzan · Yen-Wei Chen · Masatoshi Hori · Yoshinobu Sato · Noriyuki Tomiyama

Received: 9 November 2013 / Accepted: 31 March 2014

Abstract
Purpose Statistical shape models (SSMs) represent morphological variations of a specific object. When there are large shape variations, the shape parameters constitute a large space that may include incorrect parameters. The human liver is a non-rigid organ subject to large deformations due to external forces or body position changes during scanning procedures. We developed and tested a population-based model to represent the shape of liver.
Methods Upper abdominal CT-scan input images are represented by a conventional shape model. The shape parameters of individual livers extracted from the CT scans are employed to classify them into different populations. Corresponding to each population, an SSM model is built. The liver surface parameter space is divided into several subspaces which are more compact than the original space. The proposed model was tested using 29 CT-scan liver image data sets. The method was evaluated by model compactness, reconstruction error, generality and specificity measures.
Results The proposed model is implemented and tested using CT scans that included liver shapes with large shape variations. The method was compared with conventional and recently developed shape modeling methods. The accuracy of the proposed model was nearly twice that achieved with the conventional model. The proposed population-based model was more general compared with the conventional model. The mean reconstruction error of the proposed model was 0.029 mm while that of the conventional model was 0.052 mm.
Conclusion A population-based model to represent the shape of liver was developed and tested with favorable results. Using this approach, the liver shapes from CT scans were modeled by a more compact, more general, and more accurate model.

Keywords Statistical shape model · Population-based shape model · Shape representation · Medical image analysis

Electronic supplementary material The online version of this article (doi:10.1007/s11548-014-1000-5) contains supplementary material, which is available to authorized users.

A. H. Foruzan (✉)
Department of Biomedical Engineering, Engineering Faculty, Shahed University, Tehran, Iran
e-mail: a.foruzan@ece.ut.ac.ir; a.foruzan@shahed.ac.ir

A. H. Foruzan · Y.-W. Chen
Intelligent Image Processing Lab, College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan

Y.-W. Chen
e-mail: chen@is.ritsumei.ac.jp

M. Hori · N. Tomiyama
Department of Radiology, Graduate School of Medicine, Osaka University, Osaka, Japan
e-mail: mhori@radiol.med.osaka_u.ac.jp

Y. Sato
Division of Image Analysis, Graduate School of Medicine, Osaka University, Osaka, Japan
e-mail: yoshi@image.med.osaka-u.ac.jp

Introduction

Background

Statistical shape models (SSMs), introduced by Cootes et al. [1], have been used as a powerful tool to model variations of both 2D and 3D shapes. In medical image analysis, the SSM has been proved to be useful in a wide range of applications including segmentation [2], shape variations analysis [3], treatment planning [4], identification and shape recovering [5]. Concerning body organs, the shape of a tis-

Published online: 20 April 2014

sue does not follow a unique pattern. The SSM considers shape variations and is, therefore, the preferred tool in such cases. The robustness of the technique against noise makes it an appealing method for segmentation. Typical applications include segmentation of heart, liver, bones, and spleen [2]. The SSM was also used to study shape variations of the wrist bones, diagnose bone pathologies, and making prosthesis [4]. Pepe et al. [3] employed the SSM to study asymmetries in the normal human brain. Buchaillard et al. [5] employed the SSM to reconstruct the shape of teeth given crown information. Chen et al. [6] used SSM for diagnosis and quantification of cirrhotic livers.

Modeling objects with large shape variations require a large number of modes and a large number of training datasets. The compactness of the model is an important evaluation metric and is used to compare models together. Researchers proposed different methods to build models employing as few as number of modes by finding corresponding points more accurately. To solve the correspondence problem, Cootes and Taylor [7] employed finite element models (FEMs) to describe shapes more accurately. However, the computational cost of the FEM model is high. Lemecker et al. [8] manually divided the surface of the liver into four patches, mapped each patch into a disk, and found the corresponding points on the disks. Their goal was to minimize local distortion of the patches. Okada et al. [9] employed inter-organ relationships to deal with large variations of shape and embedded it into a hierarchically organized statistical atlases. They used their proposed method to segment liver, gallbladder, and vena cava more accurately, compared with their previous methods. Feng and Ip [10] proposed a multi-resolution approach to build a model using low number of training samples. However, they applied their method on a few number of test data, and they did not tell anything about variation diversity of the test images. Davatzikos et al. [11] also considered a hierarchical shape model using wavelet transform to overcome the inability of active shape models to capture large shape variability of soft tissues. They used their method to segment corpus callosum and hand images which were considered as relatively rigid shapes.

The generalization of the model is another challenge in building SSM models. The eigenvalues and eigenvectors are combined through the shape parameters to resemble new shapes. However, in some cases where a new shape has not already been seen in the training set, the reconstruction is not accurate enough. If the number of training images is large, the number of salient eigenvalues and eigenvectors will be more. In the case of liver where shape variation is large, modeling the shapes by a single SSM model is a challenging task [10]. If a single model is used, minor shape differences will be neglected.

The main focus of this paper is to model the shape of the liver which is considered as a soft tissue and has complex shape variations. We classify input shapes into subgroups and build a model for individual subgroups. However, other researchers have dealt with the idea of modeling patient or group-specific shape prior including sparse shape composition (SSC) [12–15]. SSC, proposed by Zhang et al. [14], is a novel idea to overcomes three challenges concurrently: complex shape variations, gross error due to image appearance cues, and loosing local details in shape models. Their method produced more accurate segmentation results in lung and liver images. They did not compare their methods with other researched using compactness, generality and specificity measures. However, our main focus is to develop a general, specific, and compact model for liver.

In this paper, we propose a new approach to capture large shape variations using a hierarchical model. The training sets are initially employed to build a general model. The shape parameters corresponding to the general model are used to cluster input images into different populations. Then, individual models are build for the populations. Using the proposed multi-layer model, shape variations are captured better compared with other shape models. Evaluation of the proposed method by compactness, reconstruction accuracy, and generality of our model proves this fact. In "The proposed method" section, we describe the proposed method. The results are shown in "Experiments and results" section and we discuss the results in "Discussion" section. Finally, the paper is concluded in "Conclusion and future works" section.

