Computerized Medical Imaging and Graphics xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

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

From medical images to minimally invasive intervention: Computer assistance for robotic surgery

Su-Lin Lee, Mirna Lerotic, Valentina Vitiello, Stamatia Giannarou, Ka-Wai Kwok, Marco Visentini-Scarzanella, Guang-Zhong Yang*

Royal Society/Wolfson Foundation Medical Image Computing Laboratory, and the Institute of Biomedical Engineering, Imperial College London, London SW7 2AZ, United Kingdom

ARTICLE INFO

Article history:
Received 13 February 2009
Received in revised form 26 June 2009
Accepted 17 July 2009

Keywords:
Minimally invasive surgery
Medical imaging
Image guidance
Feature tracking
Augmented reality
3D reconstruction
Active constraints
Haptic feedback
Robotic assisted surgery

ABSTRACT

Minimally invasive surgery has been established as an important way forward in surgery for reducing patient trauma and hospitalization with improved prognosis. The introduction of robotic assistance enhances the manual dexterity and accuracy of instrument manipulation. Further development of the field in using pre- and intra-operative imaging guidance requires the integration of the general anatomy of the patient with clear pathologic indications and geometrical information for preoperative planning and intra-operative manipulation. It also requires effective visualization and the recreation of haptic and tactile sensing with dynamic active constraints to improve consistency and safety of the surgical procedures. This paper describes key technical considerations of tissue deformation tracking, 3D reconstruction, subject-specific modeling, image guidance and augmented reality for robotic assisted minimally invasive surgery. It highlights the importance of adapting preoperative surgical planning according to intra-operative data and illustrates how dynamic information such as tissue deformation can be incorporated into the surgical navigation framework. Some of the recent trends are discussed in terms of instrument design and the usage of dynamic active constraints and human–robot perceptual docking for robotic assisted minimally invasive surgery.

1. Introduction

With recent advances in imaging and surgical instrumentation, the field of surgery is entering a time of change towards safer, more consistent and minimally invasive intervention. The use of preoperative and intra-operative imaging augmented by mechatronically enhanced or robotic assisted instruments has significantly improved the perceptual-motor capabilities of the surgeon, allowing surgical procedures to be carried out with unprecedented accuracy and efficiency [1]. Despite these advances, technical challenges remain. Issues related to *in situ* 3D reconstruction, biomechanical modeling under large scale tissue deformation, recreation of tactile sensing and feedback, and instrument dexterity are some of the major obstacles to be tackled.

When performing open surgery, the surgeon operates directly on the patient with relatively unrestricted visual, force and tactile feedback. By combining the current tissue morphology with prior knowledge of the anatomical model, experienced surgeons can carry out the required procedure with ease, albeit often at a cost of large incisions and access trauma. In minimally invasive surgery (MIS), the small incision used for gaining surgical access has contributed to reduced trauma, shorter patient recovery time, and improved prognosis. However, the rigid instrument used, coupled with the fulcrum effect has resulted in poor hand–eye coordination and manual dexterity, contributing to a series of ergonomic and safety issues. Such is the nature of technological–clinical translational cycles that as the technology advances, a new clinical frontier becomes feasible and in its pursuit, a new set of technical challenges arise. With the advent of robotic assisted MIS, dexterity is enhanced by microprocessor controlled mechanical wrists, allowing for motion scaling for reduced gross hand movements and the performance of micro-scale tasks that are otherwise not possible [2]. Currently, the surgical community is tackling even greater challenges – surgery with no scar based on Natural Orifice Transluminal Endoscopic Surgery (NOTES). While the debate of its feasibility for safe, routine clinical use remains fervent, it is necessary to examine recent technological advances in imaging, modeling, mechatronics, robotics, and human–robot interaction to envisage the future directions of MIS.

Traditionally, surgical procedures are planned out using medical images reviewed off-line. Image guidance in surgery provides *in situ* visualization of either preoperative or intra-operative data.

* Corresponding author at: Royal Society/Wolfson Foundation MIC Laboratory, Department of Computing, 180 Queen's Gate, Imperial College London, London SW7 2AZ, United Kingdom. Tel.: +44 20 7594 8441; fax: +44 20 7581 8024.
E-mail address: g.z.yang@imperial.ac.uk (G.-Z. Yang).

0895-6111/$ – see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.compmedimag.2009.07.007

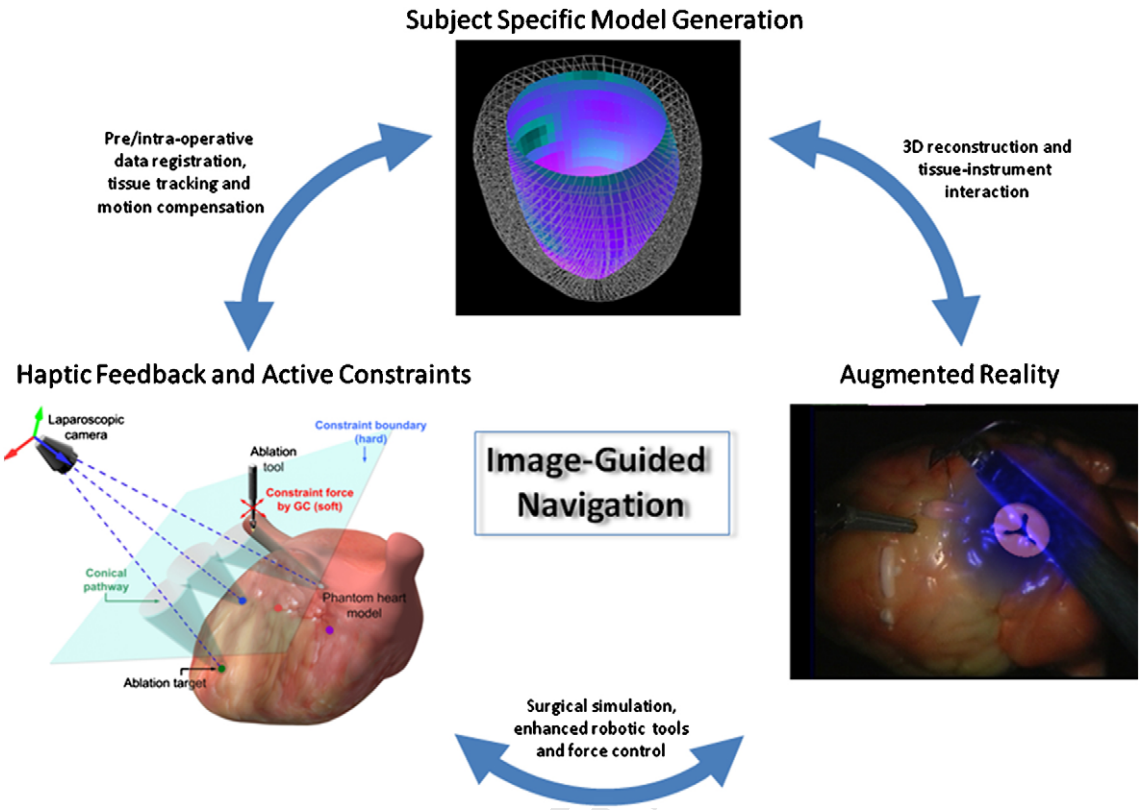

Fig. 1. A flowchart showing the key steps involved in image-guided surgical navigation. Subject-specific model generation, augmented reality, haptic feedback and active constraints all play important roles but their integration into a single platform presents many challenges.

However, real-time image guidance requires accurate registration of the medical images to the patient anatomy during surgery. Without apparent deformation or morphological changes, registration can be performed preoperatively and remain constant during surgery. If large tissue deformation is present, as in the case of cardiac and gastrointestinal surgery, it is necessary to provide on-line registration and remodeling of the data, also by taking into consideration potential topological changes of the tissue during surgery. Thus far, cardiac surgery remains a challenging problem for image guidance due to its complex structure and motion. To help illustrate the key steps involved in image-guided surgery, Fig. 1 shows a flow chart of the steps for an image-guided MIS framework, from acquisition of images through to robotic assistance, with the incorporation of haptic feedback and dynamic active constraints during surgery.

