Online Robust Projective Dictionary Learning: Shape Modeling for MR-TRUS Registration

Yi Wang, Qingqing Zheng, *Student Member, IEEE*, and Pheng Ann Heng, *Senior Member, IEEE*

***Abstract*— Robust and effective shape prior modeling from a set of training data remains a challenging task, since the shape variation is complicated, and shape models should preserve local details as well as handle shape noises. To address these challenges, a novel robust projective dictionary learning (RPDL) scheme is proposed in this paper. Specifically, the RPDL method integrates the dimension reduction and dictionary learning into a unified framework for shape prior modeling, which can not only learn a robust and representative dictionary with the energy preservation of the training data, but also reduce the dimensionality and computational cost via the subspace learning. In addition, the proposed RPDL algorithm is regularized by using the ℓ_1 norm to handle the outliers and noises, and is embedded in an online framework so that of memory and time efficiency. The proposed method is employed to model prostate shape prior for the application of magnetic resonance transrectal ultrasound registration. The experimental results demonstrate that our method provides more accurate and robust shape modeling than the state-of-the-art methods do. The proposed RPDL method is applicable for modeling other organs, and hence, a general solution for the problem of shape prior modeling.**

***Index Terms*— Dictionary learning, dimension reduction, MR-TRUS registration, online, prostate segmentation, shape modelling.**

I. Introduction

PROSTATE cancer is the most common noncutaneous cancer and the second leading cause of cancer-related deaths in males [1]. The routine clinical modality for imaging the prostate, especially for image-guided prostate biopsy and treatments, is transrectal ultrasound (TRUS) [2]. However, because of the low image quality of ultrasound, TRUS usually results in an epidemic of overdetection and overtreatment of prostate cancer [3]. On the other hand, with the development of the technical feasibility of magnetic resonance (MR)-TRUS registration [4]–[10], the MR-TRUS fusion based targeted prostate biopsy has emerged as a new standard for the current prostate biopsy schemes [3], [11], [12].

We have proposed a biomechanically constrained model-to-surface registration approach [9] in our previous work to compensate for the prostate deformation caused by the insertion of the TRUS probe during MR-TRUS registration. In our approach, based on principal component analysis (PCA) technique [13], a personalized statistical deformable model (PSDM) is constructed from numerous deformation instances simulated with the physically-plausible boundary conditions and patient-specific tissue elastic parameters. The constructed PSDM is then served as prior to constrain the devised hybrid point matching process to align the MR prostate surface with the TRUS surface. However, PCA is an unsupervised dimension reduction method that seeks to preserve the global energy of training data but not the most effective representation of them. When the training data are insufficient or contaminated by outliers, the PSDM learned by PCA may be much biased. In addition, the prerequisite for the proposed model-to-surface registration is the accurate segmentation of TRUS images. Any incorrect segmentation may degrade the surface registration performance, which is another limitation for the previous study.

Addressing the aforementioned limitations is the main purpose of this study. We propose a robust projective dictionary learning scheme for the efficient shape prior modelling. The proposed method integrates the dimension reduction and dictionary learning into a unified online framework, which is beneficial to exploit more effectively the representative information of training data via simultaneously learning the projection and dictionary. Furthermore, the jointly learned projection and dictionary is able to robustly and efficiently model the prostate shape/deformation based on projective sparse representation, despite the existence of local segmentation inaccuracies.

A. Relevant Work

Effective modelling of organ shape priors from a set of training data still remains challenging, which can be summarized in twofold. First, shape models should not only

Manuscript received October 14, 2017; accepted November 20, 2017. Date of publication November 27, 2017; date of current version April 2, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61571304 and Grant 61701312, in part by the Shenzhen Peacock Plan under Grant KQTD2016053112051497, in part by the Shenzhen Basic Research Project under Grant JCYJ20150525092940982 and Grant JCYJ20150525092940988, in part by the Research Grants Council of the Hong Kong Special Administrative Region under Grant CUHK 14225616, and in part by the Hong Kong Innovation and Technology Fund under Grant ITT/007/16GP. *(Yi Wang and Qingqing Zheng contributed equally to this work.) (Corresponding author: Yi Wang.)*

Y. Wang is with the National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Provincial Key Laboratory of Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen 518060, China (e-mail: onewang@szu.edu.cn).

Q. Zheng and P. A. Heng are with the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2017.2777870

retain the global energy of training data, but also preserve the discriminative local details of them. Second, since the organ shapes are usually extracted from the segmentation results, they may contain some gross errors or outliers due to the erroneous segmentation. Thus shape models have to be capable of robustly and efficiently approximating a new instance despite the existence of gross errors or outliers.

Many approaches have been proposed to address some of above challenges for robust and efficient shape/deformation modelling. Several dimension reduction techniques, such as PCA [13], kernel PCA [14], linear discriminant analysis (LDA) [15], have been employed to reduce the dimension of training data and also retain useful information of them. As one of the most widely used dimension reduction techniques, PCA projects the high dimensional training data into a linear subspace spanned by a few dominant eigenvectors of the covariance matrix of training data. PCA-based models have shown promise for modelling prostate shape/deformation prior thus constraining MR-TRUS registration [7]–[10], [16], [17]. PCA is simple to implement and efficient in reducing Gaussian noise appeared in training data. However, PCA-based models may be biased and misrepresentative when training data are insufficient or contaminated by non-Gaussian noises. In addition, as an unsupervised dimension reduction method, PCA only aims to preserve the global energy but not local details, especially when such details are not predominant in training data.

On the other hand, sparse representation based modelling has been employed to robustly approximate a new shape instance despite its gross errors [18], [19]. Such sparse representation based methods have been widely used in computer vision society [20]–[23], exploiting a sparse linear combination of training data to approximately represent a query instance. In the context of shape modelling, sparse representation refers an input shape to a sparse linear combination of shapes (atoms) in the shape repository (also known as dictionary). The shapes of human organs usually contain strong shape priors. Therefore, each shape often lies in a small subset of shapes in the shape repository, namely the shape can be characterized by a feature vector of much lower dimensionality. This implies that atoms in the dictionary are not incoherent but highly correlated [24]. While parts of the shape may contain gross errors due to the erroneous measurement/segmentation, such gross errors are usually sparse and different from each other [18]. Benefiting from the sparsity techniques, sparse representation is robust to gross errors and is able to represent the clean component of a shape by the sparse coding on the shape repository [25]. Zhang *et al*. [18] employed sparse shape composition to model organ shapes. The original training data (the segmented organ shapes) are used as the dictionary atoms to represent the newly input shape. Intuitively, a more robust and efficient representation could be obtained if the dictionary is learned from the original training data [19]. However, sparse representation based dictionary learning may suffer from high dimensionality but small training set problem, especially when exploring the high-dimensional volumetric data, which limits the practical application of dictionary learning.

In general, the dimension reduction and dictionary learning techniques are mostly independently investigated. Conventionally, dimension reduction (random projection, PCA, etc.) is performed to original training data with high dimensionality, and then the dimensionality reduced data are employed for dictionary construction [26]. However, such dimensionality reduced data obtained for dictionary learning may not capture the most representative training information, thus still limiting the performance of dictionary learning. In principle, the dimension reduction should aim to preserve representative information of training data, and meanwhile remove unreliable dimensions for robust dictionary learning. Therefore, the dimension reduction and dictionary learning have to be jointly performed to fully exploit the representative information of training data. To this end, we propose an online robust projective dictionary learning (RPDL) scheme, which integrates the dimension reduction and dictionary learning into a unified online framework. The joint learning studies more expressive projection and dictionary from each other, thus a more effective shape representation can be obtained by using projective sparse representation with a learned dictionary. In the proposed RPDL, the projection and dictionary can be simultaneously solved using an efficient optimization scheme. Furthermore, the joint learning is embedded in an online framework so that our scheme is of memory and time efficiency.

B. Contributions

The main contributions of our work are threefold.

1) We propose a novel sparse representation scheme, called robust projective dictionary learning (RPDL) to address the shape modelling problem. Specifically, we integrate the dimension reduction and dictionary learning into a unified framework, which can not only learn a robust and representative dictionary with the energy preservation of the training data, but also reduce the dimensionality and computational cost via the subspace learning. To our best knowledge, the proposed RPDL is the first to address the shape modelling problem with the simultaneous projection and dictionary learning.
2) We propose a novel energy function for RPDL and embed it in an online mechanism. We model the sparse projective data fitting term in robust dictionary learning by using the ℓ_1 norm to handle the outliers and noises. To solve the resulting optimization problem, we also develop an efficient algorithm based on the alternating direction framework. Benefitting from the online learning mechanism, the proposed RPDL is of memory and time efficiency, which notably enhances its practicality and usability.
3) We extensively evaluate the proposed RPDL method on the application of MR-TRUS registration. The experimental results show that our method achieves state-of-the-art performance. Due to its robustness, the newly estimated shape with RPDL could be regarded as refinement of the incorrectly segmented shape.