The proposed method

Motivation

To describe the motivation behind our method, we visualized the segmented livers in 5 CT datasets in Fig. 1.

As can be seen in Fig. 1, it is confirmed that the liver has a large shape variations which makes it very difficult to be captured by the conventional method of shape modeling. Another important point which can be observed in Fig. 1 is that some shapes are nearly similar to each other. This gives us the motivation for dividing available images into different populations and building individual models for each group. We show that the population-based model results in a more compact, more accurate, and more generalized shape model. Input images may be clustered using the shape parameters of a conventional SSM model built using the whole training set. In this research, we divide the input dataset into two groups.

Basic concepts of statistical shape models

A brief description of the SSM is shown in Fig. 2. In the SSM, a shape is conventionally represented by a set of landmarks located on the boundary or surface of the shape, in case of

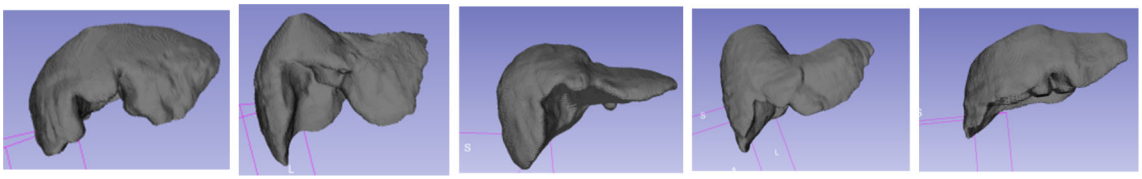

Fig. 1 Typical shape variations of the available liver dataset

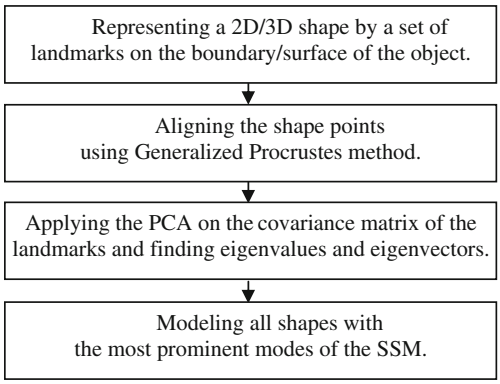

Fig. 2 Steps of building a statistical shape model

2D or 3D shapes, respectively. To compensate for the effect of any transformation (scale, rotation, and translation) on the analysis, the shapes have to be normalized. The Generalized Procrustes analysis normalizes a set of shapes on an optimum mean shape which is built using the input shapes [16]. Then, the mean and covariance matrix of the landmarks' coordinates are found and the principal component analysis (PCA) is applied on the covariance matrix. The mean shape together with the most prominent eigenvalues and eigenvectors of the matrix is used to model the shape variations.

Landmark selection is a crucial step in the whole process. The main assumption in this step is that the sets of points in all shapes correspond to each other. In real cases, finding correspondence is a difficult task. In 2D cases, e.g., in modeling the shapes of hands, this can be done interactively by a user or automatically through finding prominent points of the boundary. In 3D shapes, however, point selection is a tedious and not a repeatable task which is usually preferred to be done automatically than manually. In the case of objects consisting non-rigid shapes, the problem gets worse since a single object may have different shape when force is applied on it. In medical image processing, the shape of liver is considered as a non-rigid shape which has not a unique shape in different patients (Fig. 1). Also, it deforms easily when the patient is positioned differently during scanning procedure. Thus, it is considered as a challenging task to capture large shape variations of the liver by a single statistical shape model [10].

The best solution to finding corresponding points is the "Minimum Description Length" (MDL) algorithm proposed by Davies et al. [17,18]. In this method, input shapes are represented by sets of points using conventional algorithms such as "Marching cube" [19]. Then, the mean shape of the input dataset is found and normalized. Input shapes are rigidly aligned to the mean shape. Next, the covariance matrix of the aligned shapes is calculated and PCA is applied on it to find eigenvalues of the shape modes. A cost function is defined based on the eigenvalues of the model (Eq. 1) which tells us how to realign the input shapes to get a compact model. In Eq. (1), ΣL_m is called the "Description Length" and corresponds to the compactness of the model.

$$L_m = 1 + \log(\lambda_m/\lambda_{cut}) \quad \text{for} \quad \lambda_m \geq \lambda_{cut}$$
$$L_m = \lambda_m/\lambda_{cut} \quad \text{for} \quad \lambda_m < \lambda_{cut} \tag{1}$$

In Eq. (1), λ_m is the eigenvalue corresponding to the m-th mode and λ_{cut} is a cut-off value corresponding to noise which is set by the user.

Description of our method

As shown in Fig. 3, our method consists of training and test steps. In the training step, input images are read and converted to meshes. The corresponding points are found using the MDL method. Then, an SSM model is built using all available images. In this paper, we call this model as the general model. The general model is employed to find the shape parameters of the training set which are later used as features to cluster the shapes into N groups. These groups represent different populations. Corresponding to a population, a new SSM model is built. Therefore, we call the proposed model the "Population-Based Statistical Shape Model" (PBSSM) in this paper.

