Medical Image Analysis 16 (2012) 265–277

Contents lists available at SciVerse ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

# Towards robust and effective shape modeling: Sparse shape composition

Shaoting Zhang [a,b,*], Yiqiang Zhan [a], Maneesh Dewan [a], Junzhou Huang [c], Dimitris N. Metaxas [b], Xiang Sean Zhou [a]

[a] CAD R&D, Siemens Healthcare, Malvern, PA, USA
[b] Department of Computer Science, Rutgers University, Piscataway, NJ, USA
[c] Department of Computer Science & Engineering, University of Texas at Arlington, TX, USA

ARTICLE INFO

*Article history:*
Received 21 February 2011
Received in revised form 22 August 2011
Accepted 22 August 2011
Available online 5 September 2011

*Keywords:*
Shape prior
Medical image segmentation
Sparse learning
Lung localization
Liver segmentation

ABSTRACT

Organ shape plays an important role in various clinical practices, e.g., diagnosis, surgical planning and treatment evaluation. It is usually derived from low level appearance cues in medical images. However, due to diseases and imaging artifacts, low level appearance cues might be weak or misleading. In this situation, shape priors become critical to infer and refine the shape derived by image appearances. Effective modeling of shape priors is challenging because: (1) shape variation is complex and cannot always be modeled by a parametric probability distribution; (2) a shape instance derived from image appearance cues (input shape) may have gross errors; and (3) local details of the input shape are difficult to preserve if they are not statistically significant in the training data. In this paper we propose a novel Sparse Shape Composition model (SSC) to deal with these three challenges in a unified framework. In our method, a sparse set of shapes in the shape repository is selected and composed together to infer/refine an input shape. The a priori information is thus implicitly incorporated on-the-fly. Our model leverages two sparsity observations of the input shape instance: (1) the input shape can be approximately represented by a sparse linear combination of shapes in the shape repository; (2) parts of the input shape may contain gross errors but such errors are sparse. Our model is formulated as a sparse learning problem. Using $L1$ norm relaxation, it can be solved by an efficient expectation–maximization (EM) type of framework. Our method is extensively validated on two medical applications, 2D lung localization in X-ray images and 3D liver segmentation in low-dose CT scans. Compared to state-of-the-art methods, our model exhibits better performance in both studies.



## 1. Introduction

Shape is a distinctive characteristic of many human organs and plays a critical role in various medical image analysis tasks. Although image appearance cues provide low-level evidence to derive organ shapes, the derived shape instances may be incomplete in the presence of weak (missing) appearance cues, or incorrect when misleading appearance cues are present. Fig. 1 shows a 3D image data from whole body CT with low dose and large slice thickness, which result in low contrast and fuzzy boundaries between organs. The boundary information is weak in between the liver and the kidney (Fig. 1, Left). Furthermore, there is motion artifact induced by breath around the boundary between the spleen and the lung (Fig. 1, Right). Fig. 2 shows some challenging cases of another imaging modality, chest X-ray. In the marked regions, lung boundaries become broken due to the medical instruments. In these scenarios, segmentation methods solely relying on appearance cue may fail because of the missing or "false" boundaries. On the other hand, however, human organs usually contain strong shape priors. It provides opportunities to use shape models to infer and refine the organ shape in an optimal sense based on high-level shape priors. The success of these models is highly dependent on the way shape priors are modeled and on the optimization method used. One of the seminal work in this area, "Snake" (Kass et al., 1987), models the shape prior as a general regularity term in the optimization, which assumes that the shape should deform like a membrane or a thin plate. Subsequently, more object-specific shape priors became prevalent, where the shape priors are learned from a set of training samples, such as Active Shape Model (ASM) (Cootes et al., 1995) and Level Set representation with priors (Rousson and Paragios, 2002). Many adaptations of these algorithms have been proposed over the years and some of them are successfully applied in tasks of medical image segmentation (see Section 2 for details). These methods often confront the following three challenges. First, shape variations are usually complex and therefore difficult to model using a parametric probability distribution. Second, image appearance information can be highly misleading and non-Gaussian errors frequently appear in the input shape. Shape models have to be robust to handle these errors. Third, shape

* Corresponding author at: Department of Computer Science, Rutgers University, Piscataway, NJ, USA.
*E-mail address:* shaoting@cs.rutgers.edu (S. Zhang).

1361-8415/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2011.08.004

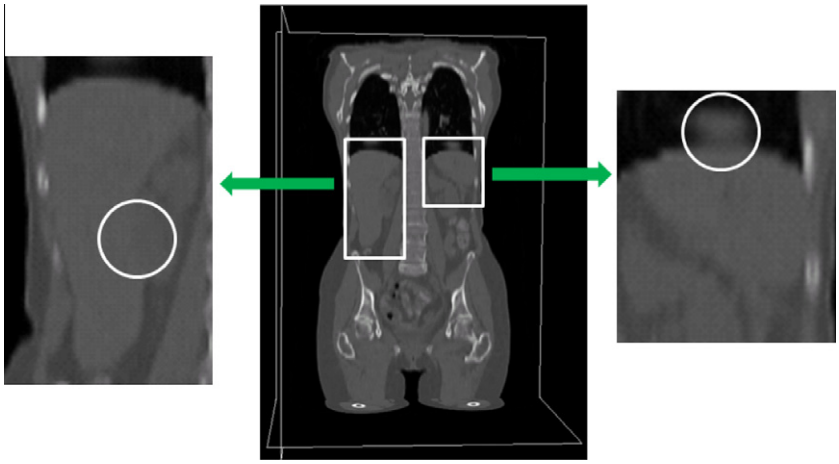


**Fig. 1.** Middle: whole body low-dose CT data. Left: zoom in of the liver and lung. In the marked region the boundary between the liver and the kidney can hardly observed. The appearance cue is weak because of the low contrast around the boundary. Right: zoom in of the spleen and lung. In the marked region, there is artifact induced by breath. It is part of the lung but not the spleen. Since the image information is misleading here, segmentation methods solely relying on appearance cue may accidentally include this region as spleen.

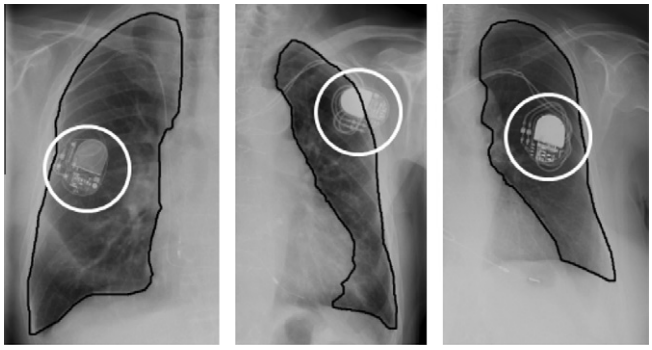


**Fig. 2.** Chest X-ray images with annotated boundaries. The appearance cue is misleading because of the instruments in the marked regions.

models should be adaptive in order to preserve local detail information in the input shape, provided such details are present in the training data, even if they are not statistically significant. While several methods have been proposed to address one or two of those challenges, it remains that, to the best of our knowledge, none of them can tackle all three challenges simultaneously.

In this paper, we propose a novel Sparse Shape Composition model (SSC) to address the aforementioned challenges in a unified framework. Given a shape repository consists of a number of annotated shape instances, instead of explicitly learning shape priors from these shapes offline, we propose to adaptively approximate the input shape on-the-fly, by a sparse linear combination of a subset of shapes in the shape repository. Hence, the shape prior constraint is implicitly applied. Our method is inspired by recently proposed sparsity theories in the compressive sensing community, i.e., the problem of computing sparse linear representations with respect to an overcomplete dictionary of base elements (Candes and Tao, 2006; Donoho, 2004). It has been successfully applied in many computer vision applications, such as, but not limited to, face recognition (Wright et al., 2009), image restoration (Mairal et al., 2009), background subtraction (Huang et al., 2009) and image annotation (Zhang et al., 2010). Yet, to our knowledge, such techniques have not been used in the context of shape priors.

There are two "sparsity" observations behind our method. First, given a large enough training dataset (repository), an instance can be approximately represented by a sparse linear combination of instances in the shape repository. Similarly, in our application each given shape is approximated by a sparse linear combination of annotated shapes. Without any assumption of a parametric distribution model (e.g., a unimodal distribution assumption in ASM), it becomes general to objects whose shape statistics can be very complex. Moreover, such a setting is able to recover detail information even if the detail of the input shape is only present in a small proportion of the shape repository and is not statistically significant. Second, the given shape information may contain gross errors, but such errors are often very sparse, e.g., there is an object occluded in the image or a point missing in the input shape. Combining these two, we formulate the shape prior task as a sparse learning problem, and efficiently solve it by an expectation–maximization (EM) type of framework. Furthermore, we explicitly model the nonlinear shape transformation in the optimization framework without assuming the misalignment is small, which is different from the sparse learning method in face recognition (Wright et al., 2009). In our experiments, we validate our method on two applications of medical image analysis, and the proposed SSC shows improved accuracy and robustness compared to some widely used approaches.

The main contributions of our work are threefold: (1) SSC is proposed to model shapes and implicitly incorporate the shape prior constraint effectively. It is based on sparse representations and our unified framework is able to handle non-Gaussian errors, multimodal distributions of shapes and detail information recovery. (2) The problem is efficiently solved by an EM type of framework. (3) It is successfully applied to two medical applications, 2D lung localization from X-ray image and 3D liver segmentation from low-dose CT. The extensive experiments demonstrate the superior performance of our method.