The remainder of this paper is organized as follow. Section II presents the details of the proposed online robust

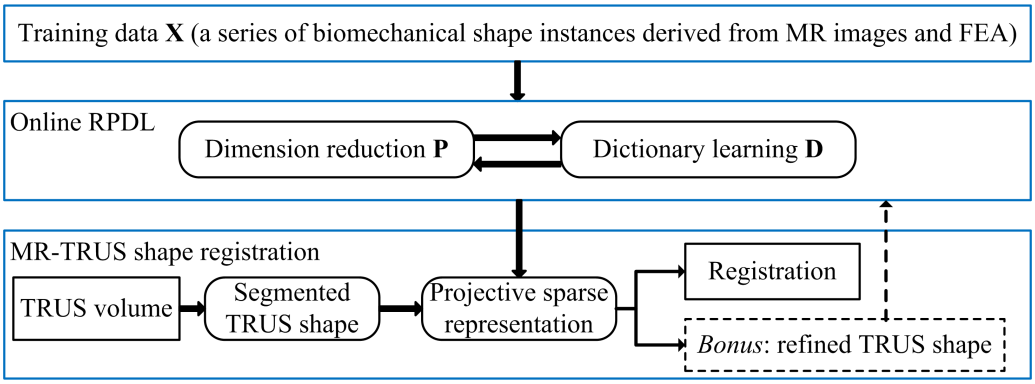


Fig. 1. Flowchart of the proposed registration framework.

projective dictionary learning scheme. Section III presents the experimental results of the proposed RPDL for the application of MR-TRUS registration. Section IV elaborates the discussion of the proposed method, and the conclusion of this study is given in Section V.

II. Method

The effective shape modelling from a set of training data is difficult since the shape variation is complicated, and shape models should preserve local details as well as handle shape noises. To address these issues, we propose a novel robust projective dictionary learning (RPDL) scheme for shape modelling and employ it to model prostate shape prior for the application of MR-TRUS registration (Fig. 1). In the proposed RPDL, the dimension reduction and dictionary learning are jointly performed to fully exploit the representative information of training shapes. Then the jointly learned projection and dictionary is able to robustly and efficiently represent the query TRUS shape based on projective sparse representation, namely register the MR shape with the TRUS shape.

We discuss the proposed RPDL for shape modelling and shape registration in detail in the following three subsections: (1) training data preparation based on biomechanical modelling, (2) shape prior modelling based on online robust projective dictionary learning, and (3) robust shape estimation based on projective sparse representation.

A. Training Data Preparation

For each patient, we prepare training shape instances according to the procedure described in our previous work [9]. More specifically, a series of patient-specific TRUS prostate shapes are simulated with the anatomical meshes derived from MR images and various physically-plausible boundary conditions, based on finite element analysis (FEA) technique (Fig. 2).

To obtain the anatomical meshes, we segment the prostate in MR images via an interactive segmentation software: SmartPaint [27]. Then, 50 evenly spaced points from the segmented contour in each slice are sampled using a custom interface developed using MATLAB (The MathWorks, Natick, MA, USA). All these points are further converted into triangulated anatomical meshes using an adaptive skeleton climbing method [28], which can overcome the gap-filling problem in

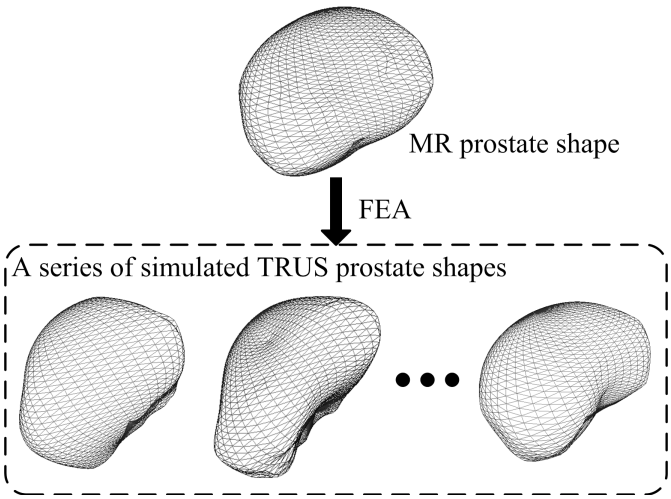


Fig. 2. Training data preparation: a series of TRUS prostate shapes are simulated based on MR prostate shape and FEA.

traditional marching cubes algorithm. All generated meshes are imported into the commercial software COMSOL Multiphysics (COMSOL, Inc., LA, USA) for the FE analysis.

When undergoing TRUS imaging, the prostate often deforms, primarily due to the insertion of the TRUS probe. To obtain training shape instances that are representative for clinical cases, we conduct a series of simulations with biomechanical parameters obtained from shear wave elastography [9] and different settings for the probe insertion conditions, via the FEA software - COMSOL. Specifically, we perturb the pose parameters, i.e., orientation, displacement, and original position, of the virtual 3D TRUS probe within the specific ranges defined in [9] to provide various boundary conditions for each time of simulation. All the simulated TRUS prostate shapes are calculated by the non-linear FE solver embedded in the solid mechanical deformation module of COMSOL. To represent the ith simulated shape, the coordinates of all its m mesh vertices are concatenated into a vector $\mathbf{x}_i \in \mathbb{R}^q$, where $q = m \times 3$.

B. Shape Prior Modelling via Online Robust Projective Dictionary Learning

1) Problem Formulation: Given the input training data $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, where $\mathbf{x}_i \in \mathbb{R}^q$ denotes each training shape,

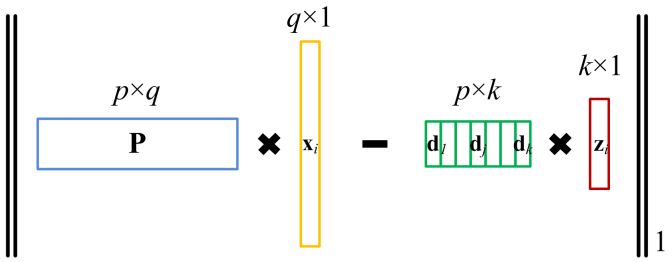


Fig. 3. Illustration of the term $||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1$ in our RPDL scheme.

we assume that the dictionary $\mathbf{D}$ belongs to a closed, convex and bounded set $\mathcal{D}$ as:

$$\mathcal{D} \triangleq \{\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_k], \\ s.t.\ \mathbf{d}_j^T \mathbf{d}_j \leq 1, \forall j = 1, \cdots, k\}, \tag{1}$$

where $\mathbf{d}_j$ is j_{th} atom in $\mathbf{D}$.

In general, traditional dictionary learning [19], [25], [29]–[31] is based on the following sparse representation function:

$$\min_{\mathbf{D}\in\mathcal{D},\mathbf{Z}} \frac{1}{n}\sum_{i=1}^{n}\{||\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_2^2 + \lambda||\mathbf{z}_i||_1\}, \tag{2}$$

where $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n]$ and each $\mathbf{z}_i \in \mathbb{R}^k$ encodes the ith input $\mathbf{x}_i$ with the dictionary $\mathbf{D}$; $||\mathbf{z}_i||_1$ is the sparsity regularization and λ is a penalty parameter. In such a way, each $\mathbf{x}_i$ is represented as a sparse linear combination over a set of basis vectors from $\mathbf{D}$ with its corresponding sparse coefficient $\mathbf{z}_i$.

Different from conventional dictionary learning with ℓ_2 minimization, we employ the ℓ_1 norm $||\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1$ to denote the loss between the input data and the sparse dictionary representation. The use of ℓ_1 norm enables the minimization of the loss term more robust to noises and outliers [23]. Simultaneously, we also aim to learn a projection matrix $\mathbf{P} \in \mathbb{R}^{p\times q}(p < q)$, which could preserve the energy of the training data as a robust and representative subspace. Moreover, it is helpful to reduce the computational cost by dimension reduction. To this end, we propose an online robust projective dictionary learning (RPDL) scheme, which is beneficial to exploit more effectively the representative information of training data via simultaneously learning the projection and dictionary. We formulate the RPDL for sparse representation by minimizing the following objective function:

$$\mathcal{L}(\mathbf{P}, \mathbf{D}, \mathbf{Z}) = \frac{1}{n}\sum_{i=1}^{n}\{||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1 + \lambda||\mathbf{z}_i||_1\}, \\ s.t.\ \ \mathbf{P}\mathbf{P}^T = \mathbf{I}, \mathbf{D} \in \mathcal{D}, \tag{3}$$

where the fitting term $||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1$ is illustrated in Fig. 3.