The general model

To build the general model, input CT images of livers are segmented manually by a physician. Preprocessing is an important step when we want to find corresponding points. The segmented livers are processed to remove unwanted small objects produced during segmentation step and to fill the holes inside the objects. Liver masks are represented by meshes using the Marching cube algorithm [19]. To reduce the run-time of the method, output meshes are decimated to include 642 vertices. We used the code presented in

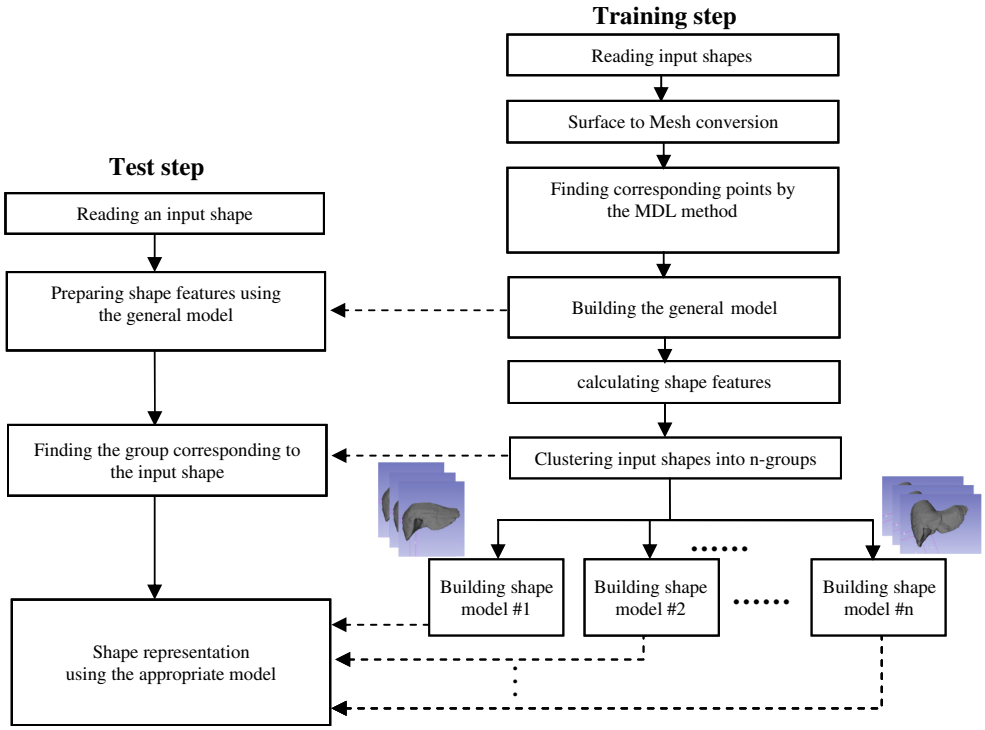

Fig. 3 Flowchart of the proposed method

ITK journal to build the model [20]. To find the correspondence between the vertices of the input meshes, "Minimum Description Length" is employed [17]. The run-time of building one model is up to 1 h. The bottle neck relates to the MDL algorithm which finds corresponding points in a set of training data. In this step of our research, we constructed surfaces using 642 points to assess the effectiveness of the proposed method. In the next steps of our research, we may increase number of points representing a shape to get more accuracy.

Then, the shapes are normalized by the Procrustes method [16]. The PCA is used to find the directions of the shape variations (eigenvectors) together with their scores (eigenvalues). This model is used both in the training and test steps. It is used to cluster the training dataset into different populations and also it is used as shape features to choose the appropriate model for a test data.

Liver shape clustering

To more accurately represent liver shapes, we propose to build separate models for similar liver shapes. Thus, we have to cluster input shapes. In an SSM model, the eigenvectors show the principal directions of the shape points distribution while the eigenvalues show the scores of the corresponding eigenvectors. A shape in this model is represented by Eq. (2) [1].

$$X = \bar{X} + \sum_{m=1}^{M} b_m \varphi_m \tag{2}$$

In Eq. (2), X and $\bar{X}$ are input shape and mean shape vectors, and b_m and φ_m are shape parameters and eigenvectors, respectively. In Eq. (2), it is assumed that the eigenvalues are sorted in descending order and the first M eigenvalues are used to reconstruct a shape. By changing the shape parameters, new shapes, which the current model represents, are reconstructed. Thus, the distance between two shape parameters may be regarded as a metric of the similarity between the two shapes. Using the general model, we employ shape parameters of the input data as a feature vector to cluster the training dataset. In statistical shape models, the cumulative sum of eigenvalues is also used as a metric for the compactness of the model which may be considered as a measure for the quality of correspondence between shape points. Thus, by using shape parameters as the metric for shape similarity, we build more compact model in the second layer of our multi-layer model.

To find the shape parameters of an input data, an iterative approach is employed. The input shape (Y) is first approximated by Eq. (2) in which b_m is initially selected to be a zero vector. Then, the input shape is aligned with the reconstructed shape using the Procrustes method (Eq. 3).

$$y = T^{-1}_{t,\theta,S}(Y) \tag{3}$$

In Eq. (3), t, θ, and S are translation, rotation, and scale parameters, respectively. Next, the aligned shape is projected into the tangent plane to the mean shape $\bar{X}$ (Eq. 4).

$$y_p = y/(y \cdot \bar{X}) \tag{4}$$

Finally, the shape parameters are updated using Eq. (5).

$$b = \varphi^T (y_p - \bar{X}) \tag{5}$$

The updated parameters have to be checked against constraints that bound them by $\pm 3\sqrt{\lambda_i}$ (Eq. 6).

$$|b_i| \leq 3\sqrt{\lambda_i} \tag{6}$$

The stopping criterion can be the maximum number of iterations, the tolerance of the parameter values, or both of them.

In this research, we employed K-means and Fuzzy c-means clustering algorithms to divide input meshes into two groups. Studying the effect of clustering algorithms on the final shape modeling results is left as our future works.

Experiments and results

Database

The input dataset included 29 CT images of the abdominal region. The images had a size of 512 × 512 × 159 and a spacing of 0.625 × 0.625 × 1.25 mm^2. The datasets were stored in DICOM image format with a depth of 12 bits per pixel. Input images were acquired by LightSpeed Ultra GE scanners with eight detectors at Osaka University Hospital. Datasets of the first phase were used. Datasets were acquired from normal cases between 20 and 75 years old including 15 males and 14 females. The patients, who were suspected having diseases other than a liver donor, metastatic liver cancer, or chronic liver disease, were scanned, but the results are normal. Therefore, there was not any information on the age and sex of the patients. The use of the scans is approved by the University Ethics Committee. Liver was segmented in each dataset by a specialized physician. A can be seen in Fig. 1, input images contained a large shape variations which could represent the whole population of liver shapes.