## 2. Relevant work

As discussed earlier, many approaches have been proposed in different contexts to incorporate the shape prior constraint. In the context of medical image analysis, ASM (Cootes et al., 1995) and its variations (Heimann and Meinzer, 2009) are probably the most widely used approaches. Many methods have been proposed to improve the shape prior module of ASM. They mainly focus on three aspects:

*Modeling complex shape variations:* A significant effort has been put on handling multimodal distribution of shapes, which cannot be represented by their mean shape and variations. A classical solution is to use a mixture of Gaussians to represent shape variation (Cootes and Taylor, 1997). Manifold learning techniques (Etyngier et al., 2007) can also be used to learn a non-linear shape

prior to alleviate this problem. For example, Zhang et al. (2011b) employed manifold learning method to overcome the limitation of ASM on statistical constraint. However, it is still possible that shape variation is too complex to model with a parametric probability distribution. Therefore, some researches proposed to decompose the shape space into multiple sub-spaces. Representative studies include patient-specific shape statistics (Shi et al., 2008; Yan and Kruecker, 2010) or subject-specific dynamical model (Zhu et al., 2008a; Zhu et al., 2008b; Zhu et al., 2009; Zhu et al., 2010) to constrain the deformable contours. Since shape distributions in these sub-spaces are often more compact, a particular shape might be better approximated by mean shape and its variations in a sub-space. It is also worth mentioning that there are some successful work in the face alignment field that attacks similar problems, such as the multi-level generative model (Gu and Kanade, 2008) and subspace constrained mean-shift (Saragih et al., 2009). Shape inference is also a potential solution. It constructs a surface from a set of 3D points. In (Georgescu et al., 2005), to infer the shape, a nearest-neighbor approach is used by finding the closest instance in a database, and the database is based on the expert's structure annotations.

*Handling non-Gaussian errors:* A large number of proposed modifications on the original ASM algorithm tries to improve the robustness against outliers (erroneous landmarks/boundaries). Duta and Sonka (1998) propose detecting and correcting outliers by using the variance information from the PDM. If a point is considered an outlier, it is corrected based on the position of its neighbors. Lekadir et al. (2007) employ a shape metric based on the ratio of landmark distances to detect outliers. Other methods try to decrease outliers' influence using the weighting of residuals. Rogers and Graham (2002) evaluate the use of M-estimators, image match and random sample consensus (RANSAC) (Fischler and Bolles, 1981) for this purpose. In a concluding evaluation, RANSAC was the most effective of these three methods. Nahed et al. (2006) proposed to use a robust point matching algorithm (Chui and Rangarajan, 2003) which rejects outliers and finds the best-fitting model. Missing landmarks/boundaries is a special case of outliers. Yan et al. (2010) tried to use partial ASM to address this problem of missing boundary in image segmentation.

*Preserving local detail information:* Another difficulty is to preserve local details of the input shape when such details are also present in the training data but not statistically significant. PCA performs eigen-analysis and extracts eigenvectors with the largest eigenvalues. The discarded eigenvectors are statistically insignificant, but they may contain important local details. Some relevant work can alleviate this problem. Sparse PCA (Sjostrand et al., 2007) obtains sparser modes and produces near-orthogonal components. Thus each mode only affects locally clustered landmarks and captures more detail information. Some other methods divide the shape model into several independently modeled parts, such as the hierarchical approach (Davatzikos et al., 2003). Since the smaller parts exhibit less variation, they can be captured with fewer training samples than the variations for the full shape.

However, most discussed methods focus on solving one or two limitations. It is not trivial to handle all of them simultaneously. In our work, we address these challenges in a unified framework as outlined in the next sections.

## 3. Sparse learning based shape prior modeling

### 3.1. Problem formulation

In this study, we aims to model the shape of an object using a set of existing training shape instances.

*Notations and basic framework:* Explicit parametric shape representation is employed to model a shape instance, i.e., a curve (2D) or a triangular mesh (3D) consisting of a set of vertices. To describe the $i$th shape in the training data, the coordinates of all its vertices are concatenated into a vector $d_i \in \mathbb{R}^n$, where $n$ is the product of the number of vertices in each shape by the dimension. Thus the training repository can be represented as a matrix $D = [d_1, d_2, \ldots, d_k] \in \mathbb{R}^{n\times k}$, where $k$ is the number of shapes. In our framework, all $d_i$, $i = 1,2,3,\ldots,k$ are pre-aligned using generalized Procrustes analysis (Goodall, 1991). $y \in \mathbb{R}^n$ is the vector of a newly-input shape which needs to be constrained or refined. Our basic framework assumes that after proper alignment, any input shape $y$ can be approximately represented as a weighted linear combination of existing data $d_i$, $i = 1,2,3,\ldots,k$, and the parts which cannot be approximated are noises. We denote $x = [x_1, x_2, \ldots, x_k]^T \in \mathbb{R}^k$ as the coefficients or weights. Thus the value of $x$ for the linear combination is found by minimizing the following loss function:

$$\arg\min_{x,\beta} \ \|T(y,\beta) - Dx\|_2^2, \tag{1}$$

where $T(y,\beta)$ is a global transformation operator with parameter $\beta$. It aligns the input shape $y$ to the mean shape of existing data $D$. $x$ and $\beta$ are computed by solving (1). Thus the input shape $y$ is constrained or refined as $Dx$ and transformed back by the inverse of the transformation matrix using parameter $\beta$.

*Sparse linear combination:* The limitations of (1) are twofold. First the data matrix $D$ may be overcomplete ($k > n$) when the number of shapes is larger than the length of $d_i$. Thus the system may not have a unique solution. More constraints of the coefficient $x$ are needed. Second, the input shape, including the noises, may be perfectly represented if any linear combination can be used. A more appropriate assumption is that the input shape can be approximately represented by a *sparse* linear combination of existing data. This way, the problem is reformulated as:

$$\arg\min_{x,\beta} \ \|T(y,\beta) - Dx\|_2^2, \tag{2}$$
$$s.t. \ \|x\|_0 \leqslant k_1$$

where $\|\cdot\|_0$ is the $L^0$ norm counting the nonzero entries of a vector, $k_1$ is the pre-defined sparsity number. Such formulation ensures that the number of nonzero elements in $x$ is smaller than $k_1$. The value of $k_1$ depends on specific applications, and is discussed in Section 5.

*Non-Gaussian errors:* The formulation (2) works well for many scenarios. However, there is still one limitation in (2). Since the loss function is based on $L2$ norm, it assumes that the error model follows a Gaussian distribution. Thus it is sensitive to large noises or gross errors of the input shape, caused by image occlusion or points missing. Such problem happens frequently in many applications. In these cases, some errors can be very large, but they are relatively sparse compared to the whole data. To alleviate this problem, we explicitly model the error as a sparse vector $e \in \mathbb{R}^n$ by reformulating the problem as:

$$\arg\min_{x,e,\beta} \ \|T(y,\beta) - Dx - e\|_2^2, \tag{3}$$
$$s.t. \ \|x\|_0 \leqslant k_1, \|e\|_0 \leqslant k_2$$

where $k_2$ is the sparsity number of $e$. When solving (3), $e$ captures sparse but large errors which are caused by occlusion or point missing. When there is no such error, the $L2$ norm loss function can deal with it well and $e$ will be all zeros. Thus $e$ is a good supplement which specifically handles non-Gaussian and sparse errors. Note that unlike the formulation in the robust face recognition (Wright et al., 2009), we do not assume that the misalignment is small and thus explicitly model the transformation with parameter $\beta$ in (3).

*Convex relaxation:* The constraints in (3) are not directly tractable because of the nonconvexity of $L^0$ norm. Greedy algorithms can be applied to this NP-hard $L^0$ norm minimization problem, but

there is no guarantee to capture the global minima. In the general case, no known procedure can correctly find the sparsest solution more efficiently than exhausting all subsets of the entries for $x$ and $e$. Furthermore, in practice the sparsity numbers $k_1$ and $k_2$ may change for different data in the same application. For example, some data have errors while others do not. Fortunately, recent developments in sparse representation provide a theorem to efficiently solve this kind of problems through $L1$ norm relaxation (Starck et al., 2004). Thus (2) is reformulated as:

$$\arg\min_{x,\beta} \|T(y,\beta) - Dx\|_2^2 + \lambda_1\|x\|_1, \tag{4}$$

which is denoted as SSC (4), and is evaluated in the experiments section. Similarly, (3) is reformulated as:

$$\arg\min_{x,e,\beta} \|T(y,\beta) - Dx - e\|_2^2 + \lambda_1\|x\|_1 + \lambda_2\|e\|_1, \tag{5}$$

where $\lambda_1$ and $\lambda_2$ control how sparse $x$ and $e$ are, respectively. After relaxation, $\lambda_1\|x\|_1 + \lambda_2\|e\|_1$ is non-smooth but continuous and convex. (5) is our objective function of SSC (Zhang et al., 2011a). The deviation from (3)–(5) relaxes the absolute sparseness constraints of the objective function ($L0$ norm to $L1$ norm). From the shape modeling perspective, we might use more shape instances for shape composition by optimizing (5). However, since this deviation converts a NP hard problem to a continuous and convex optimization problem which can be solved efficiently, it paves the way for a feasible shape composition procedure as described in Section 3.2.

Although our focus is on the shape prior modeling instead of sparse learning methods, it is still worth mentioning that many other methods can also achieve sparsity, such as Bayesian variable selection (Dellaportas et al., 2002; Kuo and Mallick, 1998; George and McCulloch, 1993). In our model, we choose $L1$ norm based sparse representation because it is a convex optimization problem, which can be effectively solved by many convex techniques and solvers. Furthermore, Donoho (2004) provides theoretical proofs that the $L1$ relaxation can preserve the sparsity property of $L0$ norm constraint.

*Connections to other methods:* It is interesting to look into (5) by adjusting $\lambda_1$ and $\lambda_2$ into some extreme values.

- If $\lambda_2$ is extremely large, $e$ will be all zeros. Thus SSC is similar to methods which do not model non-Gaussian errors.
- If both $\lambda_1$ and $\lambda_2$ are large enough, $e$ will be all zeros and $x$ may have only one nonzero element. Thus SSC becomes the nearest neighbor method.
- If $\lambda_2$ is extremely large and $\lambda_1$ is small, a dense linear combination of shapes is used, which is able to perfectly approximate the transformed input shape. Thus SSC degenerates to the Procrustes analysis.

The insight of (5) indeed reveals the connections of our SSC with some other popular methods. Those methods can be regarded as special cases of SSC. In other words, SSC provides a unified framework to deal with different challenges of shape prior modeling simultaneously. SSC can also provide flexibility to meet the requirements of different applications by adjusting the sparsity of $x$ and $e$.