2) Optimization Framework: The traditional way to solve (3) involves steps to update $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{Z}$ iteratively with the batch learning procedure, assuming that all training data in $\mathbf{X}$ can be provided at one time. However, this assumption may not be satisfied when dynamic training instances arrive in sequence or change over time. In addition, batch learning approach could be infeasible for processing large-scale training data, because the computational cost for batch learning is dependent on the training data size. To address these issues, we propose an online algorithm that processes a mini-batch (or even one instance) of the training instances at a time, which is particularly useful for processing large-scale and dynamic data. Let h denote the size of mini-batch ($1 \leq h \leq n$) and l is the index for the coming mini-batch instances, the proposed online algorithm incrementally learns the $\mathbf{P}_l$, $\mathbf{D}_l$ and $\{\mathbf{z}_i\}_{i=(l-1)h+1}^{lh}$ given the newly input $\{\mathbf{x}_i\}_{i=(l-1)h+1}^{lh}$, based on the current best represented dictionary $\mathbf{D}_{l-1}$ and projection matrix $\mathbf{P}_{l-1}$.

Algorithm 1 Online Robust Projective Dictionary Learning

Input : Training data $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n]$, regularization parameter λ, and mini-batch training data size h.
Output: Dictionary $\mathbf{D}$ and projection $\mathbf{P}$.

```
Initialize: D_0, P_0
for l ∈ {1, ··· , L} do
    Initialize: D ← D_{l−1}, P ← P_{l−1}
    for i ∈ {(l − 1)h + 1, ··· , lh} do
        update z_i = arg min_z ||Px_i − Dz||_1 + λ||z||_1
    end
    update D_l by combining D_{l−1} and
    {z_{(l−1)h+1}, ··· , z_{lh}}, with fixed P_{l−1} ;
    update P_l by combining P_{l−1} and {z_{(l−1)h+1}, ··· , z_{lh}}
    with fixed D_{l−1};
    l ← l + 1
end
return D = D_l, and P = P_l
```

When processing each mini-batch training data, simultaneously solving $\mathbf{P}$, $\mathbf{D}$ and $\mathbf{Z}$ in (3) is difficult, because both the fitting term and the sparsity penalty are non-smooth and non-differentiable, hence resulting in no closed-form solution. To tackle this issue, we partition the whole optimization problem in (3) into two subproblems, and solve them by an iterative alternative direction method:

$$(\mathbf{D}^{t+1}, \{\mathbf{z}_i\}^{t+1}) = \arg\min_{\mathbf{D},\{\mathbf{z}_i\}} \mathcal{L}(\mathbf{P}^t, \mathbf{D}, \{\mathbf{z}_i\}) \tag{4a}$$

$$\mathbf{P}^{t+1} = \arg\min_{\mathbf{P}} \mathcal{L}(\mathbf{P}, \mathbf{D}^{t+1}, \{\mathbf{z}_i\}^{t+1}), \tag{4b}$$

where the superscript t denotes the iteration. Note that joint minimization of $\mathbf{D}$ and $\{\mathbf{z}_i\}$ in (4a) is a non-convex problem due to the coupled fitting term. We solve (4a) by applying the strategy similar to [32]. Dictionary learning updates $\mathbf{D}$ when fixes $\{\mathbf{z}_i\}$, and updates $\{\mathbf{z}_i\}$ when fixes $\mathbf{D}$. We summarize the whole steps for solving the optimization problem in (3) in Algorithm 1. The whole optimization is presented in detail as follows.

a) Update Sparse Coefficients $\mathbf{Z}$*:* Given the dictionary $\mathbf{D}$, projection matrix $\mathbf{P}$ and input data $\mathbf{x}_i$, we can derive the update of $\{\mathbf{z}_i\}$ by solving the following ℓ_1 measured and ℓ_1 regularized convex optimization problem:

$$\mathbf{z}_i = \arg\min_{\mathbf{z}_i} ||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1 + \lambda||\mathbf{z}_i||_1. \tag{5}$$

We rewrite (5) into an equivalent optimization problem as below:

$$\mathbf{z}_i = \arg\min_{\mathbf{z}_i} \left\| \begin{pmatrix} \mathbf{P}\mathbf{x}_i \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{D} \\ \lambda \mathbf{I} \end{pmatrix} \mathbf{z}_i \right\|_1, \tag{6}$$

which is a robust regression problem. Equation (6) cannot be solved in the same way as ℓ_2 norm regression due to its non-differentiability. Thus we resort to iterative reweighted least squares (IRLS) algorithm [33] to update it.

Let $\mathbf{b} = \begin{bmatrix} \mathbf{P}\mathbf{x}_i \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{(p+k)\times 1}$ and $\mathbf{A} = \begin{bmatrix} \mathbf{D} \\ \lambda \mathbf{I} \end{bmatrix} \in \mathbb{R}^{(p+k)\times k}$, and b_j denote the jth element of $\mathbf{b}$. The IRLS algorithm solves the following two problems iteratively:

$$\mathbf{z}_i \approx \arg\min_{\mathbf{z}} \sum_{j=1}^{p+k} \omega_{ij} (b_j - \mathbf{A}(j,:)\mathbf{z})^2, \tag{7}$$

$$\omega_{ij} = \frac{1}{\sqrt{(b_j - \mathbf{A}(j,:)\mathbf{z}_i)^2 + \delta}}, \tag{8}$$

where δ ($\delta = 1\text{e}{-6}$ in our implementation) is a small positive constant that prevents zero divisor. Each iteration of (7) involves minimizing a quadratic objective function. The global optimum can be reached by setting the derivative to be 0. This leads to solve $\mathbf{z}_i$ in the linear programming as:

$$[\sum_{j=1}^{p+k} \omega_{ij} b_j \mathbf{A}(j,:)^T] = [\sum_{j=1}^{p+k} \omega_{ij} \mathbf{A}(j,:)^T \mathbf{A}(j,:)]\mathbf{z}_i. \tag{9}$$

We can denote $\mathbf{s}_i \in \mathbb{R}^{k\times 1}$ as $\mathbf{s}_i = \sum_{j=1}^{p+k} \omega_{ij} b_j \mathbf{A}(j,:)^T$, and $\mathbf{C}_i \in \mathbb{R}^{k\times k}$ as $\mathbf{C}_i = \sum_{j=1}^{p+k} \omega_{ij} \mathbf{A}(j,:)^T \mathbf{A}(j,:)$. Then the sparse coefficients $\mathbf{z}_i$ can be updated as:

$$\mathbf{z}_i = \mathbf{C}_i^{-1} \mathbf{s}_i. \tag{10}$$

b) Update Dictionary $\mathbf{D}$*:* Given the projective input $\{\mathbf{P}\mathbf{x}_i\}_{i=1}^n$ and the sparse coefficients $\{\mathbf{z}_i\}_{i=1}^n$. If without the consideration of the constraint $\mathbf{D} \in \mathcal{D}$, the objective function can be expressed as:

$$\min \mathcal{L}(\mathbf{D}) = \frac{1}{n} \sum_{i=1}^{n} ||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1, \tag{11}$$

which is also an ℓ_1 regression problem. Different from [32], which stores intermediate parameters and treats all samples equally hence resulting in expensive computational cost and unstable optimization, we resort to stochastic gradient descent (SGD) algorithm [34] to update $\mathbf{D}$ iteratively. In each iteration, given the projective input, the gradient with respect to dictionary $\mathbf{D}$ can be derived as:

$$\nabla_{\mathbf{D}} \mathcal{L}(\mathbf{D}) = \frac{1}{h} \sum_{i=(l-1)h+1}^{lh} sgn(\mathbf{D}\mathbf{z}_i - \mathbf{P}\mathbf{x}_i)\mathbf{z}_i^T, \tag{12}$$

where $sgn(\cdot)$ is the sign operator. Then the update of $\mathbf{D}$ can be calculated by

$$\mathbf{D}^{t+1} = \mathbf{D}^t - \eta \nabla_{\mathbf{D}^t} \mathcal{L}(\mathbf{D}), \tag{13}$$

where η represents the learning rate.