Typically, the first step of an image-guided surgical navigation system consists of gathering relevant preoperative data. In the case of robotic assisted MIS, it is necessary to combine three sources of information: data describing the general anatomy of the patient with clear pathologic indications, a kinematic model of the robot, and geometrical information for co-registration of preoperative planning and simulation of alternative intra-operative approaches. In using this information, additional constraints also need to be taken into account. For example, for robotic assisted MIS, remote centre of motion of the instrument at the trocar must be satisfied. The second step of surgical guidance is adapting the preoperative surgical plan according to intra-operative data. To this end, dynamic information such as tissue deformation and topological change, is important. These can be used for co-registration of pre- and intra-operative data and facilitate the implementation of augmented reality (AR) for mapping imaging data to the operative field-of-view. If advanced features such as virtual fixtures (active constraints) are to be incorporated, haptic boundaries are defined and updated at this stage. Although not yet available routinely in the clinic, procedure simulation can also be provided to assess the potential effect of the surgical step before its execution. The overall goals of image-guided surgery are therefore to provide a fully planned and rehearsed procedure before its execution; integrate real-time intra-operative imaging for enhanced accuracy and surgical vision; track and model deformation to adapt to anatomical changes during surgery; incorporate robotic assistance for enhanced dexterity and motion control; and recreate haptic and tactile sensing with dynamic active constraints to improve the consistency and safety of the surgical procedures. It is evident that the pre-requisite of achieving these goals is to accurately track the deformation of the tissue such that motion adapted visualization and control can be applied. In the following sections, we will describe key technical considerations of tissue deformation tracking, 3D reconstruction, subject-specific modeling, image guidance and augmented reality. We will also present some of the recent trends in instrument design and in using dynamic active constraints and human–robot perceptual docking for robotic assisted MIS.

2. Tissue deformation tracking during surgery

In addition to facilitating pre- and intra-operative image integration, tissue deformation tracking in robotic assisted MIS also serves the purpose of visual servoing for motion stabilization. In robotic assisted beating heart TECAB (Totally Endoscopic Coronary Artery Bypass), for example, this provides virtual immobilization of the heart and forgoes the need of mechanical stabilizers. In general, the landmarks to be tracked can be divided into two main categories, natural and artificial landmarks. Natural landmarks are prominent local features on the tissue surface, whereas artificial landmarks (also known as *fiducials*) are external markers with distinctive shape or color placed *in vivo*. These landmarks pro-

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

vide a stable frame of reference for making image registration possible.

2.1. Tracking of artificial landmarks

In image-guided intervention, it is common to use artificial landmarks, particularly in established clinical systems. An initial study on motion compensation of the beating heart in robotic assisted interventions has been performed by Nakamura et al. [3]. They proposed a high speed visual servoing system which uses a 4-DOF (degrees of freedom) robotic finger to track fiducials placed on the epicardial surface. *In vivo* experiments performed on porcine data showed good 2D trajectory tracking. Ginhoux et al. [4] used active optical markers rather than fixed fiducials to measure cardiac deformation. By decoupling cardiac and respiratory induced deformation, an adaptive model for predictive control in visual servoing was established. The performance of the method was evaluated with *in vivo* porcine data and demonstrated that the model can cancel out physiological motion with a residual tracking error up to 1.5 mm. The main drawback of this approach is that due to the large distance between the fiducials, the recovered motion is heavily interpolated across the surface. Sauvée et al. [5] addressed this issue by combining fiducials with surface textures. The calibrated fiducials allowed for 3D pose estimation of the laparoscopic camera whereas the surface texture permitted denser motion recovery.

Dey et al. [6] used fiducial markers to register freehand 2D endoscopic images to preoperative 3D computed tomography (CT) models of the brain. A marker based system for referencing the vertebrae using an endoscope was presented by Thoranaghatte et al. [7]. Also, Siddique and Jaffray [8] used fiducials in the thorax during respiration under X-ray fluoroscopy. They proposed an estimator of the location of the landmarks based on a particle filter approach that estimates their location and the associated spatial uncertainty. In terms of consistency, the tracking of soft tissue based on artificial landmarks is attractive in that it gives full control of the placement of the fiducials. However, it can also interfere and significantly complicate the actual surgical procedure itself. To circumvent this problem, these markers often need to be placed away from the operating site, thus diminishing its accuracy and practical value.

2.2. Tracking of natural landmarks

With the maturity of computer vision techniques, the use of natural landmarks, such as vessel junctions and surface textures, has attracted significant interest in recent years. Existing research has shown that direct application of existing tracking techniques in computer vision to MIS has significant problems, largely due to free-form tissue deformation and contrastingly different visual appearances during changing surgical scenes. The main difficulty involved is in the consistency of the natural landmarks as they deform with the surrounding structure, a situation that is not encountered when using rigid, high-contrast external fiducials. Desirable properties of an ideal MIS feature tracking technique include high repeatability under rotation, translation, scaling and affine transformation, as well as robustness to illumination changes and local deformations. To this end, Giannarou et al. [9] presented a comprehensive study on feature detection in soft tissue tracking. They introduced an affine invariant feature detector based on anisotropic features for reliable MIS feature tracking in the presence of laparoscopic camera motion and tissue deformation. Experimental results on *in vivo* MIS data showed that the anisotropic detector performed favorably as compared to existing state-of-the-art techniques and the method was capable of detecting a greater number of corresponding features in the presence of significant geometric and photometric image transformations and tissue deformation. Example results from the technique are shown in Fig. 2.

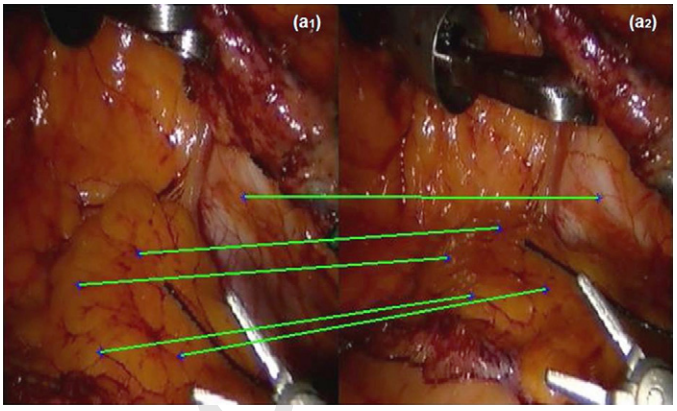

Fig. 2. Example image frames taken from an *in vivo* MIS image sequence with significant tissue deformation coupled with illumination changes: (a1) the first frame of the footage and (a2) a subsequent frame of the same sequence that presents a large degree of global motion. Example matched points between the two image frames are highlighted in blue and the green lines show the point correspondences extracted using an affine anisotropic feature detector [9]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.) Q5

Mountney et al. [10] evaluated the performance of state-of-the-art feature descriptors in computer vision when applied to soft tissue tracking in MIS. A novel probabilistic framework was developed for selecting the most discriminative descriptors for consistent deformation tracking. The accuracy and temporal persistency of the tracking approach was further enhanced by employing a Bayesian fusion method. Stoyanov et al. [11] introduced an approach for computing 3D structure and motion from a set of sparse salient features on the soft-tissue surface with a pre-calibrated stereo laparoscope. The temporal behavior of the extracted features was detected with stereo-temporal constraints by using an iterative registration algorithm. In order to improve the robustness of the algorithm in detecting the landmarks, Ortmaier et al. developed a prediction algorithm based on ECG and respiratory signals [12]. For the evaluation of the proposed approach, the tracking of the detected landmarks was restricted in a mechanically stabilized area of the beating heart.