*Parameter settings:* Eq. (5) has two user tunable parameters $\lambda_1$ and $\lambda_2$, which are usually are crucial to the performance and convergence. From a practical point of view, it is desirable that the parameters are easy to tune and not sensitive to different data in one application. If these parameters have physical meanings, it is straightforward to adjust them. Fortunately, the parameters of our algorithm also have such a property. $\lambda_1$ controls the sparsity of $x$. The length of vector $x$ is equal to the number of shapes in the shape repository. It is usually larger than 100. To generate a sparse coefficient $x$, a large $\lambda_1$ is necessary. $\lambda_2$ controls the sparsity of $e$. The length of vector $e$ is equal to the number of vertices (multiplied by the dimension), which ranges from around 20 to 2000. $e$ should not be too sparse, otherwise it cannot capture any errors. Thus $\lambda_2$ should be relatively small. Both parameters are straightforward to tune given their meanings. Furthermore, the experiments in Section 5 show that the same group of parameters works well for all data in one application.

### 3.2. Optimization framework

To solve (5), we need to simultaneously recover the alignment parameter $\beta$ and error $e$. It is a typical Chicken-and-Egg problem. Furthermore, to efficiently optimize (5), we need to deal with the nonlinearity of $T(y,\beta)$ if the transformation is rigid or a similarity. A notable approach is to use iterative linearization and optimize all variables simultaneously, which was proposed and successfully applied in image alignment (Peng et al., 2010). However, this algorithm assumes that the initial misalignment is not too large, which may not be held in our problem. Furthermore, it focuses on rigid transformation in 2D images, while we deal with nonrigid transformation in arbitrary dimensions for shapes. The efficiency of the optimization framework is also important.

Our solution is to use EM types of algorithms (or alternating minimization). (5) is divided into two sub-problems: (1) estimating $\beta$ and computing $T(y,\beta)$, (2) efficiently minimizing this simplified linear inverse problem. In the "E" step, $\beta$ is estimated using Procrustes analysis, which aligns the shape $y$ to the mean shape. Then vector $y' = T(y,\beta)$ is obtained. The initial value of beta is obtained by registering the average shape of the database to the subject space. In the "M" step, the following simplified problem is minimized:

$$\arg\min_{x,e} \|y' - Dx - e\|_2^2 + \lambda_1\|x\|_1 + \lambda_2\|e\|_1, \tag{6}$$

which is now a linear inverse problem. It is then efficiently solved using the Fast Iterative Shrinkage Thresholding Algorithm (FISTA) (Beck and Teboulle, 2009). Two procedures are iteratively employed to obtain $x$, $e$ and $\beta$. Then $Dx$ is computed as the approximated shape and is transformed to its original coordinate system. The framework is detailed in Table 1. Note that theocratically this EM algorithm might lead to local minima. However, in our extensive experiments (Section 5.1), we did not observe this situation yet. We also observe that our results are quite stable with respect to the local random perturbations of subject space (e.g., auto-detected landmarks). This proves that slight differences of beta do not affect the final results.

## 4. Shape inference and refinement for organ localization and segmentation

Due to imaging artifacts and diseases, appearance cues in medical images might be unreliable or misleading. On the other hand, however, strong shape priors of human anatomy provide opportunities to shape prior-based methods. To evaluate the capability of the proposed shape prior modeling, we apply it to two tasks: (1) organ localization using shape inference and landmark detection, and (2) segmentation using shape refinement and deformable model. The first task organ localization can be used as an initialization step of the segmentation framework (Fig. 3).

*Organ localization based on landmark detection and shape inference:* The positions and orientations of the same organ vary significantly in medical image data. Quickly and accurately locating the organ is crucial to image segmentation. One approach is to find a similarity transformation matrix, and then use this matrix to align a mean shape to the organ (Zheng et al., 2008). Generally it achieves good performance. However, similarity transform only has nine degrees of freedom. Thus it may not be able to represent

**Table 1**
Optimization framework to solve (5).

**Algorithm 1.** Optimization

**Input:** Data matrix of shape repository $D \in \mathbb{R}^{n\times k}$, where $k$ is the number of shapes, and each column is a training shape $d_i \in \mathbb{R}^n$
**repeat**
"E" step: $\beta$ is estimated using Procrustes analysis, which is a similarity transformation and aligns the input shape $y$ to the mean shape of $D$. $y' = T(y,\beta)$
"M" step: (6) is efficiently minimized use FISTA Beck and Teboulle, 2009. $x$ and $e$ are computed
**until** halting criterion true
Compute $y_{refined} = Dx$
Compute $y'_{refined} = T(y_{refined}, \gamma)$, where $\gamma$ is the parameter corresponding to the inverse of the transformation matrix using $\beta$
**Output:** $y'_{refined}$

some specific data or shapes by transforming a mean shape (in Section 5.1).

To solve this problem, we propose a landmark detection and shape inference based localization method. A learning-based method is employed for landmark detection (Zhan et al., 2008). Detected landmarks can be very sparse compared to the whole shape. Furthermore, there may be gross errors or point missing from the detection results. We use SSC to infer a shape based on these detected landmarks. Compared to solely using similarity transformation to transform the mean shape, our approach has more degrees of freedom and is able to better fit to the actual shape in the image.

In this framework, we assume that ground truths for shapes are available from training data, and the one-to-one correspondence is already obtained. We then automatically or manually choose some specific points (e.g., corner points or high curvature points) as landmarks on the shape of each data. Such training landmarks and shapes are fed into data matrices denoted as $D_L$ and $D_S$, respectively. Given a testing image, its landmarks $y_L$ are detected using a learning based method (Zhan et al., 2008). Then $x$ and $\beta$ is computed by optimizing (5) with $D_L$ and $y_L$. Finally $D_S x$ is computed as the refined shape and transformed back to its coordinate system using inverse of the transformation matrix with parameter $\beta$. Such localization can also be used as the initialization of many segmentation algorithms.

In Section 5.1, this framework is employed to locate 2D lung from X-ray image. In Section 5.2, this framework is used to initialize the 3D deformable model.

*Organ segmentation based on deformable model and shape refinement:* Curve or surface based deformable models have been widely used for organ segmentation (Kass et al., 1987; Cootes et al., 1995; Xu and Prince, 1998; Huang and Metaxas, 2008; Shen and Davatzikos, 2000; Li et al., 2005; Zhuang et al., 2010). Many deformable models consist of two iterative steps, local deformation based on low level image information and global shape refinements based on high-level shape priors. For those applications that have noisy image information, high level shape information becomes especially critical. Our sparse shape prior modeling can be used for shape refinement method as a regularization step during deformation. An initialized shape is deformed following the image gradient information. During the deformation procedure, the shape refinement is employed as high level constrains to avoid getting stuck in local minima of the image information. Denote the training shape matrix as $D_S$, and the intermediate deformation result as $y_S$. Then $x$ is computed by solving (5) with $D_S$ and $y_S$. $D_S x$ is used as the refined shape and transformed back. In this refinement procedure, $e$ may not have large values since the model is already roughly aligned after initialization. However, modeling $e$ is still necessary to capture small errors not following Gaussian distribution.

In Section 5.2, this framework is used to segment 3D liver. The whole framework is shown in Fig. 3.

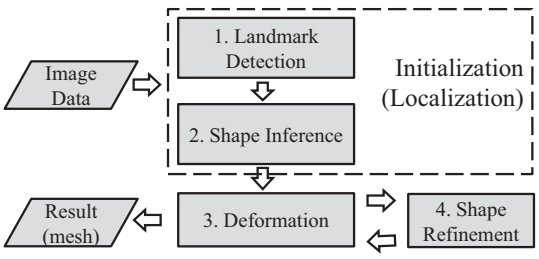



**Fig. 3.** The flowchart of the segmentation framework which consists of landmark detection, shape inference, model deformation and shape refinement. Using landmark detection and shape inference, we are able to do model initialization (i.e., organ localization). After the shape model is initialized, the shape is deformed and refined alternatively to fit the image boundary.

## 5. Experiments

### 5.1. 2D lung localization from X-ray image

#### 5.1.1. Clinical background

Radiography (X-ray) is the most frequently used medical imaging modality due to its fast imaging speed and low cost. About one third radiograph exams are chest radiographs. It is used to reveal various pathologies including abnormal cardiac sizes, pneumonia shadow and mass lesions. The automation of pathology detection often requires robust and accurate lung segmentation. The major challenges of lung segmentation in radiography come from large variations of lung shapes, lung disease and pseudo-boundary close to diaphragm. In chest X-ray, the position, size and shape of lungs often provide important clinical information. Therefore, in this experiment we try to locate the left or right lung using landmark detection and shape inference. Out of 367 X-ray images (all images are from different patients), 200 are used as training data, and the rest 167 are used for testing purpose. In this study, we select training samples to ensure a good coverage of different ages and genders (according to information from DICOM header.) The number of training samples is determined empirically. The ground truths are binary masks of manual segmentation results. A 2D contour is extracted from each mask. To obtain the landmarks for training purpose, we manually select six specific points (e.g., corner points) on the contour, and then evenly and automatically interpolate a fixed amount of points between two neighboring landmarks along the contour. Thus a rough one-to-one correspondence is obtained for both landmarks and shapes. Since the detected landmarks may not be accurate or complete, shape prior is necessary to infer a shape from them. When applying this model, we constantly use the same parameter values for all X-ray images, i.e., $\lambda_1 = 50$ and $\lambda_2 = 0.15$.