Then considering the constraint $\mathbf{D} \in \mathcal{D}$, we conduct explicit scaling with

$$\mathbf{d}_j = \frac{\mathbf{d}_j}{max(1, ||\mathbf{d}_j||_2)}. \tag{14}$$

c) Update Projection $\mathbf{P}$*:* Given the dictionary $\mathbf{D}$ and the sparse coefficients $\{\mathbf{z}_i\}_{i=1}^n$, the objective function with respect to $\mathbf{P}$ is:

$$\begin{aligned} &\min \ \mathcal{L}(\mathbf{P}) = \frac{1}{n} \sum_{i=1}^{n} ||\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i||_1, \\ &s.t. \ \mathbf{P}\mathbf{P}^T = \mathbf{I}. \end{aligned} \tag{15}$$

If without the consideration of the orthonormal constraint on $\mathbf{P}$, the update of $\mathbf{P}$ can also be derived by SGD algorithm with

$$\mathbf{P}^{t+1} = \mathbf{P}^t - \theta \nabla_{\mathbf{P}^t} \mathcal{L}(\mathbf{P}), \tag{16}$$

where θ represents the learning rate and $\nabla_{\mathbf{P}} \mathcal{L}(\mathbf{P})$ is the gradient, which can be calculated with

$$\nabla_{\mathbf{P}} \mathcal{L}(\mathbf{P}) = \frac{1}{h} \sum_{i=(l-1)h+1}^{lh} sgn(\mathbf{P}\mathbf{x}_i - \mathbf{D}\mathbf{z}_i)\mathbf{x}_i^T. \tag{17}$$

To preserve the orthonormal constraint, we employ the Gram-Schmidt process to orthogonalize $\mathbf{P}$ with the following four steps: i) rewrite $\mathbf{P}$ as $\mathbf{P} = [\mathbf{p}_1; \cdots; \mathbf{p}_r; \cdots; \mathbf{p}_p]$ and initialize the first vector in the orthogonal subspace as $\mathbf{v}_1 = \frac{\mathbf{p}_1}{||\mathbf{p}_1||_2}$, ii) update $\mathbf{p}_r$ by projecting it onto the null space of $[\mathbf{v}_1; \mathbf{v}_2; \cdots; \mathbf{v}_{r-1}]$ by

$$\mathbf{p}_r = \mathbf{p}_r - \sum_{i=1}^{i<r} \mathbf{v}_i \mathbf{p}_r^T \mathbf{v}_i, \tag{18}$$

iii) obtain $\mathbf{v}_r$ by projecting $\mathbf{p}_r$ onto the unit sphere, iv) finally obtain $\mathbf{P} = [\mathbf{v}_1; \cdots; \mathbf{v}_r; \cdots; \mathbf{v}_p]$.

C. Robust Shape Estimation via Projective Sparse Representation

In practical, the query shape (i.e., segmented TRUS prostate shape) $\mathbf{y} \in \mathbb{R}^q$ may be partially unfaithful or corrupted resulting from the segmentation inaccuracy. Let $\mathbf{y}_0$ denote the ground truth of the query shape $\mathbf{y}$ and $\mathbf{e}$ denote the shape error, we have

$$\mathbf{y} = \mathbf{y}_0 + \mathbf{e}. \tag{19}$$

The nonzero entries of $\mathbf{e}$ model the shape error in $\mathbf{y}$. To sparsely code the partially inaccurate $\mathbf{y}$ over the dictionary, previous methods [18], [21] used to modify the sparse representation scheme by

$$\arg\min_{\mathbf{z},\mathbf{e}} ||\mathbf{y} - \mathbf{D}_0\mathbf{z} - \mathbf{e}||_2 + \lambda_1||\mathbf{z}||_1 + \lambda_2||\mathbf{e}||_1, \tag{20}$$

where $\mathbf{D}_0$ is the dictionary constructed by the original training data or the learned atoms from the original training data; $||\mathbf{e}||_1$ approximates the ℓ_0 norm to account the number of nonzero entries in $\mathbf{e}$; λ_1 and λ_2 are regularization parameters which control the sparsity of $\mathbf{z}$ and $\mathbf{e}$, respectively. Note that the shape error $\mathbf{e}$ can be different for various query shapes,

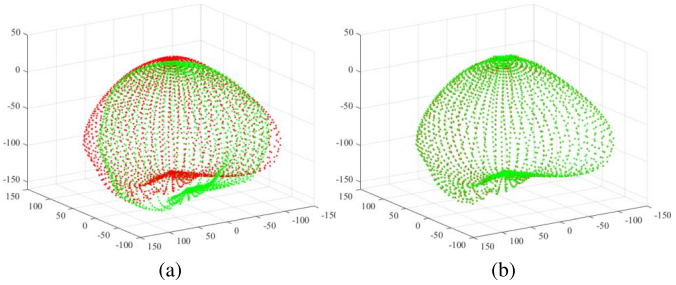


Fig. 4. The shape registration result on one patient dataset: (a) the TRUS shape (red) and MR shape (green) before registration, (b) registration using the projective sparse representation. It can be observed that the MR shape can be well registered to the TRUS shape.

and hence it is difficult to determine the free parameter λ_2 in advance. In addition, the iterative estimation for $\mathbf{z}$ and $\mathbf{e}$ is computationally expensive, especially for high dimensional query sample $\mathbf{y}$.

To address above issues, we leverage the learned subspace $\mathbf{P}$ and project the query shape $\mathbf{y}$ into a lower dimensional space by $\mathbf{Py}$. The subspace $\mathbf{P}$ is more representative than the original high-dimensional space since it is learned simultaneously with the dictionary and sparse representation given the training data. Therefore we implicitly depress the outliers and noises by the projective sparse representation scheme with

$$\arg\min_{\mathbf{z}} ||\mathbf{Py} - \mathbf{Dz}||_1 + \lambda||\mathbf{z}||_1, \tag{21}$$

which is similar to (5) and can be efficiently solved.

D. Implementation Detail

The MR images are firstly resampled to the same dimensionality and voxel size as the TRUS images. The MR and TRUS images are then segmented with the method [27] and [35], respectively, and further refined by an experienced clinician for the correctness assurance. Note that after segmentation, the TRUS prostate shape is constructed in a similar way as described in Section II-A. Thus a rough one-to-one point correspondence is obtained for both MR and TRUS prostate shapes. Then our RPDL method is performed to model the shape prior by jointly learning the projection and dictionary from a series of simulated TRUS shape instances (200 in our implementation). Once obtaining the optimal projection $\mathbf{P}$ and dictionary $\mathbf{D}$, the projective sparse representation (Equation (21)) is directly used to constrain the MR shape to register with the TRUS shape. Fig. 4 illustrates one typical shape registration result obtained from one patient dataset. After the shapes are registered, the volumetric displacement field inside the deformed shape is computed by interpolating the shape displacement via thin-plate spline (TPS) method [36], which can finally warp the MR images to the TRUS images.

In our RPDL algorithm, we employ PCA to initialize $\mathbf{P}$ and select the number of principal components as $p = 20$. We initialize $\mathbf{D}$ randomly and choose the number of atoms as $k = 100$. We set the size of mini-batch as $h = 10$, the penalty parameter $\lambda = 0.1$, the learning rates $\eta = 0.001$ and $\theta = 0.001$. For the stopping criterion of each subproblem, if the maximum iteration number (we set $maxIter = 100$) is reached or the difference between the objective value in adjacent iterations is smaller than a preset threshold $\epsilon = 1\text{e}{-6}$, algorithm stops and outputs the corresponding parameter. Thanks to the warm start of $\{\mathbf{D}, \mathbf{P}\}_{l-1}$, the update of $\{\mathbf{D}, \mathbf{P}\}_l$ for the online training converges in several iterations.

TABLE I
THE MAIN IDEA OF EACH COMPARED METHOD

Method	Main idea
PCA	dimension reduction
SSC	dictionary learning with sparse representation, where dictionary consists of original training data
DL	dictionary learning with sparse representation, where dictionary consists of learned atoms from training data
PCA+DL	independent dimension reduction and dictionary learning
Our RPDL	online joint dimension reduction and dictionary learning

III. Experiments and Results

A. Materials

Experiments were carried on the datasets obtained from eighteen patients with suspected prostate cancer at the First Affiliate Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China. The study protocol was reviewed and approved by the Ethics Committee of Sun Yat-Sen University and informed consent was obtained from all patients. The methods were carried out in accordance with the approved guidelines.

We acquired one set of MR and TRUS volumes from each patient. The T2-weighted MR images were acquired using a 3.0 Tesla Siemens TrioTim MR scanner (Erlangen, Germany) with a voxel size of $0.625 \times 0.625 \times 3.6$ mm^3. TRUS data were obtained by Mindray DC-8 ultrasound system (Shenzhen, China) with an integrated 3D TRUS probe. These data were then reconstructed into a TRUS volume with a voxel size of $0.5 \times 0.5 \times 0.5$ mm^3. For each patient, the MR scanning was performed several days ahead from the acquisitions of TRUS; both MR and TRUS images were acquired in a patient-centered coordinate system.[1]

B. Registration Performance

1) Compared Methods: To demonstrate the advantages of the proposed RPDL on shape modelling, we compared the RPDL with other four state-of-the-art shape modelling methods: PCA, sparse shape composition (SSC) [18], dictionary learning (DL) [19], and PCA+DL. The main idea of each compared method is listed in Table I, respectively. For each patient data, all the five methods were employed to model the TRUS prostate shape prior based on the same training data. Then the shape prior models were used to constrain the MR-TRUS registration.