With endoscopic procedures exploring the natural cavities of the patient, the paucity of salient landmarks makes region-based rather than feature-based tracking methods more suitable for motion tracking. Groeger et al. [13] recovered heart motion by a region-based tracking using natural epicardial features. An affine motion model was used to track local regions on the tissue, showing model trajectories strongly correlated with the dynamic characteristics of the cardiac and respiratory motion. The limitation of this approach is that motion is only recovered in the 2D image space and the ability of the algorithm to recover 3D measurements has not been investigated. For sinus examination, Burschka et al. [14] also used region-based tracking for intra-operative registration directly from monocular endoscopic images.

In order to increase the robustness of region-based tissue tracking approaches in the presence of large motion and variability between successive image frames, motion models have been employed. Hager and Belhumeur [15] proposed region tracking combined with linear 2D motion models. The method integrated both geometric and photometric cues based on an illumination model. To make the tracking robust to occlusions, they also developed a regression technique whereby occlusions were treated as outliers. In a different approach, a model-free visual servoing method was introduced by Bourger et al. [16] based on the Efficient Second-order Minimization (ESM) tracking algorithm. In

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

order to tackle occlusions, a pixel color selection method based on histograms matching was applied prior to tracking. The method performed the ESM tracking with a 4-DOF surgical robot by incorporating a vision-based robot control strategy, showing the robustness of the method even in the presence of significant occlusion. The main drawback in using the ESM tracking algorithm is its high computational complexity.

3. 3D deformation recovery

The above step of tissue deformation tracking only involves tracking distinctive features within the image plane. For recovering 3D surface deformation, it is necessary to combine these features with the relative configuration of the camera(s) and constraints imposed by the biomechanical properties of the soft tissue. To this end, techniques traditionally employed in computer vision applications have found limited success in MIS applications. Due to complex deformation of the tissue, the research effort on 3D reconstruction in MIS has been shifted from transposing methods developed for generic environments with strong *a priori* constraints to developing solutions specifically tailored to the properties of endoscopic environments. Such systems are characterized by weak constraints, an emphasis on stochastic modeling, and strict runtime requirements. These can be grouped according to the visual cues employed during the reconstruction phase, which include stereo vision, photometric stereo and structure from motion.

3.1. Visual cues for 3D deformation reconstruction

In general, stereoscopic methods for 3D reconstruction have the advantage of not relying on assumptions of the image formation model beyond the inherent geometric relationships between the stereo cameras. Thus far, stereo laparoscopes are part of the standard equipment for robotic assisted MIS systems. The major hurdles faced by stereo based deformation recovery in MIS are related to feature detection and matching, with the calculated disparity between features matched in the left and right visual channels being sought for the 3D triangulation in a calibrated system. In these tasks, the average density of the detected features dictates which areas can be reconstructed, whereas their distinctiveness determines the ambiguities encountered during the feature matching process. Traditional methods [17] rely on either global smoothness constraints or adaptive support windows during the feature matching process. For MIS, global smoothness constraints are unsuitable for dealing with separate organs and occlusions due to surgical tools. More recently, the success of Markov Random Fields (MRFs) for dense stereo reconstruction has attracted interests in the computer vision community but its real-time application to 3D tissue deformation has yet to be established [18,19].

In MIS images, the homogeneity of visual features is manifested in a general decrease both in the number and distinctiveness of detected features. As a result, stereo approaches will initially yield at best a semi-dense surface. In order to yield dense reconstructions, the most common approach is to impose local smoothness constraints given a sparse set of strong features in different areas of the MIS image. In cardiac surgery, for example, Lau et al. [20] used a B-spline representation of the disparity map for the reconstruction of the epicardial surface surrounding the coronary artery, allowing a compact algebraic representation and real-time performance of the algorithm. Richa et al. [21] used thin-plate splines (TPS) made of three-stacked Radial Basis Functions (RBF) in order to represent a target epicardial surface area under projective transformation. The surface is characterized by its pre-defined and tracked control points, and the dense stereo reconstruction problem is reduced to finding the optimal warping parameter of the TPS between the left and the right stereo channels. Finally, Stoyanov et al. [11] employed a Normalized Cross-Correlation (NCC) measure to match features detected in both channels with a Maximally Stable Extremal Regions (MSER) detector, thus achieving desirable properties such as rotational and monotonic illumination invariance. The matched features are then tracked with the Lucas–Kanade (LK) tracker extended by including epipolar constraints.

While the aforementioned approaches have shown to be suitable for recovering the motion of small regions and an approximation of their structure, the smoothness constraints introduced make them less suitable for a complete reconstruction of the entire field-of-view [21]. Moreover, their implicit dependency on the tracker performance makes them particularly susceptible to occlusions and specular highlights. An approach addressing the issue of global dense 3D reconstruction has been proposed recently by Stoyanov et al. [22], where the feature matching process was reduced to a 1D search problem through planar rectification of the stereo image pair, followed by a constrained dense disparity registration using hierarchical Piecewise Bilinear Maps (PBM). While achieving a dense reconstruction of the complete field-of-view, the hierarchical approach causes early mismatches to be propagated through all remaining scales, making it particularly sensitive to specular highlights.

Other approaches addressing the reconstruction problem while avoiding dependency on image features include photometric methods such as Shape-From-Shading (SFS). The original algorithm of SFS stemmed from the seminal work of Horn [23] as a way of reconstructing the shape of an object given its reflectance properties and the measured perceived brightness. The problem was originally expressed as finding the unique solution of a non-linear first-order Partial Differential Equation (PDE) resulting from an image formation model consisting of orthographic cameras with a light source located at infinity and perfectly Lambertian surfaces. One of the first attempts to find practical applications for SFS in MIS was presented by Rashid and Burger [24], who formulated a direct algorithm for SFS under perspective projection with the light source located at the optical centre. The approach is based on a linearization of the SFS equation and its locality allows for efficient parallel implementation; however, it requires very smooth surfaces in order for the assumed linearity to hold. A different approach to the same environment was presented by Deguchi and Okatani [25], who demonstrated more practical results from the GI tract. However, the proposed method did not attempt to solve the PDE, rather it evolved an initial level set made of points located at the same distance from the projection centre. Apart from practical difficulties in finding such points, with initial inaccuracies propagated to the subsequent stages of the reconstruction process, this method like most others suffered from the inability of solving the Bas-Relief ambiguity or disambiguating between concave and convex structures.

More recently, mathematical advances in viscosity solution theory [26] allowed for the development of methods performing well in realistic conditions for endoscopic images. Tankus et al. [27] extended the fast marching formulation of the SFS problem by Kimmel and Bruckstein [28] to the perspective case, showing good results with data extracted from a GI tract endoscopy sequence. While the practical applicability of the method suffered from (1) requiring the exact locations of all points of local depth minima in the input image (manually labeled in the study); and (2) the assumption of an infinitely distant light source, it demonstrated the potential of the SFS approaches. Yuen et al. [29] further extended the method by providing a faster single-pass solution. At the same time, Prados et al. [30], provided a control formulation for the solution of the SFS problem based on the notion of discontinuous viscosity solutions using realistic assumptions and showed that the problem can be well-posed in the case of a light source at the optical centre and perspective projection, a scenario well suited to endoscopy

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

applications, thus removing the need of providing values for critical points in order to resolve ambiguities.

Despite these recent developments, SFS can suffer from several shortcomings. First, they require input images to represent objects with a uniform albedo, a requirement reflected in the usage of test images from the GI tract. Also, almost all proposed methods tend to be iterative, with run-times still not suitable for real-time implementation. Furthermore, the underlying assumption of Lambertian surface characteristics makes them sensitive to specular highlights frequently found in MIS. Most importantly, these methods do not provide absolute depth information, as non-Lambertian surfaces, errors in the light source intensity and surface albedo estimation result in an unknown scale factor between the reconstructed surface and the ground truth.

3.2. Depth cue fusion

For MIS, since none of the single modalities of depth reconstruction can provide consistent results, fusion of multiple depth cues becomes important. Multi-cue fusion schemes have seen a recent increase in popularity among the computer vision community in an attempt to improve the overall robustness of the algorithms. This is designed to emulate the structure of the human vision system, conjectured to follow a Bayesian inferencing mechanism [31]. Among recent studies, multiple cues have been fused for feature tracking using graphical models [32], depth reconstruction from texture and stereo cues [33], and SFS ambiguity resolution using texture cues [34].