#### 5.1.2. Compared methods

In this study, we compare the proposed sparsity-based shape prior modeling with other state-of-the-art methods. For a fair comparison, we intentionally embed different shape models to the same organ localization framework. (It is not fair to compare completely different end-to-end systems, e.g., our system vs. ASM

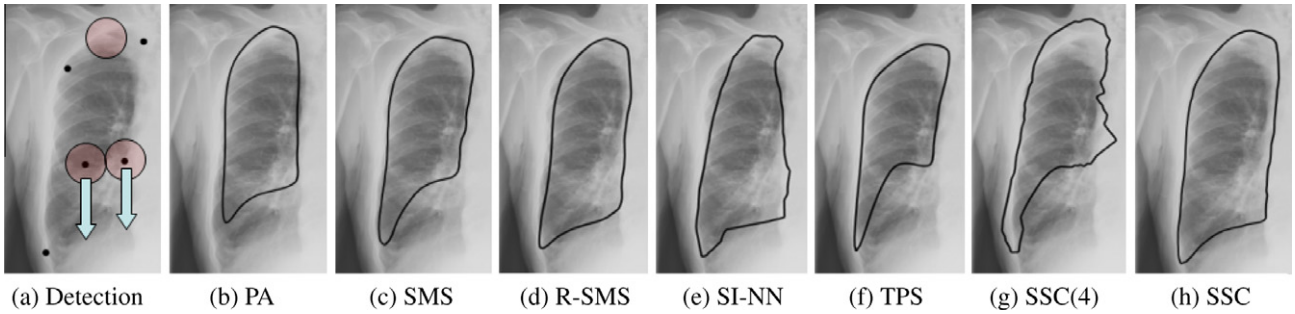



**Fig. 4.** Comparisons of the right lung localization. (a) Detected landmarks are marked as black dots. There are two detection errors and one point missing (marked as circles, and the arrows point to the proper positions). (b) Similarity transformation from Procrustes analysis. (c) Shape Model Search module in ASM, using PCA based method. (d) Shape Model Search in Robust ASM, using RANSAC to improve the robustness. (e) Shape inference method using nearest neighbors. (f) Thin-plate-spline. (g) Sparse representation without modeling $e$, by solving (4). (h) The proposed method by solving (5).

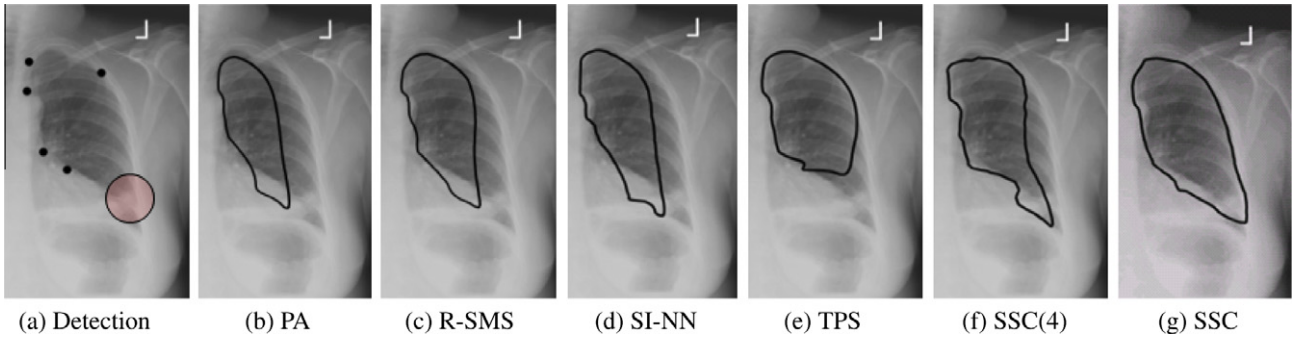



**Fig. 5.** Comparisons of the left lung localization. There is one point missing (marked by a circle), and this lung has a very special shape, which is not captured by the mean shape or its variations. Compared methods are the same as Fig. 4.

system, since the performance difference, if any, cannot be solely attributed to shape prior modeling.) More specifically, the same learning-based algorithm (Zhan et al., 2008) is used to detect landmarks for shape inference (organ localization). Furthermore, the shape is inferred from the detected landmarks directly without iteratively deforming and fitting to the image boundary. The reasons of this setting are twofold. First, enough times of iterative deformations may eventually bring the shape to the image boundary accurately because of the deformation strategy. It is then difficult to evaluate the performance of shape prior models. Thus we apply the inference method only once without the deformation. Second, such one-step inference process is very fast and already good enough as the input for some clinical applications, such as a computer-aided diagnosis (CAD) program. The compared methods are listed as the following:

1. PA: Procrustes analysis is used to find a similarity transformation to fit a mean shape to detected landmarks.
2. SMS: It is the Shape Model Search module in ASM, which employs the PCA method to refine the input shape. Note that we are not using the entire ASM framework including boundary detection and iterative fitting. We focus on the key module of ASM inducing shape prior information.[1]
3. R-SMS: The shape model search step in the robust ASM (Rogers and Graham, 2002) method uses the RANSAC framework to remove the influence of erroneous detections.
4. SI-NN: It stands for shape inference using $k$ nearest neighbors. It is similar to (Georgescu et al., 2005), which uses nearest neighbors to find the closest prototypes in the expert's structure annotations. The distance metric we used is based on the $L2$ distance between corresponding points.
5. TPS: Thin-plate-spline (Bookstein, 1989) is used to deform the mean shape to fit detected landmarks. TPS is a non-rigid and local deformation technology and has been used in robust point matching application (TPS-RPM) (Chui and Rangarajan, 2003).
6. SSC (4): It is the sparse learning shape method without modeling $e$. The result is computed by solving (4).

[1] Since the detected landmarks are very sparse compared to the whole contour, directly fitting ASM to the small number of landmarks result in poor performance. To achieve reasonable performance, a contour is approximated by interpolating points in-between landmarks. When there is a point missing, the mean position of that point is used instead.

#### 5.1.3. Visual comparisons

Some representative and challenging cases are shown in Figs. 4–6. In Fig. 4, there are some mis-detections which are considered as gross errors. The Procrustes analysis, SMS method, SI-NN algorithm and TPS cannot handle such cases. R-SMS is not sensitive to outliers and performs better. SSC (4) also fails to handle such non-Gaussian errors since $e$ is not modeled. SSC can successfully capture such mis-detected points in $e$ and generate a reasonable shape.

In Fig. 5, the underlying shape of the lung is special and different from most other lung shapes (see the mean shape in Fig. 5b). Furthermore, there is a missing point. Neither a transformed mean shape or its variations can represent such shape. TPS is very flexible and able to generate special shapes. However, it fails to handle the missing point. SSC roughly captures the correct shape and generates a better result than the others.

In Fig. 6, all six detections are correct. However, the shape's details are not preserved using the mean shape or its variations. Fig. 7 shows the first five modes (i.e., the five largest shape

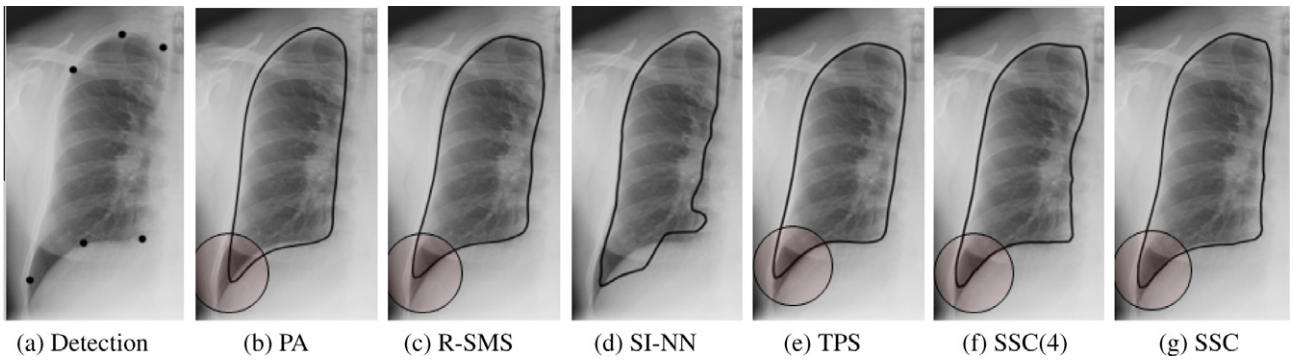



**Fig. 6.** Comparisons of the right lung localization. All six detections are roughly accurate. Thus there is no gross error. The regions marked by circles show the difference of preserved details. Compared methods are the same as Fig. 4.

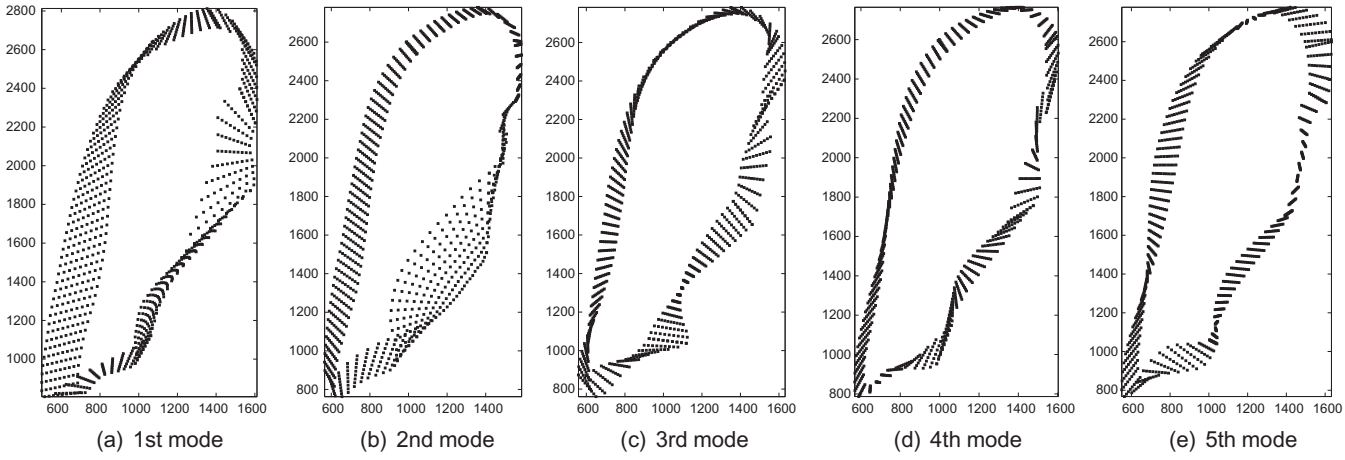



**Fig. 7.** Five largest ASM modes of the right lung. In each mode, we plot shapes from $-3\sigma$ to $3\sigma$ variance. The thickness of boundaries represented shape variation. For example, the major variation of the second mode is in the bottom right, which is the boundary between the right lung and the cardiac. Note that the variance of the bottom left tip is not the major variation in any mode. Thus ASM does not generate accurate results for the case in Fig. 6.

variations) of ASM. The thickness of boundaries represents major variations in that mode. Note that since the variations of the bottom left tip are not the major variation modes, ASM does not able to preserve these local shape details for this testing case in Fig. 6. Both SSC (4) and SSC discover more detail information than other methods. Thus a sparse linear combination is sufficient to recover such details even the gross error $e$ is not modeled. Fig. 8 shows three SSC-selected shape components with largest weights which generate the result in Fig. 6. Two of them do have certain levels of detail information in the bottom left region, although they are still different from the input shape. It demonstrates that our model can discover meaningful shape components, and the final shape composition result can well approximate the testing data. Please note that the proposed model can not "create" local shape details without the support of local appearance cues. Instead, our method aims to "preserve" shape details derived by appearance cues given these details exist in our shape repository. In particular, our method is able to preserve local shape details even when they are not statistically significant in the shape space.