[1]In patient-centered coordinate system, the coordinate axes are oriented in the same direction regardless of the patient's position on the scanner.

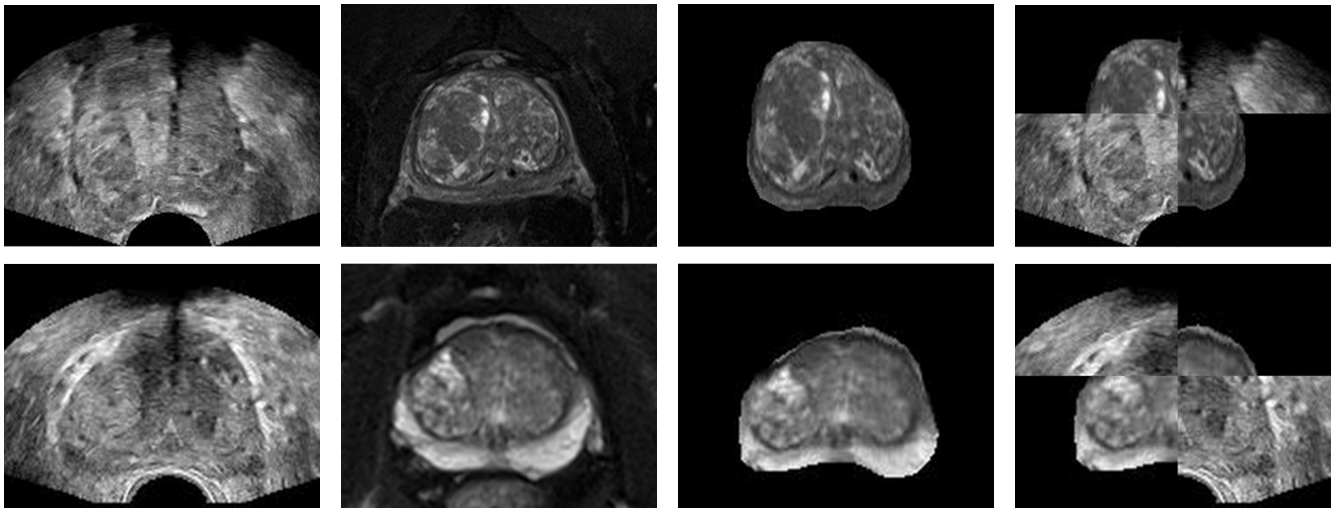


Fig. 5. The registration results of two patients obtained by the proposed RPDL method: (a) TRUS images, (b) MR images, (c) registered MR images, and (d) the checkerboard of (a) and (c) for comparison.

TABLE II
MEAN TRE AND REGISTRATION TIME FOR FIVE DIFFERENT METHODS

	TRE (mm)	Registration time (s)
PCA	2.24 ± 0.74	1.02e−3
SSC	2.13 ± 1.28	1.29
DL	1.84 ± 1.00	1.09
PCA+DL	1.85 ± 0.87	0.20
RPDL	1.51 ± 0.71	0.20

The target registration error (TRE) [37], defined as the Euclidean distance between corresponding intrinsic landmarks in MR and TRUS images, was measured to evaluate the registration performance. All the landmarks used for TRE calculation were first extracted by a urological physician with extensive experience in interpreting the prostate MR and TRUS images. The physician separately labeled each landmark (e.g., centres of small nodules, cysts, calcifications, etc.) in MR and TRUS images, and then correlated the distinctly corresponding landmarks as pairs. To ensure the anatomical consistency of landmark pairs, one more experienced physician was invited to make necessary refinement and confirmation of the landmark annotations. For each patient, 4-6 pairs of corresponding landmarks were extracted; and totally 94 pairs of landmarks were extracted from eighteen patients for the TRE calculation. Among the 94 pairs of landmarks, 20, 59 and 15 pairs were within the base, mid-gland and apex, respectively. The values of fiducial localization error (FLE) [38] were 1.26 mm and 0.92 mm for the landmarks obtained from MR and TRUS images, respectively. The low FLE values indicate the reliable registration evaluation.

2) Registration Accuracy: Fig. 5 visualizes two registration results obtained using the proposed RPDL method. Table II summarizes the numerical TRE for each of the five methods. The mean TRE was improved from the initial 6.78 mm, to 1.51 mm, obtained by the proposed RPDL method. This improvement proves the efficacy of our jointly learned projection and dictionary in representing the query shape. Specifically, our RPDL method yielded the TRE values of 1.42 mm, 1.46 mm and 1.81 mm within the base, mid-gland and apex zones, respectively. The higher TRE value in apex zone is mainly due to the large deformation in this zone caused by the insertion of the TRUS probe. As shown in Table II, the mean TRE values by performing the PCA, SSC, DL, PCA+DL methods are 2.24 mm, 2.13 mm, 1.84 mm and 1.85 mm, respectively. The TRE value by the proposed RPDL method is approximately 32%, 29%, 18% and 18% lower than the PCA, SSC, DL, PCA+DL method, respectively. These results demonstrate that the jointly learned projection and dictionary contribute to the improvement of the shape representation. And the DL method performs better than the SSC method due to the use of the more representative dictionary learned from the original training data. The PCA+DL and DL methods have similar TRE performances, which demonstrates that the dimensionality reduced data obtained before dictionary learning may not capture the most representative information of training data, and hence not quite effective for the improvement of dictionary learning.

To investigate the statistical significance of the proposed RPDL method over PCA, SSC, DL, and PCA+DL methods, a series of statistical analyses are conducted. First, the Bartlett's test [39] is used to analyze the homoscedasticity of variances for the TREs of different methods. The null hypothesis is not rejected, suggesting that the variances are statistically equal across all sets of TREs. Based on the homoscedasticity property, the one-way analysis of variance (ANOVA) [40] is further performed to evaluate if the TREs of different methods are statistically different. The resulting F $(= 6.70) > F_{critical}$ $(= 2.48)$ indicates that the differences between the TREs on 18 patients from the five methods are statistically significant. In addition, the two-sample, two-tailed t-test is employed to pairwisely compare

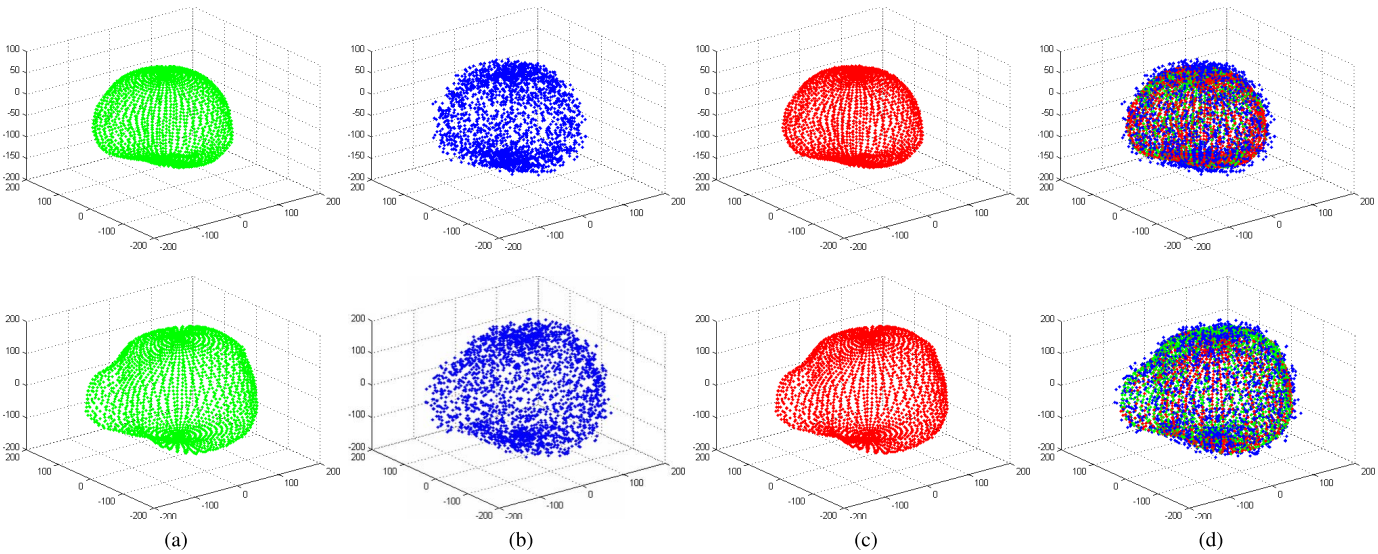


Fig. 6. Two shape representation results obtained by the proposed RPDL method on synthetic noisy query shapes: (a) original segmented TRUS prostate shapes, (b) the noise perturbed shapes of (a), (c) the corresponding RPDL represented shapes using (b) as query shapes, and (d) the comparison of (a), (b) and (c).

