Medical Image Analysis 44 (2018) 86–97

Contents lists available at ScienceDirect

Medical Image Analysis

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

A real-time and registration-free framework for dynamic shape instantiation

Xiao-Yun Zhou, Guang-Zhong Yang, Su-Lin Lee*

The Hamlyn Centre for Robotic Surgery, Imperial College London, London, UK

ARTICLE INFO

Article history:
Received 9 June 2017
Revised 16 October 2017
Accepted 22 November 2017

Keywords:
Image-guided interventions
Dynamic shape instantiation
3D shape recovery
KPLSR

ABSTRACT

Real-time 3D navigation during minimally invasive procedures is an essential yet challenging task, especially when considerable tissue motion is involved. To balance image acquisition speed and resolution, only 2D images or low-resolution 3D volumes can be used clinically. In this paper, a real-time and registration-free framework for dynamic shape instantiation, generalizable to multiple anatomical applications, is proposed to instantiate high-resolution 3D shapes of an organ from a single 2D image intra-operatively. Firstly, an approximate optimal scan plane was determined by analyzing the pre-operative 3D statistical shape model (SSM) of the anatomy with sparse principal component analysis (SPCA) and considering practical constraints. Secondly, kernel partial least squares regression (KPLSR) was used to learn the relationship between the pre-operative 3D SSM and a synchronized 2D SSM constructed from 2D images obtained at the approximate optimal scan plane. Finally, the derived relationship was applied to the new intra-operative 2D image obtained at the same scan plane to predict the high-resolution 3D shape intra-operatively. A major feature of the proposed framework is that no extra registration between the pre-operative 3D SSM and the synchronized 2D SSM is required. Detailed validation was performed on studies including the liver and right ventricle (RV) of the heart. The derived results (mean accuracy of 2.19 mm on patients and computation speed of 1 ms) demonstrate its potential clinical value for real-time, high-resolution, dynamic and 3D interventional guidance.

1. Introduction

Current clinical systems for minimally invasive procedures, such as cardiac radio-frequency ablation, image-guided needle biopsies, and endovascular interventions, typically incorporate static 3D surfaces for guidance. Real-time dynamic tracking of 3D surfaces can help to optimize the interventional procedure, especially for complex anatomical structures undergoing gross tissue deformation, bulk organ motion, and potential topological changes during interventions.

A combination of multiple imaging modalities has been used for dynamic 3D navigation. For example, a real-time registration scheme based on both spatial registration and electrocardiography was proposed to overlay pre-operative 3D magnetic resonance (MR) or computed tomography (CT) volumes onto intra-operative 2D ultrasound images for dynamic 3D navigation (Huang et al., 2009). 3D transesophageal echocardiography (TEE) was fused with 2D X-ray fluoroscopic images using image localization and calibration for dynamic cardiac navigation (Gao et al., 2012). However, based on a combination of multiple imaging modalities, the dynamic 3D shapes were either interpolated from pre-operative 3D volumes or intra-operatively collected 3D volumes with low-resolution. A 3D shape recovery scheme based on intra-operative 2D images including X-ray, ultrasound, and MR could take intra-operative information into account whilst achieving high-resolution at the same time. This kind of 3D shape recovery is termed dynamic shape instantiation. The scheme may or may not involve the use of template models (Filippi et al., 2008). Without template models used, more intra-operative information and longer image acquisition time are needed; for example, at least seven intra-operative 2D images were needed for reasonable 3D prostate reconstruction (Cool et al., 2006). In this paper, a single intra-operative 2D image is targeted and hence we focus on template-based 3D shape instantiation.

For template-based 3D shape instantiation methods, statistical shape model (SSM) (Frangi et al., 2002), free form deformation (FFD) (Koh et al., 2011), and Laplacian surface deformation (Karade and Ravi, 2015) can be used for the representation of templates. SSM (Cootes et al., 1995) is a popular technique which represents a set of 3D meshes or 2D contours with the same

* Corresponding author.
E-mail addresses: xiaoyun.zhou14@imperial.ac.uk (X.-Y. Zhou), g.z.yang@imperial.ac.uk (G.-Z. Yang), su-lin.lee@imperial.ac.uk (S.-L. Lee).

https://doi.org/10.1016/j.media.2017.11.009
1361-8415/

number of vertices and connectivities. SSM-based 3D shape instantiation learns from shape variations rather than only applying smoothness and 2D/3D similarity as the constraints. It deforms an initial 3D SSM to match intra-operative sparse inputs such as ultrasound-derived surface points (Barratt et al., 2008), digitized landmarks (Rajamani et al., 2007), or two or more calibrated X-ray images (Baka et al., 2011). These methods usually learn a model from a training set of anatomies of multiple patients and deform the learned model for a new patient, which requires a high anatomical similarity between patients. This learning is not suitable for patients with anatomical anomalies. For example, patients who have undergone liver resection have a significantly different liver shape to other subjects. A possible solution for these specific cases has been proposed in Lee et al. (2010). Here, limited optimal scan planes were determined by analyzing the pre-operative and patient-specific 3D SSM of the liver with principal component analysis (PCA). The relationship between pre-operative 3D SSM and synchronized 2D SSM constructed from 2D images at the optimal scan planes was learned by partial least squares regression (PLSR). Finally, with new intra-operative 2D images obtained at the same scan planes, the 3D shape was instantiated intra-operatively by applying the PLSR-derived relationship. However, in Lee et al. (2010), the optimal scan plane determination depended on the selected vertices that were deemed informative but were highly correlated and clustered. PLSR can only derive linear relationships while the deformations of most anatomies are non-linear. Based on Lee et al. (2010), a framework which achieves more accurate, robust, generalizable and convenient shape instantiations from a single intra-operative 2D image is proposed in this paper.