Cases in Figs. 5 and 6 are actually similar to "abnormal testing cases" since they are different from most other shapes in the database. To handle such abnormal shapes, our database needs to contain shape instances with similar local abnormalities. However, our method is still more efficient in abnormal shape modeling than other statistical shape models owing to two reasons. First, our method only requires that abnormal shape instances exist in the database. In contrast, other methods, e.g., ASM cannot model abnormal shape unless these shape instances form principal components in the shape space, which often requires a large number of abnormal shape instances in training set. Second, our method is able to approximate an abnormal shape instance that never appears in the dataset. For example, we can approximate a lung shape with abnormalities in both apex and lateral tip regions with the linear combination of a lung shape with apex abnormality and another lung shape with lateral tip abnormality.

Fig. 9 shows some results from our proposed method on challenging cases with medical instruments. Shape prior contributes to the stableness of the system. It still generates reasonable results with such misleading appearance cues.

#### 5.1.4. Quantitative comparisons

To quantitatively compare different methods, we report the mean values and standard deviations of sensitivity and specificity between binary masks in Table 2 and Fig. 10. Note that the specificity is always good in all methods. The reason is that the size of either left or right lung is relatively small compared to the whole chest X-ray image. Hence, most "true negative" can be correctly found. Thus we also report Dice Similarity Coefficient (DSC) (Popovic et al., 2007), which is a good complement to the other two measurements. DSC is defined as: $2 \times TP/(2 \times TP + FP + FN)$, where $TP$, $FP$ and $FN$ stand for true positive, false positive and false negative, respectively. Generally Procrustes analysis, TPS and SMS achieve good performances, especially when landmarks are

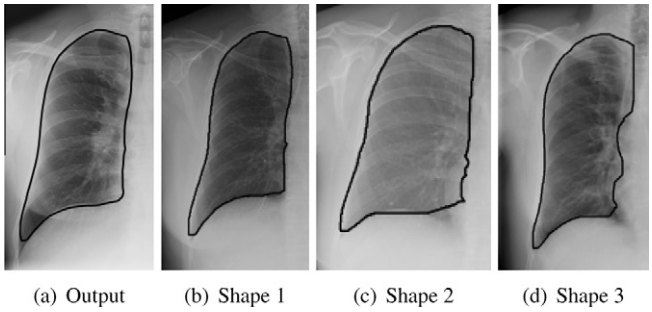



**Fig. 8.** Three shape components with largest weights from our model. "Output" means the result in Fig. 6. The three components have weights 0.5760, 0.2156, and 0.9822, respectively.

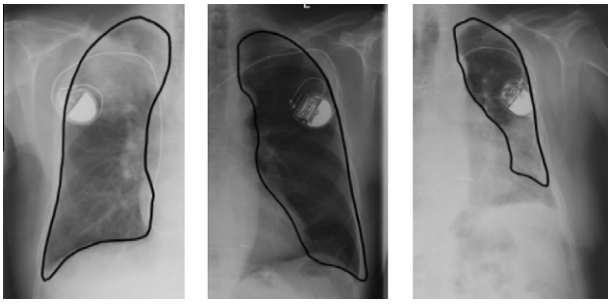


**Fig. 9.** Some localization results from our proposed method on challenging cases with medical instruments. Note that the localized shape may not be exactly on the boundary, since the shape module does not use image information. However, such results are good enough for the input of CAD program or initialization of segmentation algorithms.

correctly detected. However, they are sensitive to non-Gaussian errors. R-SMS can handle this problem because of the RANSAC method, but sometimes it fails to deal with the multimodal distribution of shapes. SI-NN is a good nonparametric method. However, it may not be able to represent shapes which do not appear in the training data. The sparse linear combination by SSC (4) can approximate such shape and it generally performs better than the others. Without modeling error $e$, this method still fails to recover a correct shape. In our proposed method, although the parameter for $e$ is set to a relatively small value, it still contributes to the performance of the model. It performs the best in terms of sensitivity and DSC, without sacrificing the specificity. The standard deviations in Fig. 10 show that SSC also achieves the best stability among all compared methods.

**Table 2**
Quantitative comparisons of seven methods. The sensitivity ($P$%), specificity ($Q$%) and Dice Similarity Coefficient (DSC %) are reported for cases in Figs. 4–6. The best performance of each column is highlighted.

| | Fig. 4 | | | Fig. 5 | | | Fig. 6 | | |
|---|---|---|---|---|---|---|---|---|---|
| Method | $P$ | $Q$ | DSC | $P$ | $Q$ | DSC | $P$ | $Q$ | DSC |
| PA | 62 | 99 | 76 | 50 | 99 | 64 | 93 | 99 | 94 |
| SMS | 66 | 99 | 78 | 61 | 99 | 72 | 93 | 99 | 95 |
| R-SMS | 81 | 99 | 88 | 61 | 99 | 72 | 93 | 99 | 95 |
| SI-NN | 81 | 99 | 87 | 63 | 98 | 73 | 87 | 99 | 90 |
| TPS | 59 | 99 | 74 | 75 | 99 | 79 | 97 | 98 | 94 |
| SSC (4) | 63 | 98 | 71 | 73 | 99 | 79 | 97 | 99 | 96 |
| SSC | **87** | **99** | **91** | **92** | **99** | **91** | **98** | **99** | **96** |

The experiments are performed on a PC with 2.4 GHz Intel Quad CPU, 8 GB memory, with Python 2.5 and C++ implementations. The whole framework is fully automatic. As it benefits from the FISTA algorithm, our algorithm is very efficient. Given this scale of data, it takes around 0.5 s to infer a shape from the landmarks.

### 5.1.5. Evaluation of parameter sensitivity

We also conducted experiments of the parameter sensitivity of $\lambda_2$ (Fig. 11), which controls the sparsity of erroneous detections $e$. A set of parameter values is evaluated for data with or without erroneous detections. Our model consistently achieves similar performance for data without erroneous detections. In these cases, the sparse linear combination of training shapes can already approximate the input shape very well. Thus values in $e$ are usually flat zeros in these cases, even with different values of $\lambda_2$. For data with erroneous detections, $\lambda_2$ is critical as it controls the number of nonzero elements in $e$. Our model still generates stable results in a reasonably wide range of parameter values, although the performance eventually goes down when $\lambda_2$ is far away from this range. This insight analysis shows that this model is not sensitive to parameter values in a certain level,

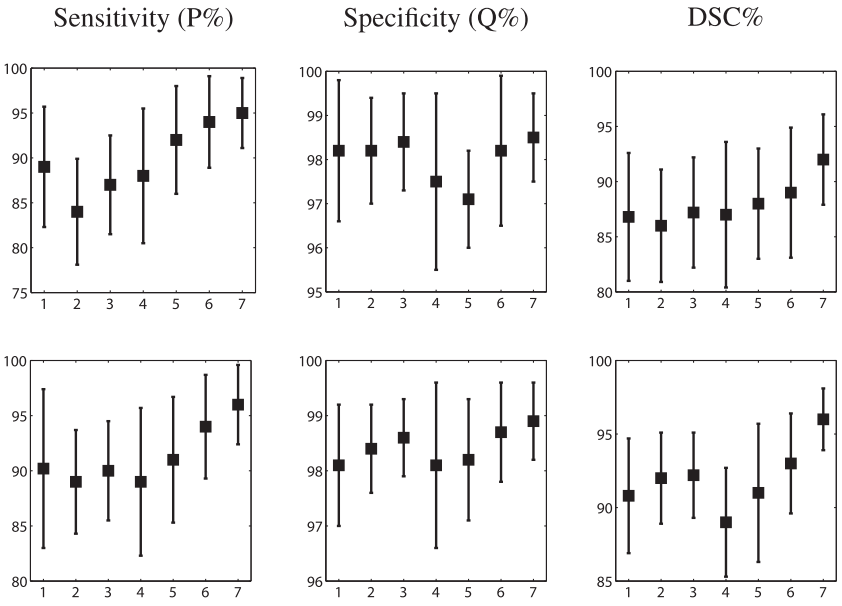
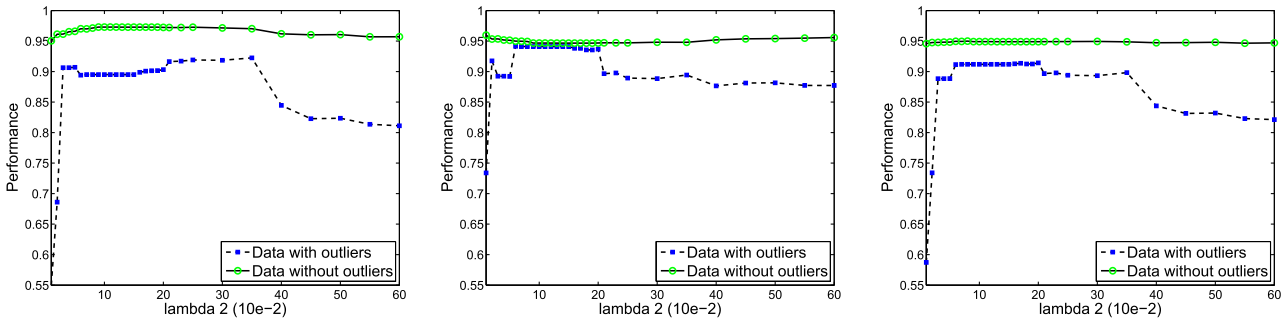



**Fig. 10.** Mean values ($\mu$) and standard deviations ($\sigma$) of P, Q and DSC from the left lung (1st row) and right lung (2nd row) of all testing data. In each figure, $y$-axis is the performance of $P$, $Q$ or DSC. $x$-axis means seven methods (from left to right) in the same order as Table 2. Squares denote $\mu$, segments denote $\sigma$.