For example, Lo et al. [35] proposed a Bayesian Network (BN) structure for fusing together stereo and SFS cues. While the complexity of the inferential process does not allow the performance to be in real-time, and the method requires an initial global alignment between the cues used, it represents a first attempt towards global dense 3D reconstruction. While the use of cues such as texture is not appropriate to MIS applications, other cues have been used towards this goal. Given the surgeon's fixation points obtained with eye-tracking equipment, Stoyanov et al. [36] combined ocular vergence with stereo information to infer depth in relatively textureless areas of the epicardial surface for focused energy ablation. While the application is not dense depth recovery of the entire field-of-view, it provides the surgeon with depth information from any given fixation point without requiring interaction with additional equipment. For surgical navigation, image guidance is necessary for *in situ* visualization of intra-operative data. While laparoscopic video depicts the organ surfaces to the surgeon, it gives no indication of the relation of internal structures to each other. Building subject-specific anatomical models combined with 3D surface deformation derived above can assist in the visualization and simulation of tissue-instrument interaction.

4. Subject-specific model generation

4.1. Pre- and intra-operative image acquisition

One of the early applications of surgical planning is based on pre-operative imaging and reconstruction. Current imaging modalities such as magnetic resonance imaging (MRI), CT, X-ray fluoroscopy and ultrasound are popular in this field. More recently, biophotonic techniques including OCT (optical coherence tomography), DOI (diffuse optical imaging), and CFLM (confocal laser microscopy) have found their way to be integrated with endoscopy for *in situ*, *in vivo* cellular level imaging. This represents the general trend in surgical technology, as we drive towards more complex procedures but with reduced invasiveness, effective use of pre- and intra-operative imaging ensures a fully rehearsed procedure before its execution and safe intra-operative navigation during its performance.

The complex motion of the heart makes image-guided cardiac surgery particularly challenging. An understanding of the underlying mechanics governing myocardial contraction is essential for the interpretation and prediction of changes induced by heart disease as well as assisting with on-line registration. In this regard, cardiovascular MR (CMR) imaging has taken a key role in preoperative simulation as it provides a non-invasive way to assess the intramural motion of the myocardium. CMR can detect gross changes in contractile behavior of the myocardium but the increased sensitivity of scanners today allows for the detection of subtle regional changes in the myocardium during early stages of cardiac dysfunction.

The study of myocardial deformation and contractility has also been possible with MR tagging, DENSE (displacement encoding with stimulated echoes), and MR velocity imaging [37]. MR tagging [38] is an imaging protocol that uses a grid of magnetic saturation, produced by a sequence of radio frequency pulses, on the myocardium, providing landmarks that are tracked through time. However, there are significant limitations on the technique, in the form of tag fading, intra-slice motion, and the complicated post-processing steps involved in tracking the landmarks, but it is still a reliable means of determining the intramural motion of the myocardium. SENC (strain encoding) imaging [39] is a related technique whereby strain is directly measured in the MR images without separate measurement of displacement or velocity. DENSE [40] is a method for myocardial displacement imaging that is performed by the manipulation of the spin phase via stimulated echoes; the phase of each pixel is modulated based on its position. Large displacements can be detected since the method can work over a long time interval but image acquisition is time consuming. Phase contrast velocity imaging [41,42] provides a velocity vector at each voxel and has traditionally been used to visualize blood flow within the heart but blood flow artefacts, limitations on velocity sensitivity and low SNR have caused difficulties in analysis. Recent work on improved pulse sequence design with blood saturation to limit the blood flow artefacts has brought its use to the myocardium, where lower velocities are expected. Huntbatch et al. [43] also developed a Bayesian motion recovery framework for the restoration of the velocities within the myocardium. Reliable contractility information can be derived from this data and direct cardiac modeling is possible. Fig. 3 displays MR images of the heart alongside phase contrast velocity images of the left ventricle as well as results of the Bayesian motion recovery framework. For imaging of the material properties, diffusion tensor imaging [44,45] has been used to investigate the underlying myocardial fiber orientation. By augmenting this data with strain rate from velocity mapping, the relationship between fiber orientation and fiber shortening may be elucidated. In the figure, the directions of the velocities show the radial thickening, circumferential twisting and longitudinal contraction as expected in a normal left ventricle.

4.2. Building 3D shape models

From the comprehensive imaging data, the immediate step before surgical navigation and simulation is to build 3D models. For cardiac surgery, detailed segmentation of the myocardium is not trivial due to the complexity of the anatomy. Manual segmentation is time consuming and impractical and much research has been directed to optimal, fully/semi-automatic techniques. In practice, segmentation can be expedited through the use of Statistical Shape Models (SSM) [46], which characterize the morphology and dynamics of the heart across a set of subjects. They are built from training sets of shapes and principal components analysis (PCA) and when applied to the set, can give information about the range of shapes. A training set requires a number of shapes with points in correspondence and this is especially difficult when dealing with 3D shapes.

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

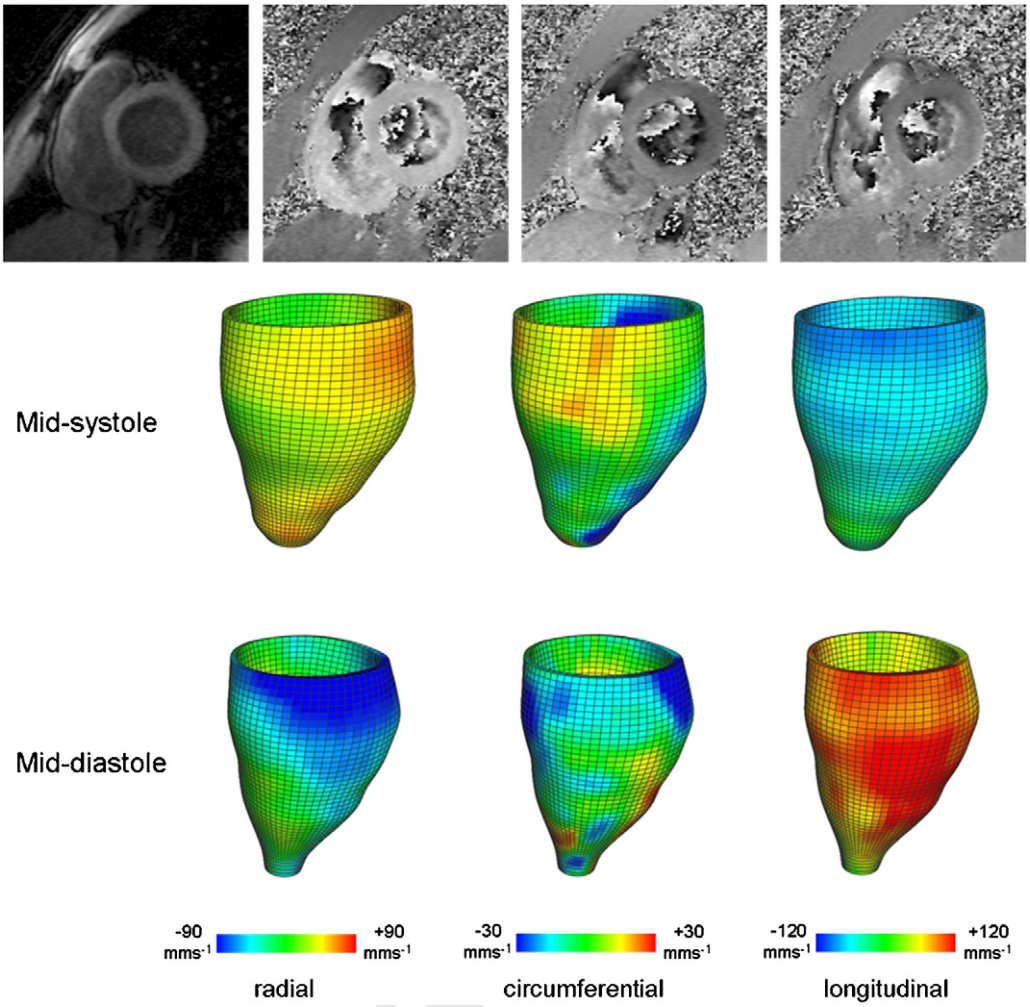

Fig. 3. (Top) Short axis images of the left ventricle of a normal subject and the associated velocity images along the *x*, *y*, and *z* directions. (Bottom) Radial, circumferential, and longitudinal velocities of the left ventricle of the same subject, at mid-systole and mid-diastole, after Bayesian motion recovery [43].