Subspace reprojection was proposed to determine an optimal scan plane for SSM-based 3D shape instantiation by fitting a plane to the most informative vertices (Lee et al., 2005). This optimal scan plane was shown to have enhanced accuracy compared to other scan planes(Lee et al., 2005). By applying PCA (Jolliffe, 2002) on the pre-operative 3D SSM, the informative vertices which contribute most to the shape variations are determined by the loadings of principal components (Lee et al., 2010). The downside of using PCA is that the derived principal components are linear combinations of multiple variables and therefore the selected informative variables are highly related and difficult to interpret. This phenomenon when reflected in our application is that the selected informative vertices are clustered and are not the real and independent informative vertices. Many methods have been proposed to solve this issue, including rotation methods (Jolliffe, 1995), limited set of integers (Vines, 2000), and simplified component technique least absolute shrinkage and selection operator (SCoTLASS) (Jolliffe et al., 2003). Simple thresholding is a common and informal method usually used in practice (Lee et al., 2010); however, this method lacks theoretical support and usually causes problems (Cadima and Jolliffe, 1995). Recently, Zou et al. proposed sparse PCA (SPCA) which reformulated PCA into a regression-type optimization problem and then added a L1 constraint to achieve sparse loadings; they demonstrated improved performance of SPCA in selecting the real informative variables over previous methods (Zou et al., 2006). A SPCA toolbox was later developed (AU: Please provide an update for reference "Sjöstrand et al.(2012), in press".[Sjöstrand et al.(2012)]Sjöstrand, Clemmensen).

PLSR is a linear regression method which has a similar prediction accuracy to ridge regression (RR) and principal component regression (PCR) (Frank and Friedman, 1993). It is more widely used than RR and PCR in medical problems, such as cardiac motion prediction (Ablitt et al., 2004) and craniofacial reconstruction (Duan et al., 2015), as it is more suitable for problems with a larger number of variables and fewer number of observations (Rosipal and Trejo, 2001). However, its accuracy for non-linear motions is limited.

Many non-linear PLSR variations have been proposed and they can be divided into two groups (Rosipal and Krämer, 2006): the first group reformulates the linear relationship into a non-linear one by polynomial functions, smoothing splines, artificial neural networks, and radial basis function networks while the second group maps the original variables into a higher dimensional space and regresses the mapped variables in the higher dimension, for example, kernel space. Kernel PLSR (KPLSR) (Rosipal and Trejo, 2001) from the second group is adopted in this paper for improved computation speed as its formulation is as time-efficient as PLSR and avoids the non-linear optimization in the first group.

In this paper, the high-resolution 3D shape of a dynamic anatomy was instantiated from a single intra-operative 2D image in real-time. Firstly, the anatomy was scanned by MR or CT pre-operatively for multiple 3D volumes along the dynamic cycle and a 3D SSM was constructed. SPCA was applied on the pre-operative 3D SSM to select the informative vertices which were used to fit an optical scan plane. Local adjustments of the scan plane parameters for better accessibility, visibility or satisfying other local constraints is possible without incurring major errors, as the later KPLSR-based 3D shape instantiation scheme is robust to optimal scan plane derivations. Secondly, 2D images synchronized with the pre-operative scanning were obtained at the approximate optimal scan plane and were sampled to generate a synchronized 2D SSM. KPLSR was applied to learn the relationship between the pre-operative 3D SSM and the synchronized 2D SSM. Finally, the high-resolution 3D shape was instantiated intra-operatively by applying the KPLSR-derived relationship onto a new intra-operative 2D image at the same scan plane. The overall framework of the proposed dynamic shape instantiation is illustrated in Fig. 1. Due to the learning of patient-specific models, the framework is applicable to any anatomy. No extra registration is needed for the pre-operative 3D SSM and the synchronized 2D SSM. Validation was performed on the liver (two digital liver phantoms, one dynamic liver phantom, one in vivo porcine liver, eight metastatic livers) and the cardiac right ventricle (RV) (18 asymptomatic RVs and 9 hypertrophic cardiomyopathy (HCM) RVs); we anticipate that potential applications of our work will include percutaneous liver biopsy, cardiac catheterization (Razavi et al., 2003), and intra-myocardial therapy (Saeed et al., 2005). For example, in cardiac ablation, the instantiated 3D RV shape can be used to help navigate the catheter tip to the target ablation area.

2. Methodology

The methods for determining the optimal scan plane are described in Section 2.1. The learning and instantiation based on KPLSR are described in Section 2.2. Finally, the data collection and detailed validation experiments are in Section 2.3.

2.1. Optimal scan plane determination

By pre-operatively scanning the target anatomy with CT or MR, a 4D volume consisting of multiple 3D volumes at different time frames along the dynamic cycle of the anatomy was obtained. These 3D volumes were represented with 3D meshes using the same number of vertices and connectivities, which created a pre-operative 3D SSM (a point distribution model) with vertices $Y_{N \times numY \times 3}$, where N is the number of time frames and $numY$ is the number of vertices. By rearranging the (x, y, z) coordinates of the vertices as independent variables, $Y_{N \times q}$ was obtained, where $q = numY \times 3$ is the number of variables. Without loss of generality, $Y_{N \times q}$ was centered and normalized as Y_{norm} with the mean and norm of each column as 0 and 1.