We implemented our method in MATLAB and C++ environments using a MS-Windows-based personal computer (Intel® CoreTM-i7 2670QM 2.2 GHz and 8GB-DRAM). The method was coded both in C++ using ITK [21] and VTK [22] and also in MATLAB. The goal was to exploit the speed of C++ programming and user-friendly developing environment of MATLAB. Visualization of the shapes were performed using 3D slicer [23].

For assessment of our method, we evaluated the compactness, reconstruction accuracy, and generalization ability of the proposed model and compared it to the other researchers' algorithms.

Table 1 Results of clustering input shapes into two groups using K-means and Fuzzy c-means algorithms

Class-01		Class-02	
K-means	Fuzzy c-mean	K-means	Fuzzy c-mean
1	2	7	1
2	4	10	7
4	8	11	10
8	15	12	11
15	16	14	12
16	20	17	14
20	23	18	17
23	24	22	18
24	29	26	22
25	30	27	25
29	31	35	26
30	32	36	27
31	33		35
32	34		36
33	37		
34			
37			

Parameter and feature selection

Regarding the proposed algorithm, shape clustering is considered as a critical step since it influences the overall performance of our method. Since our final goal is to develop an automatic algorithm to segment liver using SSM method, we did not try to divide the input datasets into subgroups. Thus, the shape characteristics were not divided into subgroups related to BMI, gender, or race. As far as we know, healthy livers may have any deliberate shapes. We also did not find any research on the relation of the liver shape and BMI, gender, or race. In this step of our research, there was a set of 29 CT images available. In future, we plan to apply our method on more datasets to more accurately evaluate our method.

To cluster input shapes, we employed K-means and Fuzzy c-means as the clustering methods. There are other advanced clustering techniques such as "Support Vector Machines" (SVM). However, to prove the feasibility of our idea, we employed conventional clustering algorithms. In this research, we used two models in the second layer of our method (Fig. 3). If we set the number of models to three or more, the available datasets would be divided into more groups. Thus, number of training images of a group would fall below 10 shapes which did not suffices to build a valid statistical shape model. In Table 1, the results of clustering input shapes into two groups using K-means and Fuzzy c-means clustering schemes are listed.

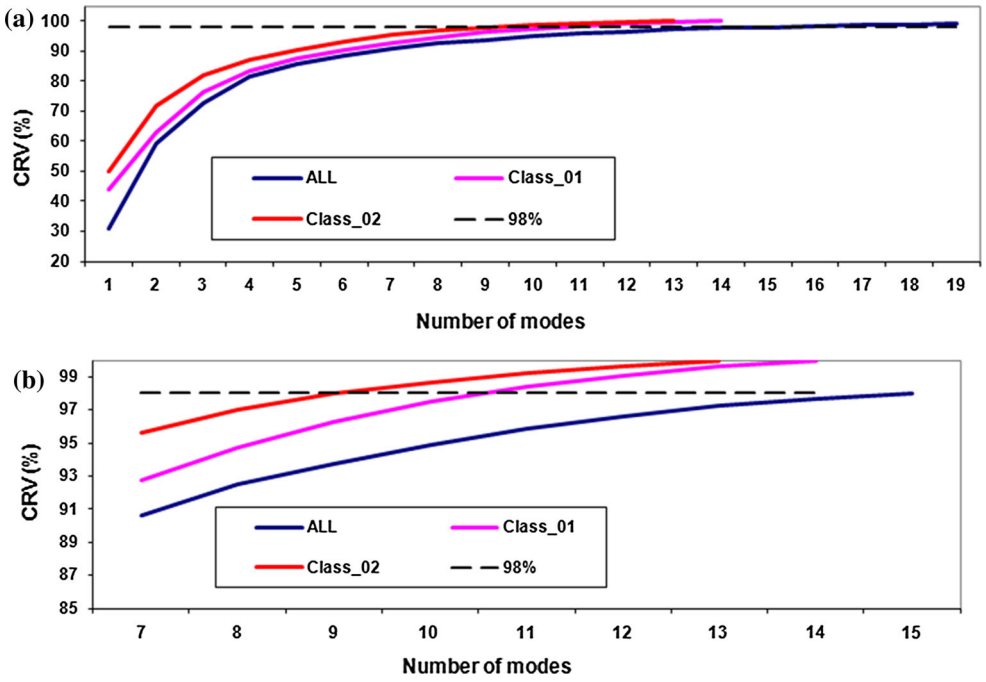

Fig. 4 Comparison of the CVR measures corresponding to the class_01, class_02, mixed, and all data models. **a** Original curve, **b** enlarged curve

Model compactness

The cumulative relative variance (CRV) is the relative cumulative frequency of the eigenvalues sorted in descending order (Eq. 7).

$$\mathrm{CVR}(m) = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{N} \lambda_i}, 1 \leq m \leq N \tag{7}$$

In Eq. (7), N is the total number of modes and CVR(m) gives the CRV up to the m-th mode. The number of eigenvalues required to reach the 98 % of the total of the CRV is used as the conventional measure of the compactness of a model. As the number of training datasets is increased the number of eigenvalues required to reach to the 98 % threshold is converged to a fixed value [8]. To compare the compactness of our model with the conventional SSM, the CRV corresponding to the class_01, class_02, and the general model schemes is plotted in Fig. 4.

The size of the class_01 and class_02 was 15 and 14, respectively. The number of input datasets of the general scheme is nearly twice those of class_01 and class_02. The number of the modes required to reach to 98 % of the total variances was 11, 9, and 15 for the class_01, class_02, and the model consisting of all available images, respectively.