Horkaew and Yang [47], for example, proposed an automatic technique for establishing the correspondences based on harmonic mapping of the surfaces of the myocardium. The Minimum Description Length (MDL) criterion was incorporated into the objective function to find the model with the most compact description of the shape variation – this indicates that the best correspondence has been found and all spurious variation removed. The technique was demonstrated on cardiac surfaces [47] and pelvic floor models [48,49]. A cardiac probabilistic atlas for the automatic segmentation of 4D cardiac images was proposed by Lorenzo-Valdés et al. [50]. The use of an atlas provides prior shape knowledge of the heart and can potentially allow for better model initialization. These techniques, however, require excessive computation time and are only suitable for off-line analysis.

Recently, Lekadir and Yang [51] have developed a robust and fully automatic segmentation technique with active shape model search. The method is based on global geometric constraints during feature point search by using inter-landmark conditional probabilities and was demonstrated on the left ventricle. Ecabert et al. [52] also presented an automatic model-based method for segmentation of the whole heart from CT images. Automatic localization of the heart is performed through a generalized 3D Hough transform and a global similarity transform. After alignment, a deformable adaptation is used to match the cardiac mesh to the boundaries in the images. It should be noted, however, that while the use of a statistical model can assist in the segmentation of the heart, building the training set from a large enough set of patients is a challenge.

4.3. Biomechanical models

For MIS, accurate, subject-specific models can greatly benefit both surgeon and patient during surgical planning, navigation, and simulation. 3D shape models, while being able to provide information about the morphology of the anatomy in question, cannot inform us as to the physical properties of the anatomy during deformation. A physical-based, or biomechanical, model, used to integrate the structural and functional information available, is required to fully simulate the deformation that the anatomy undergoes. Ideally, the model would be subject-specific, created with minimal operator intervention, able to provide realistic simulation of the motion and deformation, and able to deal with tissue–instrument interaction. Photo-realistic rendering and accurate haptic feedback would also make simulations, either preoperatively or intra-operatively, more lifelike. In a similar vein, computational fluid dynamics (CFD) can be used to further the understanding of the relationships between morphology and flow in cardiac surgical planning. For the heart, computer models of the vessels and cardiac chambers can be used to simulate blood patterns. Boundary conditions of the inflow and outflow and a dynamic volumetric mesh are both required and the flow equations

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

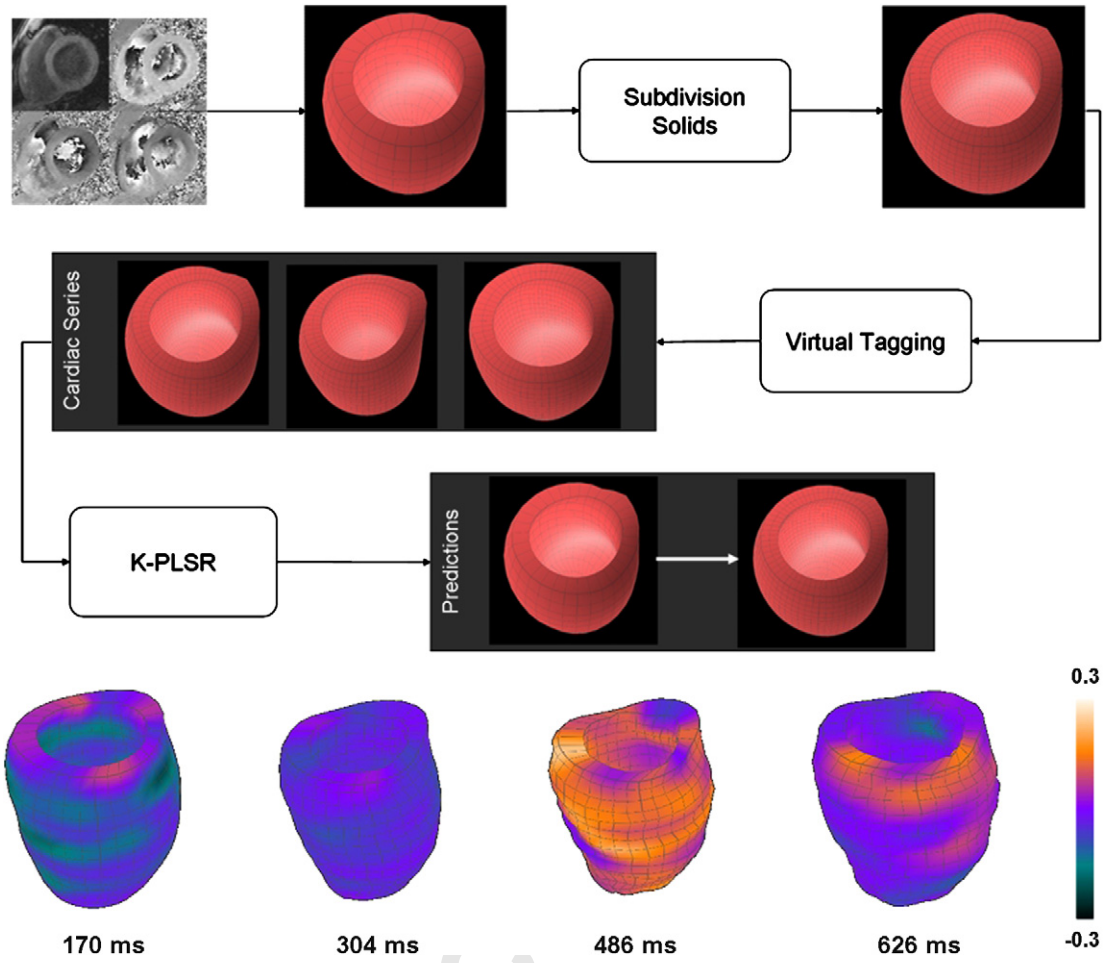

Fig. 4. (Top) A flowchart illustrating the process of K-PLSR to expedite biomechanical based modeling of the left ventricle. (Bottom) The derived longitudinal strain results from the left ventricle of a normal subject over the cardiac cycle.

are solved for each element within the mesh. CFD simulation is a promising technique for the investigation of how surgical changes to the anatomy can result in hemodynamic changes.

However, traditional biomechanical models require much *a priori* data that is difficult to obtain on a per subject basis. Parameters such as material properties and fiber orientation are crucial to the accuracy of the simulation but are difficult to acquire *in vivo*. It is therefore common for a general model to be developed and then mapped to each subject. In cardiac surgery, direct modeling of myocardial deformation is possible, for example, through virtual tagging [53], which combines the advantages of both MR tagging and velocity mapping. No *a priori* data is required as a grid is superimposed on the phase contrast velocity images and deformed based on the underlying velocity vectors. Segmentation is still needed for the epicardial and endocardial walls in the short axis slices though. The objective function to be minimized calculates the difference between the real and simulated velocity vectors and a volume conservation constraint is applied to ensure physically meaningful results.