(a) Sensitivity (P%) (b) Specificity (Q%) (c) DSC%

**Fig. 11.** Parameter sensitivity of $\lambda_2$. A set of parameter values is tested for data with or without erroneous detections. Generally data without erroneous detections (green line) consistently achieves similar performance for different parameter values, and data with erroneous detections (blue line) also produce stable results in a wide range. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

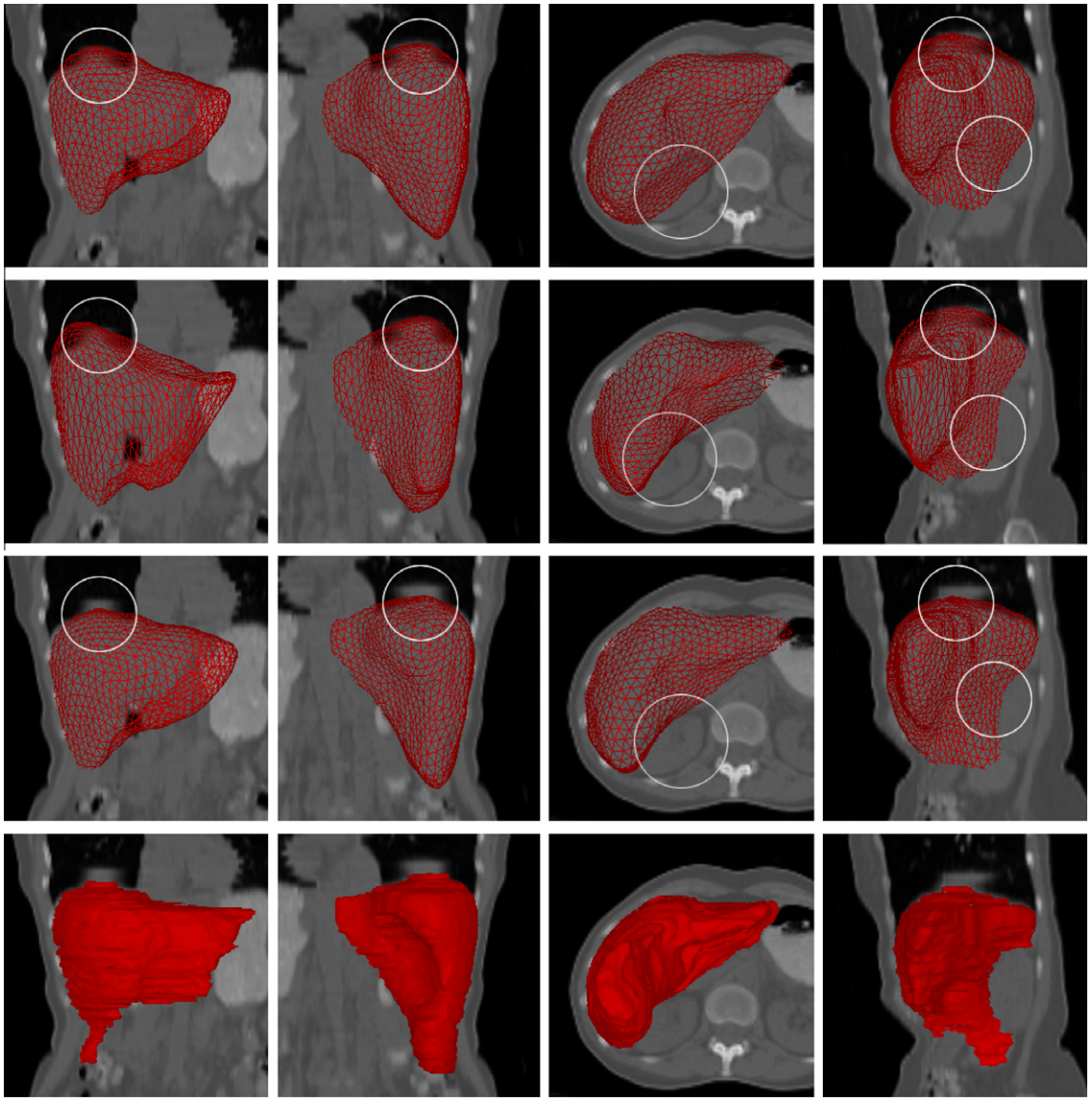


**Fig. 12.** 3D initialization results for the segmentation framework. First row: using the global transformation based on Procrustes analysis. Second row: using TPS which is a local and nonrigid deformation technique. Third row: using SSC to constrain the shape. Fourth row: ground truth from manual segmentation. Its surface mesh is obtained by applying iso-surface method on the binary image. The results from global and local deformation incorrectly include part of the liver or the artifact caused by breath. The differences are marked by circles. Results from SSC is more similar to the ground truth.

which makes it general for different data in the same application. Another observation is that when there is no outlier, this model usually generates slightly better results than having erroneous detections, which means that erroneous detections still adversely affect the model performance even with this sparse gross errors constraint.

### 5.2. 3D liver segmentation from low-dose CT

#### 5.2.1. Background and experimental setting

Whole body PET-CT is an emerging medical imaging modality that combines a Positron Emission Tomography (PET) and an X-ray Computed Tomography (CT) scan. The co-registered anatomical (CT) and functional (PET) information benefits various clinical practices, especially for oncology studies. Due to the high variations of F-fluorodeoxyglucose (FDG) uptakes across different organs, the preferred way to interpret PET-CT images is in an organ-specific fashion, which requires organ segmentation. In traditional CT images, organ segmentation such as liver segmentation (Ling et al., 2008) has been extensively investigated. However, to decrease radiations to patients, CT images in PET-CT scans usually have low dose and large slice thickness, which result in low contrast and fuzzy boundaries between organs. Hence, organ segmentation in whole body PET-CT becomes more challenging than traditional CT. In this experiment we try to segment the liver from the low-dose whole body CT, using deformable models and shape refinement. The 3D ground truth of low-dose CT is manually segmented by multiple clinical experts. Sixty-seven scans are annotated. Forty are used as training data to train the landmark detector and also used to construct the data matrix $D_S$, the rest 27 are testing. To obtain the one-to-one correspondence for vertices among all shapes, we choose one shape as a reference and register it to all the others using adaptive-focus deformable model (AFDM) (Shen and Davatzikos, 2000). The shape has around 1000 vertices, and 20 are selected as landmarks. The initialization step is based on landmark detection and similar to Section 5.1. A surface model is then fitted into the image, and is hierarchically deformed to the image gradient information (Zhan and Shen, 2006). During the deforma-

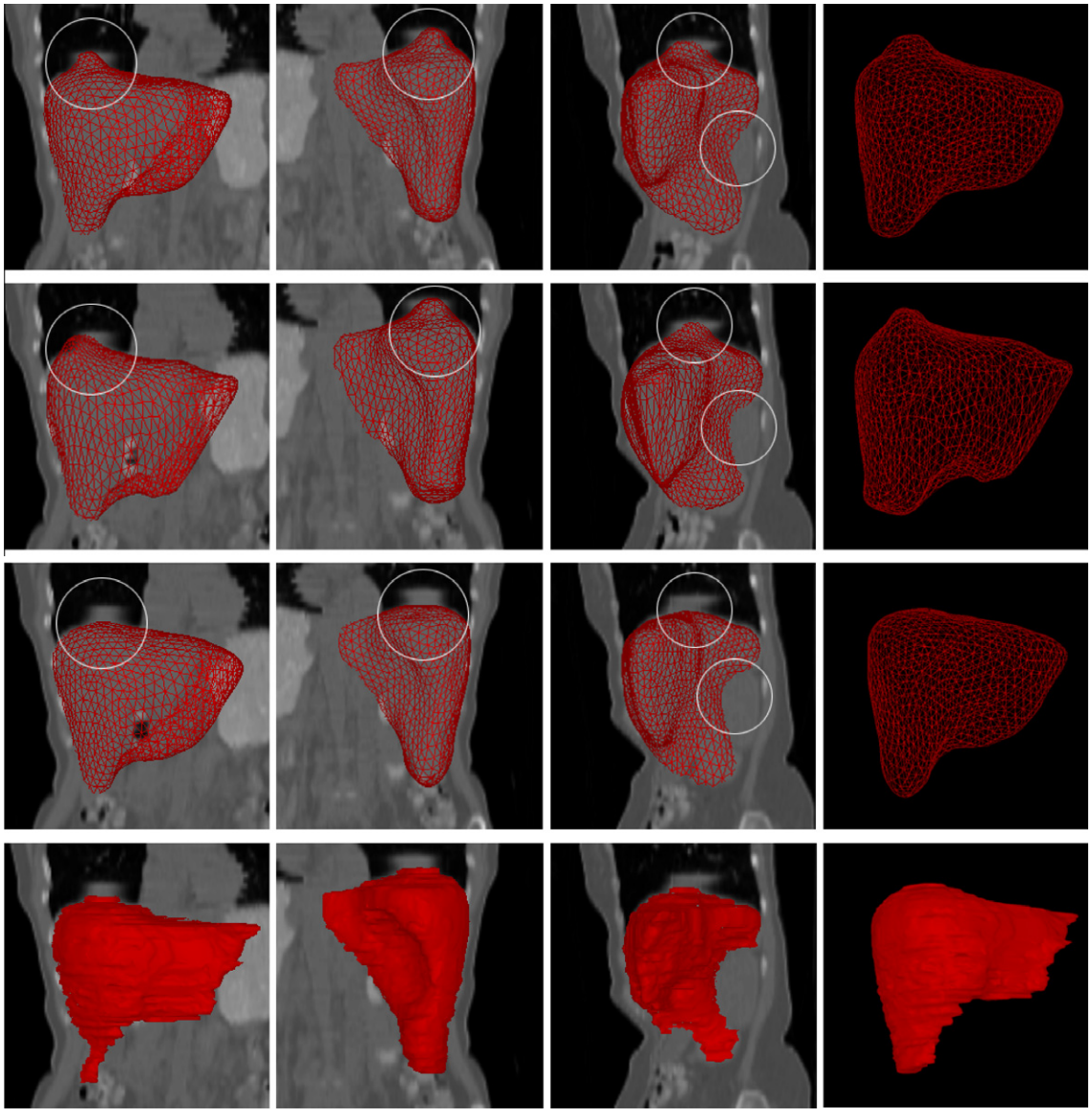


**Fig. 13.** Visual comparisons of deformation results using the initialization from Fig. 12. Note that the deformation modules are the same. First row: using initialization from the global transformation. Second row: using initialization from TPS. Third row: using initialization from SSC. Fourth row: the same ground truth as in Fig. 12. Generally better initialization ensures fewer iterations of deformation and more accurate results. The differences are highlighted by circles.