TABLE III
P-VALUES OF STUDENT TESTS BETWEEN DIFFERENT METHODS

***P*-value**	**SSC**	**DL**	**PCA+DL**	**RPDL**
PCA	0.38	1.02e−2	1.42e−3	1.01e−6
SSC	-	0.16	0.14	7.12e−3
DL	-	-	0.99	2.14e−2
PCA+DL	-	-	-	2.98e−3

the performances between the RPDL and the other four methods (see Table III). The resulting p-values of the pairs RPDL-PCA, RPDL-SSC, RPDL-DL and RPDL-PCA+DL are 1.01e−6, 7.12e−3, 2.14e−2 and 2.98e−3, respectively. It can be concluded that the null hypotheses for the four comparing pairs are not accepted at the 0.05 level. As a result, our method is significantly better than the other four compared methods. Table III further reports the p-values between the four compared methods. It can be observed from Table III that DL method is significantly better than the PCA method, which demonstrates the effect of the sparse representation based dictionary learning on accurate shape representation. In addition, the performance difference between DL and PCA+DL methods is not statistically significant, whereas the differences of RPDL-DL and RPDL-PCA+DL are significant, which proves the effect of joint projection and dictionary learning on effective shape representation.

For each patient data, the average time for each method to register MR and TRUS shapes is listed in Table II, respectively. On a workstation with Intel Xeon E5-1620 3.70 GHz CPU and 16.0 GB RAM, and with a MATLAB implementation, it takes averagely 0.20 s for our RPDL method to sparsely represent the query TRUS shape over the dimensionality reduced dictionary. The PCA+DL method achieves the similar time cost due to the same dimension reduction and sparse representation principle. Both SSC and DL methods are slower than the proposed RPDL method, because they perform sparse representation in higher dimensional dictionary space. The fastest method is the PCA method that conducts least-square regression in dimensionality reduced subspace.

C. Robustness to Inaccurate Query Shape

In practical, the query TRUS shape may be partially unfaithful or corrupted due to inaccurate segmentations. In such a case, any incorrect TRUS shape representation may probably degrade the performance of shape registration by conventional point matching methods [41], [42]. Fortunately, such outliers and noises could be implicitly depressed by our projective sparse representation scheme, which makes our method robust to partially inaccurate query shape. To illustrate the robustness of our method to partially inaccurate query shape, we conducted twofold experiments by validating the synthetic noisy shape and realistic noisy shape.

1) Synthetic Noisy Query Shape: We first perform extensive experiments by systematically perturbing the positions of points from the original query shape with positive and negative offsets for the MR-TRUS shape registration on all 18 patient data. For each patient, the originally segmented TRUS shape is perturbed by five different percentages (i.e., 20%, 40%, 60%, 80% and 100% randomly selected shape points) and five different amplitudes (i.e., ±(0-5), ±(5-10), ±(10-15), ±(15-20) and ±(20-25) pixels of original point position).

Fig. 6 visualizes the shape representation results obtained using our RPDL method on synthetic noisy query shapes. It can be observed from Fig. 6 that our projective sparse representation is able to robustly and faithfully represent the query shape despite its partial inaccuracy. Table IV summarizes the registration performances with respect to the changes of noise level of the query shape. Specifically, the mean and standard deviation of TREs are reported. It can be observed

TABLE IV
IMPACT OF NOISE LEVEL OF THE QUERY SHAPE ON REGISTRATION: TREs (MEAN±SD *in mm*)

Noise amplitude	Noise percentage				
	20%	**40%**	**60%**	**80%**	**100%**
±(**0-5**)	1.51 ± 0.75	1.55 ± 0.80	1.56 ± 0.91	1.70 ± 0.90	1.86 ± 1.38
±(**5-10**)	1.59 ± 1.00	1.60 ± 0.91	1.66 ± 0.95	2.09 ± 1.03	2.13 ± 1.47
±(**10-15**)	1.61 ± 0.78	1.69 ± 0.69	1.85 ± 0.85	2.36 ± 0.92	2.62 ± 1.81
±(**15-20**)	1.63 ± 0.71	1.75 ± 0.94	2.01 ± 0.97	2.71 ± 1.15	2.90 ± 1.50
±(**20-25**)	1.66 ± 1.09	1.82 ± 0.70	2.29 ± 1.23	3.17 ± 1.69	3.68 ± 1.88

TABLE V
IMPACT OF NOISE LEVEL OF THE QUERY SHAPE ON REGISTRATION: *P*-VALUES BETWEEN TREs OF NOISY AND ORIGINAL QUERY SHAPES

Noise amplitude	Noise percentage				
	20%	**40%**	**60%**	**80%**	**100%**
±(**0-5**)	0.87	0.73	0.86	0.33	0.21
±(**5-10**)	0.61	0.64	0.44	0.01	0.03
±(**10-15**)	0.50	0.21	0.07	4.75e−4	3.94e−3
±(**15-20**)	0.43	0.22	9.21e−3	1.24e−4	1.58e−4
±(**20-25**)	0.34	0.04	2.40e−3	1.25e−4	1.67e−5

from Table IV that the mean TRE value got worse with the increase of noise level, from 1.51 mm when 20% points were with ±(0-5) offsets, to 3.68 mm when all points were with ±(20-25) offsets. To further demonstrate the statistical significance between the registration results of using noisy query shape and original query shape, the *p*-values of student tests are reported in Table V. As shown in Table V, the performance differences of image registration are statistically insignificant when inaccurate points are i) within 20% and have offsets smaller than 25 pixels, or ii) within 40% and have offsets smaller than 20 pixels, or iii) within 60% and have offsets smaller than 15 pixels, or iv) have offsets smaller than 5 pixels. The statistical analyses shown in Table V demonstrate that our proposed RPDL could be regarded as an effective method for the robust representation of the query shape, despite its partially inaccurate segmentation.

2) Realistic Noisy Query Shape: We further invited a junior physician to manually segment all TRUS images and used the segmented TRUS shapes as query shapes to conduct MR-TRUS shape registration. Because this junior physician is without too much experience in interpreting the prostate TRUS images, the segmented TRUS shapes may be partially inaccurate. Fig. 7 illustrates one typical segmentation result obtained by the junior physician (blue contour), as well as the ground truth (green contour) for comparison. Fig. 8(a) shows the 3D comparison between the accurate TRUS shape (green) and the partially inaccurate TRUS shape (blue) obtained by the junior physician. As shown in Figs. 7 and 8(a), there existed some realistic gross errors in the TRUS shape obtained by the junior physician.

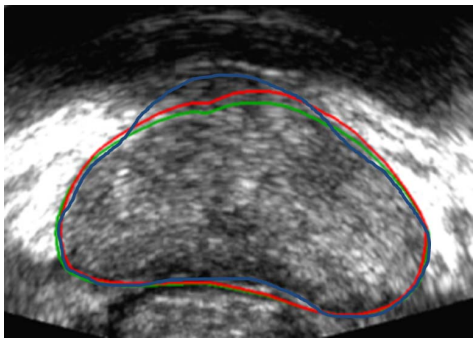
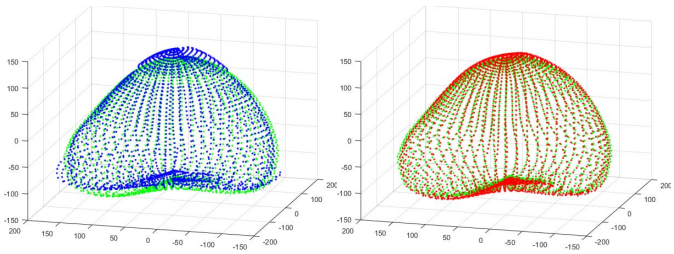


Fig. 7. The comparison results of TRUS contours. Green contour is the ground truth of TRUS prostate segmentation. Blue contour is the segmentation result by a junior physician. Red contour is the approximately represented TRUS shape using our method. Compared to the contour obtained by the junior physician (blue), ours (red) is more similar to the ground truth (green).



(a) (b)

Fig. 8. The shape representation result obtained by our method on realistic noisy query shape: (a) the comparison between the accurate TRUS shape (green) and the partially inaccurate TRUS shape (blue) obtained by a junior physician, (b) the comparison between the accurate TRUS shape (green) and our represented TRUS shape (red) using blue shape as the query shape.