For rapid modeling of contractile behavior of the myocardium, Lee et al. [54] introduced the concept of a Kernel-Partial Least Squares Regression (K-PLSR) motion modeling scheme. Initial modeling was applied to a coarse mesh of the anatomy and K-PLSR was used to predict a fine mesh of the same timeframe. The regression is trained with a set of corresponding coarse and fine meshes, the latter created by using volume subdivision to the coarse contour points. The chosen mesh also facilitates the calculation of cardiac strain in the radial, circumferential and longitudinal directions. The use of K-PLSR greatly reduced computation time of a virtual tagging analysis for a fine mesh. The flowchart representing this method, along with an example of 3D virtual tagging results on the left ventricle of a normal subject, can be seen in Fig. 4. The colormap indicates the amount of longitudinal strain in the left ventricle, characterized by left ventricle shortening and lengthening, over a single cardiac cycle. Longitudinal strain is difficult to model with other imaging modalities due to the in-plane motion of the heart during image acquisition. The method was further extended to incorporate the ordinary Kriging estimator for improved estimation of the simulated velocity vectors at the mesh points [55]. Direct contractility modeling can be advantageous to the modeling of cardiac dynamics, which would aid in the simulation of dynamic surgical procedures.

5. Image guidance and augmented reality

For image-guided surgery, a key requirement is the augmentation of the exposed surgical view with pre- or intra-operatively acquired images or 3D models. The inherent challenge is accurate registration of pre- and intra-operative data to the patient, especially for soft tissue where there is large deformation. With the increasing use of intra-operative imaging techniques, AR platforms are emerging in clinical environments [56]. Laparoscopic procedures are especially well suited to AR applications because the view of the operating scene is observed on surgical consoles. Traditionally preoperative data had to be reviewed off-line before the operation for preoperative planning or during surgery for guidance.

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

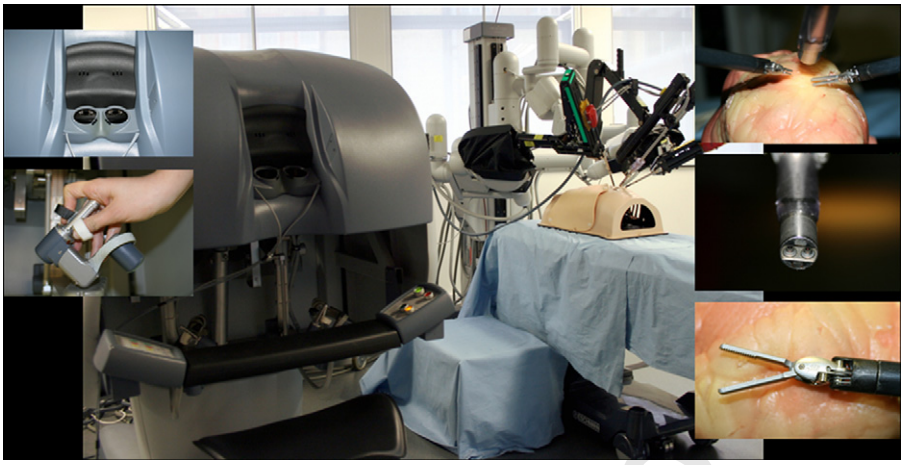

Fig. 5. The da Vinci surgical robot for minimally invasive procedures. The surgical console shown top left provides stereo views of the operating field through the use of the stereo laparoscope shown right middle, thereby providing 3D depth information to the surgeon. Master–slave endo-wrist control (left bottom) allows for intuitive motion of the instrument (right bottom) and removes the fulcrum effect present in traditional MIS.

AR in surgery [57] can provide *in situ* pre- or intra-operative data registered to the surgical field-of-view.

Early work by Bajura et al. [58] on "live" ultrasound echography data visualization within a pregnant human subject was achieved by using a small video camera mounted in front of a conventional head-mounted display worn by an observer. Abolmaesumi et al. [59] presented a robotic assistant for ultrasound based tracking and visual servoing that compensates for patient motion. Ultrasound image and AR guidance were also used for off-pump, closed, beating intracardiac surgery [60]. The method was tested on porcine models for implanting a mitral valve prosthesis via a universal cardiac introducer. The method displayed the entire prosthesis and tools without artefacts, thus providing intuitive navigation, positioning, and orientation of surgical instruments.

Preoperative data used for AR can also provide accurate subject-specific models displayed to the surgeon on demand during surgery. While being well established in neuro and orthopedic surgery, its application to cardiac and gastrointestinal surgery is still in its infancy due to the difficulty of handling large tissue deformation. Registration of preoperative cardiac images to the beating heart requires real-time tissue motion correction. For this purpose, Rickers et al. [61] developed an approach for MR guided transcatheter closure of atrial septal defects based on real-time MR fluoroscopy to track inducer sheath movement.

With the development of stereo consoles for robotic assisted surgery, such as the da Vinci robotic surgical console as shown in Fig. 5, the surgeon can regain depth perception. However, overlaying virtual objects onto the 3D scene is not straightforward [62,63]. With traditional AR, which uses surface transparency overlays, virtual objects can appear to float above the scene even though rendered at the desired depth. The reason for this is that our depth perception depends on occlusions and motion parallax. For example, if one object occludes another, it is perceived to be above it. Bichlmeier et al. [64] developed a virtual mirror that relies on motion parallax to provide accurate depth perception. The method is intuitive but requires explicit extraction of the 3D tissue geometry, and optical tracking of both the patient and the surgeon.

In order to address the issue of visual fidelity, an AR approach by Lerotic et al. [65] called *inverse realism* has been developed to provide see-through vision of the embedded virtual object while maintaining salient anatomical features of the viewed surface. The features are accentuated and occlude the virtual object therefore providing realistic depth perception in the scene. The method is based on non-photo-realistic rendering of *pq* values [66]. The *pq* values represent surface gradients of the exposed anatomical structure and the information about the surface is used to enhance salient features on the surface, while rendering smooth background regions semi-transparently. This type of rendering is perceived as "X-ray vision" where the exposed surface around the virtual object becomes transparent to provide a view through the surface.

With this technique, local surface orientation is provided by the *pq* values. The smooth background is attenuated and made transparent. Salient anatomical features are sharpened according to their *pq* values and kept non-transparent. These two images are combined together using a mask and a color look-up table. The resulting translucent rendering as shown in Figs. 6 and 7, has occlusions needed for accurate depth perception while providing views of the embedded object. It has been shown that inverse realism can provide accurate depth perception for AR in 3D surgical displays and the technique works equally well in stereo consoles and

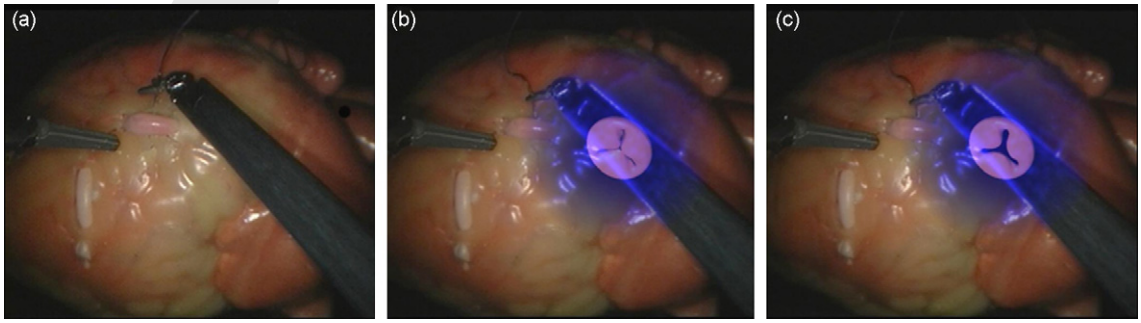

Fig. 6. (a) A phantom heart as seen through a da Vinci laparoscope. (b and c) AR visualization with inverse realism showing a superimposed artificial heart valve in relation to the cardiac structure.