tion procedure, our sparsity-based shape prior modeling is used to refine the shape, which is actually a regularization step to avoid the deformable model getting stuck in local minima of the image information. This method is fully automatic without any manual interaction.

In this study we use $\lambda_1 = 50$ and $\lambda_2 = 0.3$ for all data. For comparisons, we also employ some other shape prior modeling methods (e.g., SMS, SI-NN) to refine the intermediate deformation result after several iterations. For a fair comparison, the results for comparisons are based on the same segmentation framework, i.e., using the same landmark detection and hierarchical deformation algorithm, with different shape prior modeling.

*5.2.2. Visual comparisons*

Figs. 12–15 show some visual comparisons in 3D. Fig. 12 compares the shape inference based initialization using the same landmark detection input. Since the image contrast of low-dose CT is very low, landmark detector may easily fail to locate correct positions. SSC is less sensitive to such errors. Its initialization result is already very close to the object boundary, which can benefit the deformation procedure. Fig. 13 shows the deformation results using the initialization in Fig. 12. The deformation modules are all the same. Shape refinement is not used in this experiment. Better initialization not only ensures fewer iterations of deformation but also produces more accurate results. SSC obtains better results compared to other methods. Fig. 14 compares the shape refinement results after deformation. The refined shape may not be exactly on the image boundary since this part is just a regularization step without considering any image information. After the refinement, the segmentation framework still needs to perform some iterations of deformation to reach the image boundary. Certainly, a better refinement result can benefit the whole segmentation framework. The refined shape of ASM type method follows the mean shape and variations, but it incorrectly includes a large part of the kidney. SSC is more specific to this image and is more accurate. Fig. 15 shows the final segmentation results from a highly noisy data. Without any shape prior information, the method fails to provide an accurate result. Using proposed method, the system is more robust and less sensitive to image noise. Thus it provides more accurate results.

*5.2.3. Quantitative comparisons*

Table 3 shows the quantitative comparisons. To evaluate 3D accuracy, we report the mean value and standard deviation of the distances between shape surfaces. SSC achieves the best performance compared to other shape refinement approaches. The whole system takes around 20 s (in a Python implementation) to segment the liver, including data loading, initialization, deformation and shape regularization. SSC not only improves the robustness of the deformable model, but also decreases the iteration times of deformation since it helps avoid local minima of image information.

*5.3. Discussion*

The experimental results show the following facts.

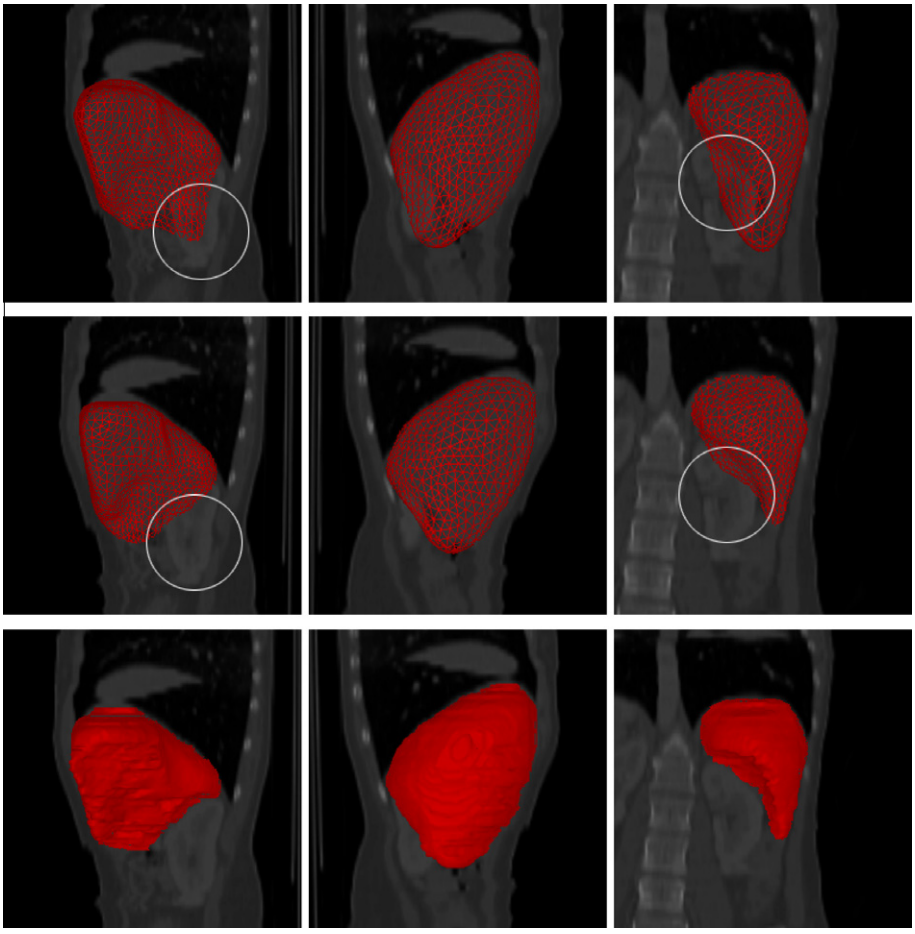


**Fig. 14.** Visual comparisons of refinement results after deformation. First row: using shape model search step in SMS as the refinement. Second row: using SSC. Third row: ground truth from manual segmentation. The refined shape may not be exactly on the image boundary since this part is just a regularization step without considering any image information. The SMS result incorrectly includes a large part of the liver (marked by circles).

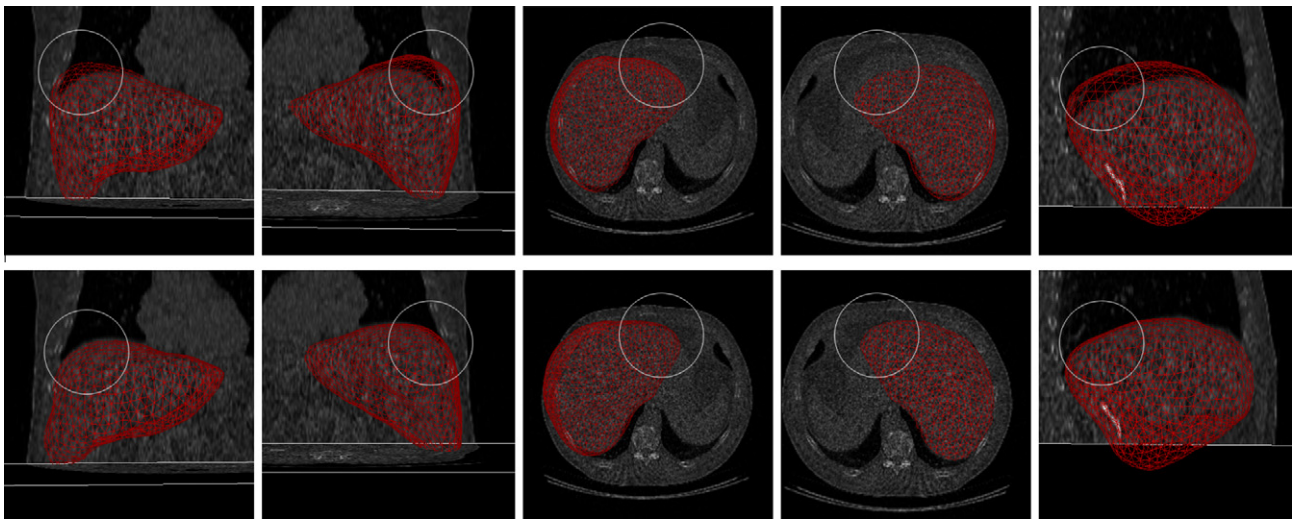


**Fig. 15.** Visual comparisons of final segmentation results from highly noisy data. First row: deformation without shape constraint. Second row: using SSC as the shape constraint. Compared to the system not using shape prior, SSC is less sensitive to image noise and is more robust. The differences are marked by circles.

1. This implicitly incorporated shape constraint benefits the interpretation of images. Such shape-based information improves the robustness and accuracy of low-level algorithms.
2. The sparse linear combination of shape repository is able to well approximate an input shape if the input does not have gross errors. The $L1$ norm constraint of $e$ handles the non-Gaussian residuals caused by occlusion, miss detection or point missing.
3. Our method is able to use a subset of vertices on the contour and surface as the input. Sparse learning is applied on the this subset and computes a group of coefficients. Such coefficients are then used to combine the whole contour or surfaces, which generate reasonable results. This property can seamlessly incorporate with multi-resolution scheme. Our method can naturally extend to multi-resolution without significant overhead of implementation or computational complexity.
4. We use it to solve two medical image applications with different modalities and different types of shapes (i.e., 2D contour and 3D surface mesh). Thus SSC is independent of the application or data dimension, and does not substantially increase the computational complexity. Existing shape based segmentation systems can be easily extended by adding this proposed method as a shape refinement module.