By using the partially inaccurate TRUS shapes as query shapes to conduct MR-TRUS shape registration, the TRE value of our method degraded to 2.63 ± 1.73 mm, from 1.51 ± 0.71 mm obtained using accurate TRUS shapes. However, the mean TRE of 2.63 mm still largely improved the initial 6.78 mm TRE value. Furthermore, because our method is, to some extent, robust to partially inaccurate query shape, it is potentially useful for the refinement of the incorrectly segmented query shape. Figs. 7 and 8 further visualize the comparisons of the ground truth shape and our represented TRUS shape using inaccurate query shape. It can be observed from Figs. 7 and 8 that our projective sparse representation is

TABLE VI
QUANTITATIVE EVALUATION OF DIFFERENT TRUS SHAPES

Metric	Shape by the junior physician	Our represented shape
Dice(%)	0.87 ± 0.04	0.91 ± 0.03
ABD(mm)	3.24 ± 1.05	1.72 ± 0.55
HD(mm)	7.28 ± 2.68	4.00 ± 2.11

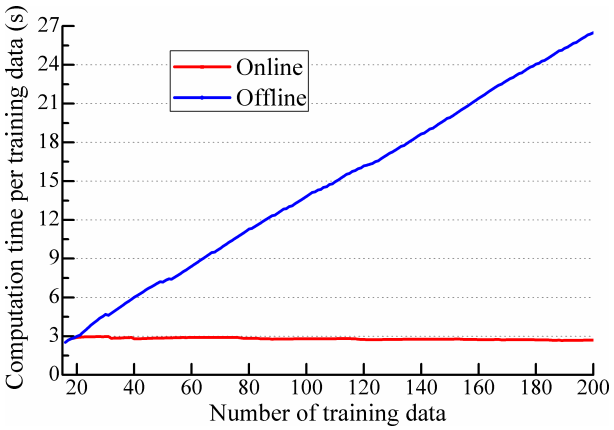


Fig. 9. Training time comparison between online and offline modes.

able to robustly represent the query shape despite its partial inaccuracy. Table VI further reports the quantitative evaluation of the shape by the junior physician and our represented shape. The Dice coefficient, average boundary distance (ABD) and Hausdorff distance (HD) [43] were calculated to evaluate the difference between ground truth and different TRUS shapes. As shown in Table VI, comparing to the shape obtained by the junior physician, our represented shape is more similar to the ground truth. Therefore the represented shape by our method could be regarded as a refinement of the inaccurate query shape.

D. Temptation of Online Mechanism

In real-world applications, the adequate number of training data usually cannot be collected all at once, also it is always better to add more newly collected data to enrich the training information, thus an efficient learning scheme is necessary for the optimization of training. We demonstrate the advantages of our online framework in terms of time and memory efficiency.

1) Training Time: Assume the training instance is collected and used for the training one by one, the computation time of the joint dimension reduction and dictionary learning with online and offline mechanisms is illustrated in Fig. 9. We can observe that the computation time for each training instance using our online setting keeps almost constant (less than 3 s), while the computation complexity of offline setting is growing linearly as the amount of training instances increases.

2) Memory Usage: Assume we process h training instances per time, when the lth mini-batch training instances arrive, our online method needs to store the newly arrived instances $\mathbf{X}_l = [\mathbf{x}_{h(l-1)+1}, \cdots, \mathbf{x}_{hl}]$, the sparse coefficients $\mathbf{Z}$, the dictionary $\mathbf{D}$ and the projection $\mathbf{P}$, which requires the memory of size qh, kh, pk and pq, respectively. For the updating of $\mathbf{Z}$, our online method needs to store the intermediate variables $\mathbf{A}$, $\mathbf{b}$ and weights $\mathbf{W}$, which needs $(p+k)k + (p+k) + (p+k)$ elements for each instance. For the updating of $\mathbf{D}$ and $\mathbf{P}$ with SGD algorithm, the online method needs to store the gradient $\nabla_{\mathbf{D}}\mathcal{L}(\mathbf{D})$ and $\nabla_{\mathbf{P}}\mathcal{L}(\mathbf{P})$, which takes the memory of size pk and pq, respectively. Finally, to preserve the orthonormal constraint of $\mathbf{P}$, the online method stores a finite and linear independent set $\mathbf{V}$, which supposes to store $(p-1)q$ elements. Thus our online method requires memory of size $[(q+k)+(p+k)(k+2)]h + 2pk + 3pq - q$.

For the offline method (here we take PCA+DL method as an example), when the Nth instance comes, it needs to save all the training instances $\mathbf{X}_N = [\mathbf{x}_1, \cdots, \mathbf{x}_N]$, which takes qN elements. For dimension reduction, we assume PCA uses a condensed SVD with p singular values, which supposes to save $pq + p^2 + pN$ elements. For the DL, it requires to store the dictionary $\mathbf{D}$, sparse atom sets $\{\mathbf{D}_\Gamma\}_{i=1}^N$ and sparse coefficients $\mathbf{Z}$, which occupies pk, pkN and kN elements, respectively. For the updating of dictionary, for each atom, it first stores an error matrix $\{\mathbf{E}\}_{j=1}^k$ and then uses SVD to find the largest singular vector, which takes memory of size $(pN + p + 1 + N)k$. Thus the PCA+DL method requires the memory of size $(p + q + 2pk + 2k)N + pq + p^2 + 2pk + k$.

In general, the comparison shows that the memory required by offline method scales with the training size N, while ours keeps constant.

IV. DISCUSSIONS

In this paper, a novel robust projective dictionary learning (RPDL) scheme for the shape prior modelling is presented. Specifically, we propose to jointly optimize the processes of dimension reduction and dictionary learning using a unified online framework. The joint learning is beneficial to make the projection and dictionary more appropriate with each other, and hence a more representative shape prior modelling of training data can be obtained. With the jointly learned projection and dictionary, the query shape can be robustly and efficiently represented based on projective sparse representation. We employ the proposed RPDL method to model prostate shape prior for the application of MR-TRUS registration. The experimental results suggest that the proposed method can achieve promising registration accuracy. The averaged target registration error obtained using our method is 1.51 mm, which meets the clinical requirement of 1.9 mm for the accurate identification of 95% targets with high Gleason score [44].

This work is an extension of our previous work [9]. The differences between [9] can be significant in twofold. First, the shape model in [9] was learned from the training data via PCA method in batch mode. In this study, the shape model is learned by the projective dictionary learning with online mechanism. Second, the registration schemes implemented in [9] and this study are distinctive. In [9], a hybrid scheme was adopted to constrain the registration by using both the learned shape model and the similarity of the modality independent neighborhood descriptor. In this study, the registration is simplified by using only the learned model to constrain the shape-based registration, thus reducing the computational

cost of the multi-modality similarity measure. Accordingly, the method in this work can be more general, efficient and clinically practical. In addition, this work is also different from [6] and [7], which rely heavily on the FEA to obtain MR-TRUS fusion. Our method only simply uses FEA [9] to prepare training shapes for the proposed projective dictionary learning and sparse representation.

To demonstrate the advantage of joint dimension reduction and dictionary learning for the shape prior modelling, we further compare our proposed method with other four state-of-the-art shape modelling methods: PCA, SSC [18], DL [19], and PCA+DL. As shown in Table II, the comparison results demonstrate that our method provides more accurate and robust MR-TRUS registration than state-of-the-art methods do. Such improvement is shown to be statistically significant in Table III. These satisfactory comparison results demonstrate that the proposed method is beneficial to exploit more effectively the representative information of training data via simultaneously learning the projection and dictionary, and the jointly learned projection and dictionary further contribute to the improvement of the shape representation. In addition to the advantages of robustness and accuracy, our proposed method also possesses fast processing speed. Our method only takes 0.20 s to sparsely represent the query shape over the dimensionality reduced dictionary. Accordingly, our method can achieve small registration error and fast registration time and can be more clinically practical.

One key limitation of conventional model constrained shape registration methods is that the query shape (i.e., the shape needs to be registered) may not be faithfully represented due to inaccurate segmentation, thus the estimated shape using learned prior model may not always be close to the real case. In the case of TRUS prostate segmentation, because of the low signal-to-noise ratio, TRUS usually falls short of clearly depicting the prostate boundary, resulting in a high variability of segmentation [45]. Due to the high variability of TRUS segmentation, the segmented TRUS shape may be partially unfaithful or corrupted, and hence degrading the performance of model constrained shape registration. Accordingly, the robust representation of TRUS shape despite its partial inaccuracy deems to be necessary in our shape registration problem. Different from conventional dictionary learning methods, we embed the ℓ_1 sparse data fitting term in our projective dictionary learning, which enables our RPDL method to handle the problem of partial inaccuracies of the query shape. Taking advantage of the proposed projective sparse representation, our method can implicitly depress noise in the query shape and further robustly and faithfully model this shape, as shown in Tables IV, V and VI. Therefore our method could help ease the TRUS segmentation process, and is more clinically practical and feasible.