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

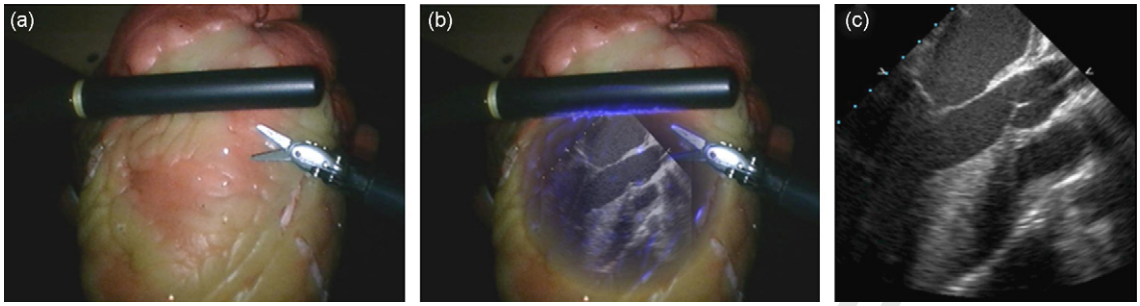

Fig. 7. (a) Laparoscopic ultrasound probe (Aloka UST-5550) used on a phantom heart, as seen through the da Vinci laparoscope. (b) Augmented view using the inverse realism method, where ultrasound images (c) are shown *in situ*.

head-mounted displays. Moreover, it removes the need for explicit extraction of the 3D tissue geometry, thus facilitating its real-time execution.

6. Improved manipulator/end-effectors for surgical navigation

The standard laparoscopic configuration used in robotic MIS features two or three manipulator arms and an laparoscope. Optimal entry points must be selected from a set of pre-defined admissible ports. However, due to the fixed port points, the robotic manipulators can only reach a restricted region inside the body. This kinematic constraint given by the fulcrum point on the patient's skin incision must be considered when developing a robot-assisted MIS system. To determine the relationship between the entry points and the operative region within the body, preoperative images are also necessary for optimal port placement. Furthermore, to access regions that are out-of direct line of sight of the surgeon, articulated instruments are essential.

6.1. Port placement

Optimal port placement requires a system to consider all components of the MIS together. For example, systems such as V-CAB [67] and STARS [68] have been developed to optimize port placement for robotic assisted MIS. Using interactive visualization and manipulation of the simulated thoracic ports in 3D within the reconstructed thoracic region, the optimal position of the ports can be determined. However, it should be noted that because of carbon-dioxide insufflation during MIS coupled with respiratory and cardiac motion, the positions of the targeted regions on the heart may change significantly compared to preoperative data.

6.2. Instrument design and articulation

The design of robotic tools for MIS is not only affected by the fixed entry ports on the body of the patient but must also respect several requirements: a relatively small size; an adequate number of DOF to keep contact with organs at a minimum and avoid dangerous regions; good velocity and force resolution; and ample workspace, in order to be able to reach all points of a given trajectory with the desired orientation. The conflicts between these contextual, mechanical and control constraints make the design of a dexterous instrument challenging. Systems such as MARGE [69] are focused on the design and control of enhanced dexterity devices for complex motion assistance in MIS, especially for coronary artery bypass grafting (CABG). A novel control algorithm based on the dynamic decoupling of the control torque in task and posture behavior control has been proposed [70]. The Steady-Hand Robot [71] has been designed to extend the surgeon's dexterity in microsurgical applications. The robotic tool is attached to a serial, six DOF manipulator arm providing smooth, tremor-free, precise positional control and force scaling. Recently, a number of research groups have focused on the development of micro-robots for enhanced control and navigation for robotic assisted MIS. For example, the MICRON project [72] focused on the design of a hand-held active manipulator to compensate for tremor during ophthalmological microsurgery. The three-DOF parallel robotic manipulator features piezo-electric actuators and a 6 DOF sensing system with inertial or optic sensors. A TEC (Tethered Epicardial Crawler) system has also been developed for minimally invasive beating heart surgery [73].

Several teams have also developed manually driven prototypes aimed to improve the accuracy of surgical manipulation of tissue, mainly by augmenting the surgeon's tactile or haptic sensing. D'Attanasio et al. [74] described a new endoscope for integration with a computer-assisted arthroscopy system. The device contained a cable-actuated steerable tip and sensors to detect the position of the tip and contact with the surrounding tissues are embedded. In addition, the steering mechanism can also be servo-controlled by the surgeon. The primary characteristic of this mechatronic tool is the semi-automatic collision avoidance to prevent the tip from touching dangerous regions detected prior to the surgery.

6.3. Surgical navigation for NOTES

Recently, there has been significant interest in pushing the frontiers of MIS to NOTES. There are three major justifications for the development of NOTES – better cosmetic appearance, ease of access, and the stimulus for a technological advancement that could enable a diminution of the pain and discomfort associated with traditional surgery [75]. With NOTES, the 3D reconstruction of the deforming scene can provide depth information that would benefit the surgeon but the limited field-of-view can make intra-operative navigation particularly difficult. In order to enhance the global awareness and spatial orientation, Lerotic et al. [76] described a dynamic view expansion scheme to provide enhanced visual cues that can aid NOTES navigation. Mechatronically, the maneuverability and stability of the instruments can be fulfilled by incorporating hyper-redundant joints with an articulated probe to provide the necessary flexibility to navigate inside complex structures, such as the GI tract, while avoiding tears and perforations. It is also important that the device can be locked in the desired position when the tool reaches the operating area. It is evident that navigation in a NOTES environment can greatly benefit from the incorporation of intra-operative imaging techniques. The imaging and visualization techniques described above will therefore play a role in the future development of NOTES navigation platforms.

7. Haptic feedback and active constraints

With image-guided surgery, incorporating haptic feedback into the master control interface can provide surgeons with the required

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

perceptual information for navigation and optimal force application. The human haptic sense can be divided into two separate sensory channels [77]: kinaesthetic force and cutaneous tactility. Kinaesthetic perception refers to the sensations of positions, velocities, forces, constraints, and inertia, which all actively resist contact motion. Cutaneous stimulation can be further classified into the sensations of pressure, stretch, vibration, geometry, roughness, slippage, and temperature, all of which arise through direct contact with the skin surface. However, compared to the kinaesthetic force feedback, it is still not practical for cutaneous tactility to be adopted in tele-manipulation robot systems. Not only is the present technology limited to cutaneous tactility on the fingertip but it is also very difficult to integrate a tactile display into many master control units because of its considerable size and weight and complicated mechanical requirements.

Recently, Okamura [78] performed psycho-haptic research to evaluate the influence of haptic perception on human sensory and motor capabilities for several surgical tasks. The experiments have characterized the problems with the deficiency of haptic feedback in tele-manipulation surgery. Pezzementi et al. [79] demonstrated the effect of haptic feedback on performing soft tissue suturing showing that it can reduce the number of failures, the completion time, and the magnitude of tension force required. In practice, it is difficult to obtain haptic sensing data *in vivo* during robot-assisted MIS. During surgery, the instruments are interacting with many objects such as the trocar, tissue, and other surgical instruments. Multiple miniaturized force sensors need to be embedded in the instrument requiring practical considerations of biocompatibility and sterilization [80]. These are some of the reasons why the current state-of-the-art surgical robotic systems, including da Vinci, have not yet successfully integrated the contact haptics between instruments and tissue during surgery. Because of this, the research attention has naturally shifted from haptic sensing to rendering in tele-operated surgery. Recent advances in haptic rendering technology enable collision detection and the computation of the interaction forces and simple tactile forces between modeled instruments and organs. In MIS, the simulation of realistic haptics is faced with many challenges [81]. To render in real-time with dynamic feedback, the computational cost required for both graphic and haptic rendering of high-fidelity deformable tissue models is a major bottleneck. Updates for haptic rendering have to be maintained at a rate above 1 kHz for perceptively continuous force sensing.

Although patient-specific models can be achieved accurately with advanced medical imaging technologies, empirical investigation of the *in vivo* tissue mechanics still requires improvement. As mentioned earlier, the lack of real material parameters is the main barrier to representing the inherent non-linear, anisotropic and rate-dependent behavior of soft tissue. There has been much investigation into the derivation of subject-specific material properties directly from preoperative medical images and the incorporation of this knowledge into MIS systems would control the amount of tactile feedback to the surgeon, thus providing a more realistic experience.