Reconstruction accuracy

We used reconstruction accuracy as another measure to evaluate our model. To compensate for the reduced number of training data in class_01 and class_02, which influence reconstruction accuracy, a new class called the mixed class was prepared to build a new model. The mixed class consisted of 7 images belonging to the class_01 and 7 images belonging to the class_02. To measure the accuracy of a model, we construct a shape in the training data and find the root squared of the Euclidean distance between the reconstructed shape and the input data (Eq. 8).

$$RE = \sqrt{(Y - \bar{Y})^2} \tag{8}$$

In Eq. (8), Y and $\bar{Y}$ are the input and reconstructed shapes, respectively. The reconstruction error of the class_01 and class_02 are shown in Fig. 5. The mean reconstruction error of the PBSSM and mixed models is 0.029 and 0.052, and the maximum reconstruction error of the PBSSM and mixed models is 0.280 and 0.694 mm, respectively.

Generalization

Another measure to evaluate the performance of a statistical shape model is the generalization ability which is defined as how accurately a model describes an input shape that is not already seen in the training set. The mathematical definition of the reconstruction error is given in Eq. (8). We randomly excluded 4 datasets from each group and used them as the test data to evaluate generality of the model. Images no. 4, 16, 29, and 37 were excluded from the class_01 and images no. 11, 18, 26, and 35 were excluded from the class_02. The reconstruction error of these datasets corresponding to the class_01

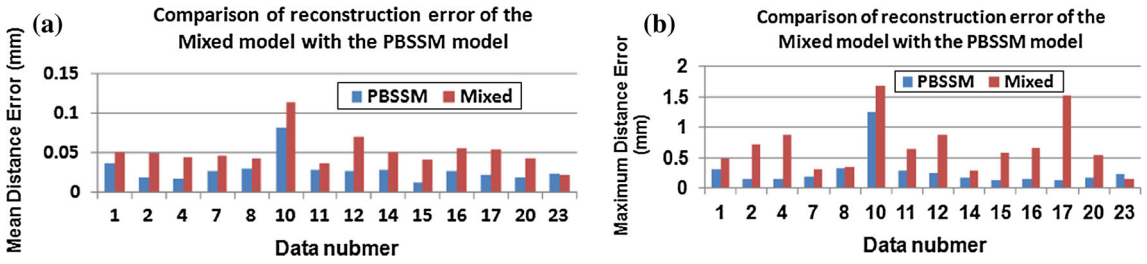

Fig. 5 **a** Mean and **b** maximum reconstruction error of PBSSM and mixed models

Table 2 Reconstruction error (mm) of leave-one-out test corresponding to the class_01 and class_02 and models

Data #	Mean distance error (mm)		Max distance error (mm)	
	Class_01	Class_02	Class_01	Class_02
24	4.03	7.65	19.04	20.38
29	4.4	7.36	17.15	23.03
30	8.22	8.8	35.07	30.98
31	4.24	9.74	18.91	31.77
Mean	5.22	8.39	22.54	26.54

Test data originally belong to class_01

Table 3 Reconstruction error (mm) of leave-one-out test corresponding to the class_01 and class_02 and models

Data #	Mean distance error (mm)		Max distance error (mm)	
	Class_01	Class_02	Class_01	Class_02
18	4.92	5.03	27.21	15.76
22	3.64	6.75	15.48	25.25
25	5.56	7.01	23.57	20.58
26	9.46	5.91	40.56	18.58
Mean	5.9	6.17	26.71	20.04

Test data originally belong to class_02

and class_02 is shown in Tables 2 and 3. As can be seen in Table 2. Figure 5, the mean and maximum reconstruction errors of the class_01 are less than the corresponding error of the class_02 model. However, in Table 3, the mean error of the class_02 is more than the corresponding errors of the class_01 model. The difference may be ascribed to the error in the clustering algorithm. This means that we may assume images no. 22 and 25 belong to the class_01.

In Figs. 6 and 7, variations of the mean shape in the first three dominant modes by $\pm 2\sqrt{\lambda_i}$ are shown for class_01 and class_02, respectively.

In Fig. 8, visualized evaluations of the proposed method are shown. The results in Fig. 8 emphasize the fact that the idea of building population-based shape models enhances reconstruction of unseen data more accurately.

Specificity

Specificity is an important metric to evaluate a statistical shape model. It measures if the model clusters in the vicinity of the training data or not. To measure specificity, M samples shapes from the model pdf are selected $\{y_A\}$. It is defined as in Eq. (9).

$$\hat{S}(n_{\mathrm{m}}) = \frac{1}{M}\sum_{A=1}^{M} \min_{i} \| y_A - x_i \| \tag{9}$$

In Eq. (9), n_{m} is the number of the modes in the model. We measured specificity of the conventional model together with the specificity of the class_01 and class_02 models. The results are 0.0012, 0.00087, and 0.0018 for conventional, class_01, and class_02 models, respectively. The results show that the specificity of the proposed models is comparable to the conventional model.

Synthetic data

For further evaluation of our method, we applied it on synthetic data to assess the potential benefits of the method to build models of tissues with gross shape variations. We prepared 50 ellipsoids and 50 cylinders with different aspect ratios and used them as synthetic data. Typical shapes of the ellipsoids and cylinders are shown in Fig. 9. The hybrid model was built using 40 cylinders and 40 ellipsoids. The remaining shapes were kept for evaluation step. Two submodels were built corresponding to cylinders and ellipsoids. The model was highly dependent on the reference shape which was selected to build it. However, in 55 % of the cases an input shape was correctly assigned to its corresponding submodel.