Our proposed method is potentially useful for the refinement of the incorrectly segmented query shape, as shown in Figs. 6, 7 and 8. Clinicians can compare and analyze both the query shape and the shape obtained using our method, thus achieving a refined shape/segmentation. And thanks to our online learning mechanism, the refined shape could be used easily to enrich the training set for the further optimization of the projection and dictionary. In addition, our proposed online framework is beneficial for joint dimension reduction and dictionary learning in terms of memory and time efficiency. Thus our framework caters to the requirement of large scale streaming medical data processing.

V. Conclusion

This work presents a pioneer study that integrates the dimension reduction and dictionary learning into a unified online scheme for shape prior modelling. The effective shape modelling from a set of training data is difficult since the shape variation is complicated, and shape models should preserve local details as well as handle shape noises. To address this difficult problem, we propose a novel robust projective dictionary learning (RPDL) scheme for jointly optimizing the projection and dictionary, which is beneficial to exploit more effectively the representative information of training data. The joint learning is embedded in an online framework so that our scheme is of memory and time efficiency. The jointly learned projection and dictionary is able to robustly and efficiently represent the query instance based on projective sparse representation. We employ the proposed RPDL to model prostate shape prior for the practical application of MR-TRUS registration. The experimental results demonstrate that our proposed RPDL method provides more accurate and robust shape modelling than state-of-the-art methods do. The proposed RPDL method is a general solution for the problem of shape prior modelling, and also is potentially useful for the shape refinement.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," *CA Cancer J. Clin.*, vol. 66, no. 1, pp. 7–30, 2016.

[2] H. Hricak, P. L. Choyke, S. C. Eberhardt, S. A. Leibel, and P. T. Scardino, "Imaging prostate cancer: A multidisciplinary perspective," *Radiology*, vol. 243, no. 1, pp. 28–53, 2007.

[3] E. A. Klein, "Prostate cancer: MR-TRUS fusion biopsy—Defining a new standard," *Nature Rev. Clin. Oncol.*, vol. 12, no. 5, pp. 253–254, 2015.

[4] Y. Sun, J. Yuan, W. Qiu, M. Rajchl, C. Romagnoli, and A. Fenster, "Three-dimensional nonrigid MR-TRUS registration using dual optimization," *IEEE Trans. Med. Imag.*, vol. 34, no. 5, pp. 1085–1095, May 2015.

[5] S. Khallaghi *et al.*, "Statistical biomechanical surface registration: Application to MR-TRUS fusion for prostate interventions," *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2535–2549, Dec. 2015.

[6] S. Khallaghi *et al.*, "Biomechanically constrained surface registration: Application to MR-TRUS fusion for prostate interventions," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2404–2414, Nov. 2015.

[7] Y. Hu *et al.*, "Mr to ultrasound registration for image-guided prostate interventions," *Med. Image Anal.*, vol. 16, no. 3, pp. 687–703, 2012.

[8] Y. Hu, E. Gibson, H. U. Ahmed, C. M. Moore, M. Emberton, and D. C. Barratt, "Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration," *Med. Image Anal.*, vol. 26, no. 1, pp. 332–344, 2015.

[9] Y. Wang *et al.*, "Towards personalized statistical deformable model and hybrid point matching for robust MR-TRUS registration," *IEEE Trans. Med. Imag.*, vol. 35, no. 2, pp. 589–604, Feb. 2016.

[10] J. A. Onofrey *et al.*, "Learning non-rigid deformations for robust, constrained point-based registration in image-guided MR-TRUS prostate intervention," *Med. Image Anal.*, vol. 39, pp. 29–43, Jul. 2017.

[11] G. A. Sonn *et al.*, "Targeted biopsy in the detection of prostate cancer using an office based magnetic resonance ultrasound fusion device," *J. Urol.*, vol. 189, no. 1, pp. 86–92, 2013.

[12] M. M. Siddiqui *et al.*, "Comparison of MR/ultrasound fusion–guided biopsy with ultrasound-guided biopsy for the diagnosis of prostate cancer," *JAMA*, vol. 313, no. 4, pp. 390–397, 2015.

[13] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2002.

[14] S. Mika, B. Schölkopf, A. J. Smola, K.-R. Müller, M. Scholz, and G. Rätsch, "Kernel pca and de-noising in feature spaces," in *Proc. NIPS*, vol. 11. 1998, pp. 536–542.

[15] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K.-R. Müllers, "Fisher discriminant analysis with kernels," in *Proc. IEEE Signal Process. Soc. Workshop. Neural Netw. Signal Process. IX*, Aug. 1999, pp. 41–48.

[16] Y. Hu *et al.*, "Modelling prostate motion for data fusion during image-guided interventions," *IEEE Trans. Med. Imag.*, vol. 30, no. 11, pp. 1887–1900, Nov. 2011.

[17] Y. Wang, D. Ni, J. Qin, M. Xu, X. Xie, and P.-A. Heng, "Patient-specific deformation modelling via elastography: Application to image-guided prostate interventions," *Sci. Rep.*, vol. 6, Jan. 2016, Art. no. 27386.

[18] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, and X. S. Zhou, "Towards robust and effective shape modeling: Sparse shape composition," *Med. Image Anal.*, vol. 16, no. 1, pp. 265–277, Jan. 2012.

[19] S. Zhang, Y. Zhan, and D. N. Metaxas, "Deformable segmentation via sparse representation and dictionary learning," *Med. Image Anal.*, vol. 16, no. 7, pp. 1385–1396, 2012.

[20] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Trans. Image Process.*, vol. 17, no. 1, pp. 53–69, Jan. 2008.

[21] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[22] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, Nov. 2010.

[23] A. Wagner, J. Wright, A. Ganesh, Z. Zhou, H. Mobahi, and Y. Ma, "Toward a practical face recognition system: Robust alignment and illumination by sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 2, pp. 372–386, Feb. 2012.

[24] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Nov. 2011, pp. 471–478.

[25] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[26] S. Liao, A. K. Jain, and S. Z. Li, "Partial face recognition: Alignment-free approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 5, pp. 1193–1205, May 2013.

[27] F. Malmberg, R. Strand, J. Kullberg, R. Nordenskjöld, and E. Bengtsson, "Smart paint—A new interactive segmentation method applied to MR prostate segmentation," in *Medical Image Computing and Computer Assisted Intervention (MICCAI) Grand Challenge: Prostate MR Image Segmentation*, vol. 8. Berlin, Germany: Springer, 2012.

[28] T. Poston, T.-T. Wong, and P.-A. Heng, "Multiresolution isosurface extraction with adaptive skeleton climbing," *Comput. Graph. Forum*, vol. 17, no. 3, pp. 137–147, 1998.

[29] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, 2009, pp. 689–696.

[30] R. Rubinstein, A. M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proc. IEEE*, vol. 98, no. 6, pp. 1045–1057, Jun. 2010.

[31] J. Mairal, F. Bach, and J. Ponce, "Task-driven dictionary learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 791–804, Apr. 2012.

[32] C. Lu, J. Shi, and J. Jia, "Online robust dictionary learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 415–422.

[33] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.*, vol. 63, no. 1, pp. 1–38, 2010.

[34] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. COMPSTAT*, 2010, pp. 177–186.

[35] X. Yang *et al.*, "Fine-grained recurrent neural networks for automatic prostate segmentation in ultrasound images," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 1633–1639.

[36] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.

[37] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, Jr., "Predicting error in rigid-body point-based registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 694–702, Oct. 1998.

[38] R. R. Shamir, L. Joskowicz, S. Spektor, and Y. Shoshan, "Localization and registration accuracy in image guided neurosurgery: A clinical study," *Int. J. Comput. Assist. Radiol. Surgery*, vol. 4, no. 1, pp. 45–52, 2009.

[39] M. S. Bartlett, "Properties of sufficiency and statistical tests," *Proc. Roy. Soc. London. Ser. A, Math. Phys. Sci.*, vol. 160, no. 901, pp. 268–282, 1937.

[40] P. Armitage, G. Berry, and J. N. S. Matthews, *Statistical Methods in Medical Research*. Hoboken, NJ, USA: Wiley, 2008.

[41] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vis. Image Understand.*, vol. 89, nos. 2–3, pp. 114–141, Feb. 2003.

[42] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.

[43] G. Litjens *et al.*, "Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge," *Med. Image Anal.*, vol. 18, no. 2, pp. 359–373, 2014.

[44] W. J. M. van de Ven, C. A. Hulsbergen-Van de Kaa, T. Hambrock, J. O. Barentsz, and H. J. Huisman, "Simulated required accuracy of image registration tools for targeting high-grade cancer components with prostate biopsies," *Eur. Radiol.*, vol. 23, no. 5, pp. 1401–1407, 2013.

[45] W. L. Smith *et al.*, "Prostate volume contouring: A 3D analysis of segmentation using 3DTRUS, CT, and MR," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 67, no. 4, pp. 1238–1247, 2007.