**Table 3**
Quantitative comparisons of the mean values and standard deviations of the distances (voxel) between surfaces. The best performance of each column is highlighted.

| Method | Fig. 12 | Fig. 14 | All data |
|---|---|---|---|
| SMS | 2.26 ± 1.72 | 1.81 ± 2.10 | 2.16 ± 1.68 |
| SI-NN | 4.88 ± 3.61 | 3.34 ± 3.78 | 3.82 ± 3.12 |
| TPS | 2.92 ± 2.19 | 5.12 ± 5.29 | 3.39 ± 3.16 |
| SSC (4) | 1.42 ± 1.02 | 2.39 ± 2.31 | 2.24 ± 1.70 |
| SSC | **1.31 ± 0.95** | **1.10 ± 0.87** | **1.13 ± 0.83** |

## 6. Conclusions

In this paper, we proposed a sparse learning based method to implicitly model shape priors. It alleviates three problems in a unified framework, i.e., modeling complex shape variations, handling non-Gaussian errors and preserve local detail information of the input shape. A segmentation framework is proposed by using this method as a shape inference module and shape refinement module. This framework is extensively validated on two medical applications, 2D lung localization in X-ray images and 3D liver segmentation in low-dose CT scans. Compared to state-of-the-art methods, our shape prior model exhibits better performance in both studies.

In the future, we look forward to receiving more shape instances to construct our training dataset. As the convex property of (4) guarantees a global optimal solution, more training samples will ensure a better composition to approximate the input shape instance. However, since the optimization procedure will take more time with more training samples, it is important to build a shape repository with the most representative shape instances. This will be an interesting topic in our future work. Potential solutions include matching pursuit, affinity propagation clustering, etc. We are also interested in the hierarchical modeling of the 3D surface shape. If the shape can be subdivided into pieces according to both the geometry and image information, we can apply SSC on each subdivision separately, and then stitch them together to be the whole surface. This approach can potentially improve the computational cost and decrease the number of required training data. Last but not least, we also plan to apply this model to more applications, such as shape matching and registration.

## References

Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences 2, 183–202.
Bookstein, F., 1989. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11.
Candes, E., Tao, T., 2006. Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Transaction on Information Theory 52, 5406–5425.
Chui, H., Rangarajan, A., 2003. A new point matching algorithm for non-rigid registration. Computer Vision and Image Understanding 89, 114–141.
Cootes, T., Taylor, C., Cooper, D., Graham, J., 1995. Active shape model – their training and application. Computer Vision and Image Understanding 61, 38–59.
Cootes, T.F., Taylor, C.J., 1997. A mixture model for representing shape variation. Image and Vision Computing, 110–119.
Davatzikos, C., Tao, X., Shen, D., 2003. Hierarchical active shape models, using the wavelet transform. IEEE Transactions on Medical Imaging 22, 414–423.
Dellaportas, P., Forster, J., Ntzoufras, I., 2002. On bayesian model and variable selection using MCMC. Statistics and Computing 12, 27–36.
Donoho, D.L., 2004. For most large under determined systems of equations, the minimal l1-norm near-solution approximates the sparsest near-solution. Technical Report. Communications on Pure and Applied Mathematics.

Duta, N., Sonka, M., 1998. Segmentation and interpretation of mr brain images an improved active shape model. IEEE Transactions on Medical Imaging 17, 1049–1062.

Etyngier, P., Segonne, F., Keriven, R., 2007. Shape priors using manifold learning techniques. International Conference on Computer Vision, pp. 1–8.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 381–395.

George, E., McCulloch, R., 1993. Variable selection via gibbs sampling. Journal of the American Statistical Association, 881–889.

Georgescu, B., Zhou, X., Comaniciu, D., Gupta, A., 2005. Database-guided segmentation of anatomical structures with complex appearance. IEEE Conference on Computer Vision and Pattern Recognition, 429–436.

Goodall, C., 1991. Procrustes methods in the statistical analysis of shape. Journal of the Royal Statistical Society 53, 285–339.

Gu, L., Kanade, T., 2008. A generative shape regularization model for robust face alignment. European Conference on Computer Vision, 413–426.

Heimann, T., Meinzer, H.P., 2009. Statistical shape models for 3D medical image segmentation: a review. Medical Image Analysis 13, 543–563.

Huang, J., Huang, X., Metaxas, D., 2009. Learning with dynamic group sparsity. International Conference on Computer Vision, 64–71.

Huang, X., Metaxas, D., 2008. Metamorphs: deformable shape and appearance models. IEEE Transactions on Pattern Analysis and Machine Intelligence 30, 1444–1459.

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: active contour models. International Journal of Computer Vision 1, 321–331.

Kuo, L., Mallick, B., 1998. Variable selection for regression models. Sankhyā: The Indian Journal of Statistics, Series B, 65–81.

Lekadir, K., Merrifield, R., Yang, G.-zhong, 2007. Outlier detection and handling for robust 3D active shape models search. IEEE Transactions on Medical Imaging 26, 212–222.

Li, C., Xu, C., Gui, C., Fox, M.D., 2005. Level set evolution without re-initialization: a new variational formulation. IEEE Conference on Computer Vision and Pattern Recognition 1, 430–436.

Ling, H., Zhou, S., Zheng, Y., Georgescu, B., Suehling, M., Comaniciu, D., 2008. Hierarchical, learning-based automatic liver segmentation. IEEE Conference on Computer Vision and Pattern Recognition, 1–8.

Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A., 2009. Non-local sparse models for image restoration. International Conference on Computer Vision, 2272–2279.

Nahed, J.A., Pierre Jolly, M., Zhong Yang, G., 2006. Robust active shape models: a robust, generic and simple automatic segmentation tool. International Conference on Medical Image Computing and Computer Assisted Intervention.

Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y., 2010. RASL: Robust alignment via sparse and low-rank decomposition for linearly correlated images. IEEE Conference on Computer Vision and Pattern Recognition.

Popovic, A., de la Fuente, M., Engelhardt, M., Radermacher, K., 2007. Statistical validation metric for accuracy assessment in medical image segmentation. International Journal of Computer Assisted Radiology and Surgery 2, 169–181.

Rogers, M., Graham, J., 2002. Robust active shape model search. European Conference on Computer Vision, pp. 517–530.

Rousson, M., Paragios, N., 2002. Shape priors for level set representations. European Conference on Computer Vision 2351, 416–418.

Saragih, J., Lucey, S., Cohn, J., 2009. Face alignment through subspace constrained mean-shifts. International Conference on Computer Vision, 1034–1041.

Shen, D., Davatzikos, C., 2000. An adaptive-focus deformable model using statistical and geometric information. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 906–913.

Shi, Y., Qi, F., Xue, Z., Chen, L., Ito, K., Matsuo, H., Shen, D., 2008. Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. IEEE Transactions on Medical Imaging 27, 481–494.

Sjostrand, K. et al., 2007. Sparse decomposition and modeling of anatomical shape variation. IEEE Transactions on Medical Imaging 26, 1625–1635.

Starck, J., Elad, M., Donoho, D., 2004. Image decomposition via the combination of sparse representations and a variational approach. IEEE Transactions on Image Processing 14, 1570–1582.

Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y., 2009. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence 31, 210–227.

Xu, C., Prince, J., 1998. Snakes, shapes and gradient vector flow. IEEE Transactions on Image Processing 7, 359–369.

Yan, P., Kruecker, J., 2010. Incremental shape statistics learning for prostate tracking in TRUS. International Conference on Medical Image Computing and Computer Assisted Intervention, 42–49.

Yan, P., Xu, S., Turkbey, B., Kruecker, J., 2010. Discrete deformable model guided by partial active shape model for TRUS image segmentation. IEEE Transactions on Biomedical Engineering 57, 1158–1166.

Zhan, Y., Shen, D., 2006. Deformable segmentation of 3D ultrasound prostate images using statistical texture matching method. IEEE Transactions on Medical Imaging 25, 256–272.

Zhan, Y., Zhou, X.S.Z., Peng, Z., Krishnan, A., 2008. Active scheduling of organ detection and segmentation in whole-body medical images. International Conference on Medical Image Computing and Computer Assisted Intervention, 313–321.

Zhang, S., Huang, J., Huang, Y., Yu, Y., Li, H., Metaxas, D., 2010. Automatic image annotation using group sparsity. IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 3312–3319.

Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D., Zhou, X., 2011a. Sparse shape composition: a new framework for shape prior modeling. IEEE Conference on Computer Vision and Pattern Recognition, 1025–1032.

Zhang, W., Yan, P., Li, X., 2011b. Estimating patient-specific shape prior for medical image segmentation. In: International Symposium on Biomedical Imaging. IEEE, pp. 1451–1454.

Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D., 2008. Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. IEEE Transactions on Medical Imaging 27, 1668–1681.

Zhu, Y., Papademetris, X., Sinusas, A., Duncan, J., 2008a. Bidirectional segmentation of three-dimensional cardiac MR images using a subject-specific dynamical model. International Conference on Medical Image Computing and Computer-Assisted Intervention, 450–457.

Zhu, Y., Papademetris, X., Sinusas, A., Duncan, J., 2008b. Segmentation of left ventricle from 3D cardiac MR image sequences using a subject-specific dynamical model. IEEE Conference on Computer Vision and Pattern Recognition.

Zhu, Y., Papademetris, X., Sinusas, A., Duncan, J., 2009. A dynamical shape prior for LV segmentation from RT3D echocardiography. International Conference on Medical Image Computing and Computer Assisted Intervention, 206–213.

Zhu, Y., Papademetris, X., Sinusas, A., Duncan, J., 2010. Segmentation of the left ventricle from cardiac MR images using a subject-specific dynamical model. IEEE Transactions on Medical Imaging 29, 669–687.

Zhuang, X., Rhode, K., Razavi, R., Hawkes, D., Ourselin, S., 2010. A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI. IEEE Transactions on Medical Imaging 29, 1612–1625.