Comparison with other methods

To evaluate our method more, we compared it with other researches. We compared our method with the conventional MDL method comprising a single class and the methods of Lemecker et al. [8], Heits et al. [24], Feng and Ip [10] and Su et al. [25]. In some of the above researches, the algorithms were applied on rigid objects such as bones or objects with

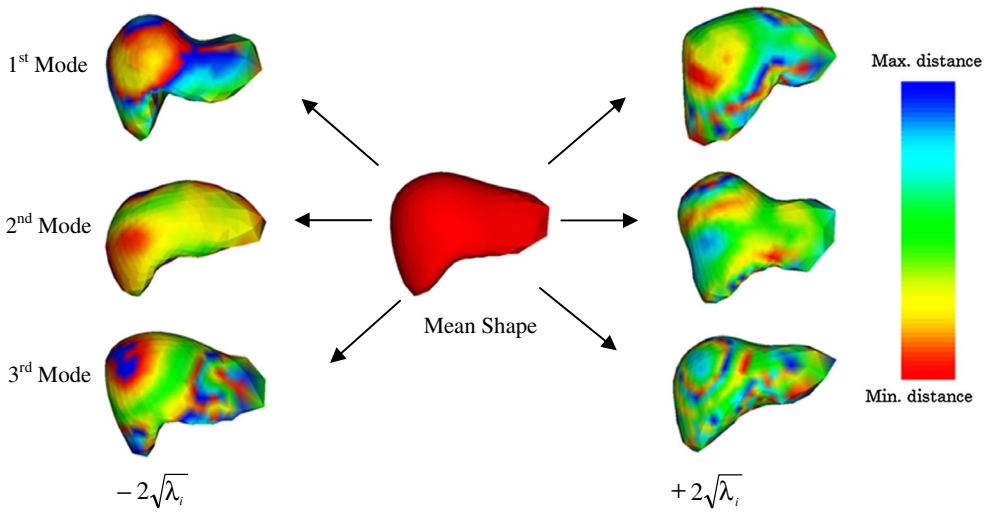

Fig. 6 Variations of the mean shape of class_01 around $\pm 2\sqrt{\lambda_i}$ the first (*top row*), second (*middle row*) and third modes (*bottom row*)

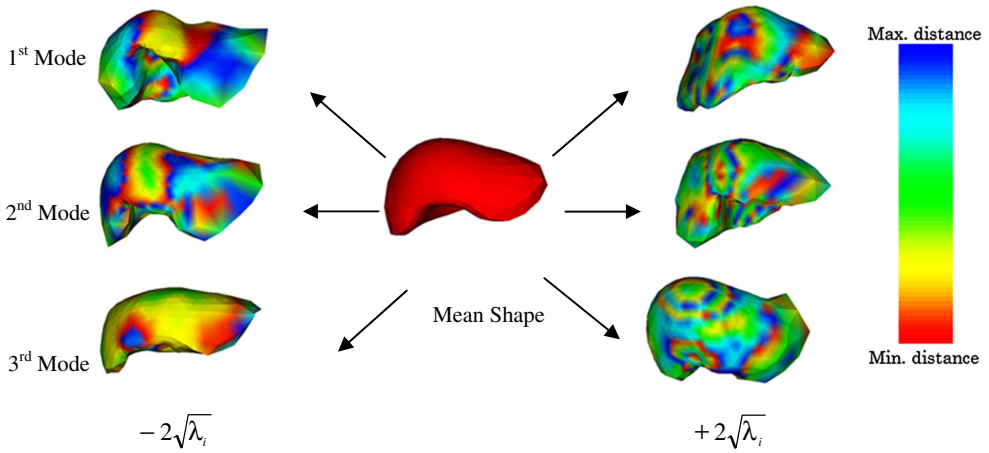

Fig. 7 Variations of the mean shape of class_02 around $\pm 2\sqrt{\lambda_i}$ the first (*top row*), second (*middle row*) and third modes (*bottom row*)

nearly a specific shape like hands. However, we used the results of these methods to compare with our model. Two types of comparisons are done. The first is the compactness of the models.

The next comparison is the generality of the models. The result of mean shape error for leave-one-out tests is shown in Table 4.

Discussion

In this paper, we proposed an automatic multi-layer approach to shape model building. The idea is motivated by the fact that shape models of rigid or nearly rigid shapes such as bones, hands, and synthetic objects are more accurate and these models are used widely in the real-world applications. Capturing shape variations of objects with elastic body needs a large number of training data. If the number of training datasets is increased, the compactness of the model will be reduced. Considering the shape of a soft tissue like liver, we notice that the wide range of liver shapes may be divided into different populations. Thus, we have nearly similar shapes in each group and the corresponding shape models are more accurate, more compact, and more general, compared with a single model comprising all available datasets. Also, finding corresponding points in a population of shapes with nearly similar shape results in better models and the MDL cost function is reduced more.

To evaluate our method, we employed a large number of liver shapes. The large variation of liver images is an important feature in the evaluation of our method. In the methods of Lemecker et al. [8], Heitz et al. [24], Feng and Ip [10], and Su et al. [25], the number of training datasets is 20, 9, 16, and 24, respectively. However, we employed 29 input images for the evaluation of our method.

While we used liver shapes which has a non-rigid shape to build our model, Heitz [24] and Su [25] used their methods

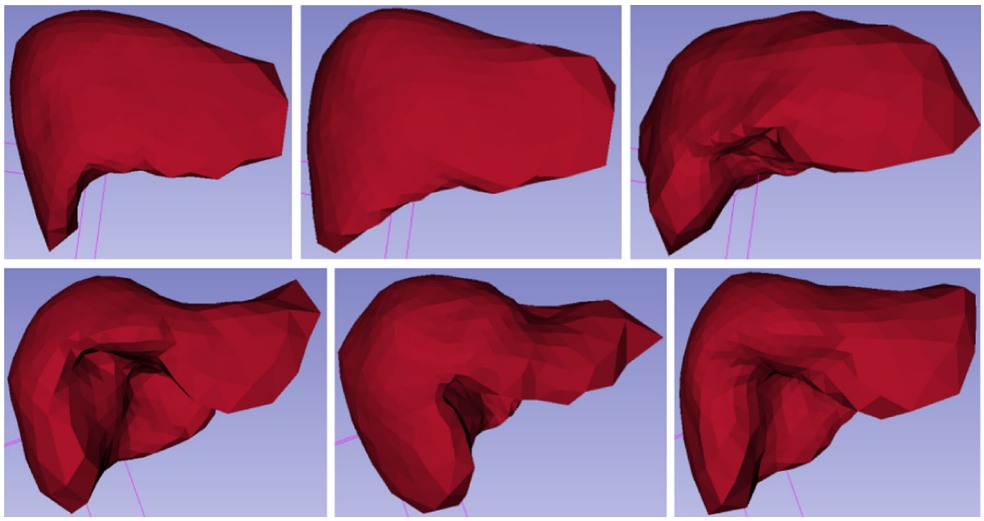

Fig. 8 Visualized evaluations of the proposed method. *Left column*: original shape, *middle column*: reconstructed liver by Group_01 model, *right column*: reconstructed liver by Group_01 model. The liver in the first row is no. 31 and is classified as Group_01 data and The liver in the first row is no. 26 and is classified as Group_02 data

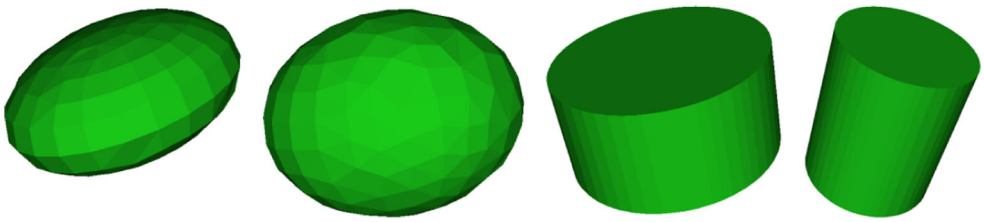

Fig. 9 Typical shapes of ellipsoids and cylinders used as synthetic images

Table 4 Comparison of the generality of the proposed method with other researches

Method	Class_01	Class_02	Single model	Lamecker et al. [8]	Feng and Ip [10]	Heitz et al. [24]	Su et al. [25]
Mean error (mm)	5	6	4	4	7	1	1

to build models for bones and hands. Liver has a soft tissue and its shape can be easily changed when forced is applied on it or the position of the patient is changed during scanning procedure.

Our method uses a general model including all available dataset and employs the corresponding shape parameters to assign a new image to a population. Regarding the clustering methods, we employed two conventional algorithms, i.e., K-means and Fuzzy c-means. In this research, our main focus was to prove the idea that the population-based model is more compact and accurate. In our next step, we will study the influence of the clustering algorithm on the proposed model. Regarding the results of the clustering, both K-means and fuzzy c-means clustered input images into nearly similar groups. The results of clustering by K-means and Fuzzy c-means have an intersection of 88.24 and 85.71 % corresponding to class_01 and class_02, respectively.

The compactness of the models compared in Fig. 4 reveals that the proposed model is more compact than the general model. Thus, we may represent a liver's shape with less modes in our model.

In Table 5, we compared the compactness of our model and the models proposed by other researchers. As shown in Table 5, the compactness of our method is better than the methods of Feng and Heitz and less than the compactness of the models proposed by Lamecker and Su. However, Su built his model using bones which have rigid shapes. In case of the Lamecker's method, the number of the training dataset is more than our method. In our next step of research, we will increase the number of the training images so that about 20 data are assigned to each group. Then, we can compare compactness of our method to the method of Lamecker. Considering the compactness of the general model and number of the input images in several researches, we believe that

Table 5 Comparison of the compactness of the proposed model and models by other researches

Method	Class_01	Class_02	Lamecker et al. [8]	Feng and Ip [10]	Heitz et al. [24]	Su et al. [25]
Number of prominent modes	11	9	12	7	12	14
Number of training data	15	14	20	9	16	24
No. of prominent modes to no. of training data (%)	73	64	60	78	75	58

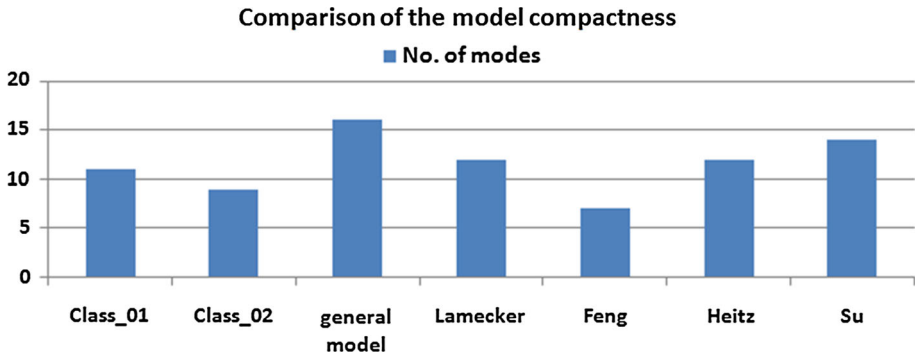

Fig. 10 Comparison of the compactness of the proposed method with other methods

increasing number of training data for class_01 and class_02 will improve their compactness.

The number of training images used to build a shape model influences reconstruction error. Thus, we compared the reconstruction error of our model with the error of a model comprising equal number of input images from class_01 and class_02. We called this group as mixed model in this paper. Reconstruction error of our PBSSM model compared with that of the mixed model (shown in Fig. 5) reveals that classification of the input data into different populations leads to better models and the training images are reconstructed more accurately using their corresponding models. The results shown in Fig. 5 prove the idea that building different shape models based on populations give us more accurate models than using a single model. The mean reconstruction error of the mixed model is about twice that of our model and the maximum error of the mixed model is more than twice of our model.

Generality of the model is an important metric for the evaluation of a shape model. In the generality test, an unvisited data are given to a model and the model tries to reconstruct it. More accurate reconstruction of the input image proves that the mode is more general and may be used to represent more new data. The results shown in Tables 2 and 3 show that unvisited data may be represented more accurately by the population model which they belong to.

In Fig. 10, we compared our method with other researches using compactness measure. The best compactness result belonged to Feng's method [10]. Our method stood in the second rank. While we employed 29 datasets in our model, Feng et al. used only nine CT volumes to construct their model [10]. Su et al. [25] and Heitz et al. [24] employed 32 and 11 training data to model bones which were considered as rigid tissues. Thus, their model did not deal with large shape variations. The method proposed by Lamecker et al. [8] used 43 liver images to build their model. They did not report about large shape variations of the input images.

In Table 4, we compared our method with other researches using generality measure. The mean error of our method was larger than the mean error of Heitz and Su. It may be explained because we modeled a soft tissue with large shape variations while they modeled rigid bones. However, the mean error of the Lamecker's method was less than ours. It may be ascribed to more accurate point correspondence algorithm which was employed by Lamecker et al.

Regarding the result of applying our method to synthetic shapes, we can conclude that the idea of population-based SSM cannot be used to model two different anatomical tissues. However, the method is beneficial to model a single tissue with large shape variations. A conventional SSM built by a large number of training shapes neglects local details. However, increasing number of training data leads to more accurate classifications in case of our method.

Conclusion and future works

In this paper, we proposed a population-based method to model the shapes of liver which is a soft tissue. The main idea of our method is that for shapes of large variation, using a multi-model approach may lead to more accurate and compact models. Our results proved this idea.

In future, we extend the input images of our method and evaluate our method on more data. Inclusion of more images into the training dataset and dividing the liver volumes into segments (anatomical segments of the liver) are considered as the next steps of our research. It may be used to find the corre-

sponding points more accurately and to build the shape model more precisely. Also, we decide to incorporate inter-organ relationship into building statistical shape models to include both the inter-patient differences by using population-based models and to consider the relationship between organs to build more accurate models.

Conflict of interest Amir H. Foruzan, Yen-Wei Chen, Masatoshi Hori, Yoshinobu Sato and Noriyuki Tomiyama declare that they have no conflict of interest.

References

1. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59
2. Heimann T, Meinzer HP (2009) Statistical shape models for 3D medical image segmentation: a review. Med Image Anal 13(4):543–563
3. Pepe, A, Zhao L, Koikkalainen J, Hietala J, Ruotsalainen U, Tohka J (2013) Automatic statistical shape analysis of cerebral asymmetry in 3D T1-weighted magnetic resonance images at vertex-level: application to neuroleptic-naïve schizophrenia. Magn Reson Imaging 31(5):676–687. doi:10.1016/j.mri.2012.10.021
4. van de Giessen M, Foumani M, Streekstra GJ, Strackee SD, Maas M, van Vliet LJ, Grimbergen KA, Vos FM (2010) Statistical descriptions of scaphoid and lunate bone shapes. J Biomech 43(8):1463–1469
5. Buchaillard SI, Ong SH, Payan Y, Foong K (2007) 3D statistical models for tooth surface reconstruction. Comput Biol Med 37(10):1461–1471
6. Chen YW, Luo J, Dong C, Han X, Tateyama T, Furukawa A, Kanasaki S (2013) Computer-aided diagnosis and quantification of cirrhotic livers based on morphological analysis and machine learning. Comput Math Methods Med 2013:264809. doi:10.1155/2013/264809
7. Cootes TF, Taylor CJ (1995) Combining point distribution models with shape models based on finite element analysis. Image Vis Comput 13(5):403–409
8. Lamecker H, Lange T, Seebass M (2002) A statistical shape model for the liver. In: Medical image computing and computer-assisted intervention—MICCAI 2002. Springer, Berlin, pp 421–427
9. Okada T, Linguraru MG, Yoshida Y, Hori M, Summers RM, Chen Y.-W, Tomiyama N, Sato Y (2012) Abdominal multi-organ segmentation of CT images based on hierarchical spatial modeling of organ interrelations. In: Abdominal imaging, computational and clinical applications. Springer, Berlin pp 173–180
10. Feng J, Ip HHS (2009) A multi-resolution statistical deformable model (MISTO) for soft-tissue organ reconstruction. Pattern Recognit 42(7):1543–1558
11. Davatzikos C, Tao X, Shen D (2003) Hierarchical active shape models, using the wavelet transform. Med Imaging IEEE Trans 22(3):414–423
12. Zhang W, Yan P, Li X (2011) Estimating patient-specific shape prior for medical image segmentation. In: Biomedical imaging: from nano to macro, 2011 IEEE international symposium on. IEEE 2011, pp 1451–1454
13. Zhu Y, Papademetris X, Sinusas AJ et al (2010) Segmentation of the left ventricle from cardiac MR images using a subject-specific dynamical model. Med Imaging IEEE Trans 29(3):669–687
14. Zhang S, Zhan Y, Dewan M et al (2012) Towards robust and effective shape modeling: sparse shape composition. Med Image Anal 16(1):265–277
15. Wang G, Zhang S, Li F et al (2013) A new segmentation framework based on sparse shape composition in liver surgery planning system. Med Phys 40:051913
16. Goodall C (1991) Procrustes methods in the statistical analysis of shape. J R Stat Soc Ser B (Methodol), 285–339
17. Davies RH, Twining CJ, Cootes TF, Waterton JC, Taylor CJ (2002) A minimum description length approach to statistical shape modeling. Med Imaging IEEE Trans 21(5):525–537
18. Styner MA, Rajamani KT, Nolte L.-P, Zsemlye G, Székely G, Taylor CJ, Davies RH (2003) Evaluation of 3D correspondence methods for model building. In: Information processing in medical imaging. Springer, Berlin, pp 63–75
19. Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. In: ACM Siggraph computer graphics, vol 21, no 4. ACM, pp 163–169
20. Heimann T, Oguz I, Wolf I, Styner M, Meinzer H.-P (2006) Implementing the automatic generation of 3d statistical shape models with ITK. In: Open science workshop at MICCAI, Copenhagen
21. Insight toolkit. http://www.itk.org. Accessed on 6 July 2011 (2013)
22. Visualization toolkit. http://www.vtk.org, Accessed on 6 July 2011 (2013)
23. 3D Slicer Software. http://www.slicer.org, Accessed on 6 July 2011 (2013)
24. Heitz G, Rohlfing T, Maurer Jr CR (2005) Statistical shape model generation using nonrigid deformation of a template mesh. In: Medical imaging. International society for optics and photonics, pp 1411–1421
25. Su Z, Lambrou T, Todd-Pokropek A (2006) A minimum entropy approach for automatic statistical model building. In: (Proceedings) IEEE 5th international special topic conference on information technology in biomedicine