7.1. Active constraint control based on preoperative data

Early work on haptic rendering was focused on active constraints. One well-known example of a haptic-enabled active constraint robotic system is the Acrobot hands-on surgical system [82,83]. It features a force feedback handle positioned close to the robot end-effector. The surgeon can differentiate between the cutting of hard or soft tissue via the handle. The method relies on preoperative registered imaging data and does not require any force sensors. By using the CT-based planning system, not only does the surgeon have a better estimation of prosthesis positioning, but he/she can also preoperatively define a safety region in which it is safe to perform bone tissue removal. The basic concept behind the active constraint control instrument is to gradually increase the haptic stiffness to the surgeon's hand via the handle when the cutter approaches the pre-defined forbidden region.

Indeed, the concept of active constraints can be deduced from the original work of Virtual Fixtures first proposed by Rosenberg [84] in 1993. The idea behind virtual fixtures is as a ruler guiding a pen and this would alleviate the workload of certain human sensory modalities on processing the remote-control task. Further work on virtual fixtures was carried out by Burschka et al. [85], who developed a closed-loop position control system to guide the user using virtual fixtures based on both tool tracking and a visual reconstruction of the surrounding environment. The latest research on virtual fixtures seeks to simulate force feedback consisting of dynamic properties. In recent work, Ren et al. [86] applied dynamic virtual fixtures on beating heart ablation procedures. The operator would sense a peg-in-hole haptic effect when reaching the ablation target using the haptic device. Therefore, the method relies heavily on the preoperative images and the accuracy of intra-operative registration.

7.2. Dynamic active constraints and gaze-contingent motor channeling

With the limitations of 3D structural recovery and registration for intra-operative guidance, the future clinical impact of active constraints relies not only on imaging techniques, but also on its ability to incorporate human interaction to enhance its flexibility. To this end, gaze-contingent information was presented as an important visual sensory channel to illustrate how the information from eye movements and ocular vergence can be used to control the instruments and update the active constraints in dynamic surgical scenes. The work on gaze-contingent depth recovery was proposed by Mylonas et al. [87]. By tracking the binocular eye movement and calculating ocular vergence, the 3D depth of the fixation point on the tissue can be determined. Given the known intrinsic and extrinsic parameters of the calibrated stereo laparoscopic camera, the 3D distance between the laparoscopic instrument and the eye fixation point can be computed. Based on the gaze-contingent framework, Mylonas et al. [88] have proposed gaze-contingent active constraints. Real-time binocular eye tracking was adopted to augment the robotic manipulation with human vision in a way that increases performance and accuracy.

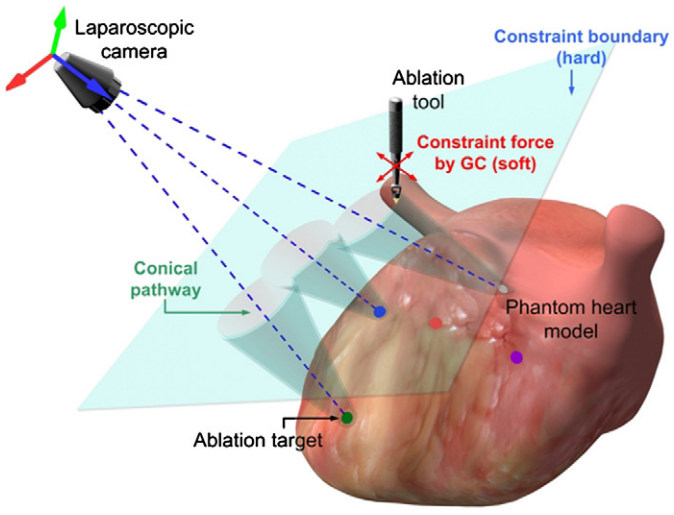

Fig. 8. A schematic illustration of dynamic active constraints by the use of conical pathways carved out from the active constraint boundary, which can be positioned towards the ablation targets based on the fixation during the experiment [88].

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

By generating a force based on the relative separation between the fixation point and the instrument tip, the operator can perceive the haptic cues and constraints for manipulating the instrument. Fig. 8 illustrates the use of dynamic active constraints through the use of conical pathways leading towards the target. This constraint comprised of a planar boundary parallel to the viewport of the virtual camera and a narrow conical channel aimed at the target as determined by the fixation point. It has been shown that by the use of this scheme, significant improvement can be achieved in terms of the confidence and accuracy of the instrument manipulation [88]. This effectively bridges the visual and motor modalities by shifting the burden of the instrument control and the computational cognition towards a perceptually enabled channel.

8. Discussion and conclusions

In this paper, we have outlined some of the major research issues related to image-guided and robotic assisted surgical navigation. Although there exists much research into robotic assisted MIS, there still remain many unresolved challenges. For tissue deformation tracking, the accurate identification of surface features is a critical issue as they can be unstable and prone to errors due to intrinsic tissue surface appearance or specular highlights. Dense 3D reconstruction constitutes one of the essential building blocks for truly general robotic assisted MIS frameworks. While current methods offer adequate performance mainly in the field of motion estimation/compensation, the ability to recover accurate dense depth information is of great importance if an integrated framework encompassing processes from preoperative and intra-operative data registration for computer-aided navigation to prescription of active constraints is to be developed.

Given the shortcomings of single-cue systems and the recent drive towards developing multi-cue integration, the future of 3D reconstruction is increasingly moving towards multi-modal approaches, with cues not limited to being visual, but possibly also inferred from the physical properties of the scene. For example, invisible structured light [89] can be used to recover depth and surface information without interfering with the surgeon's visual scene. As for the fusion mechanism, while graphical models provide great flexibility when compared to ad hoc methods, their computational complexity cannot be under-estimated, especially with MIS tools being developed with increasingly higher output resolutions. However, depth information provides no indication of the structures located below the surface visible to the surgeon. The integration of 3D imaging techniques can assist with navigation. Registration of images acquired intra-operatively and models built preoperatively are crucial for real-time guidance and intervention. It is feasible that biomechanical models obtained from the patient data can be simulated and displayed to the surgeon on demand, however, the difficulties of achieving real-time performance while maintaining the required accuracy should not be under-estimated. Increasingly, the incorporation of biophotonics can provide surgeons with details of the tissue at a cellular level. However, the issue of how to manage this information effectively presents some unique research opportunities.

In conclusion, medical simulation, manipulation and augmented reality systems have made significant inroads into their practical clinical use but significant challenges are still lying ahead. One of the challenges to be tackled is the creation of a complete multi-modal image-guided MIS system for soft tissue interventions, such as tumor excision and beating heart surgery. Ideally, a dynamic biomechanical model will have been built from preoperative data, providing a subject-specific model of the structures in question. The visualization of the model *in situ* during a procedure will greatly enhance orientation and improve accuracy and the incorporation of real-time depth recovery can be used for haptic feedback and the definition of no-go zones, thus significantly improving patient safety.

It is expected that tactile and haptic-enabled virtual fixtures or active constraints will be an important research direction with the aim of improving the fidelity of the interface between the operator and surgical environment in robotic assisted MIS. We envision that the synergistic use of dynamic active constraints associated with sensory information will enhance the surgical performance in terms of speed, safety and reliability. Research such as gaze-contingent active constraints [90] is a new way of enhancing the accuracy of instrument positioning by guiding and confining the motion of the instrument end-effector based on visual sensory information. With the emergence of hyper-redundant flexible manipulators, systematic constraint analysis in this large robot configuration space will be required. Many significant hurdles such as excessive manipulation delay due to tele-operated control will also need to be addressed.

Acknowledgement

The authors wish to thank Andrew Huntbatch for his images on myocardial velocity.

References

[1] Dario P, Hannaford B, Menciassi A. Smart surgical tools and augmenting devices. IEEE Transactions on Robotics and Automation 2003;19(5):782–92.
[2] Mack MJ. Minimally invasive and robotic surgery. JAMA 2001;285(5):568–72.
[3] Nakamura Y, Kishi K, Kawakami H. Heartbeat synchronization for robotic cardiac surgery. In: Proceedings of the 2001 IEEE international conference on robotics and automation (ICRA), vol. 2. 2001. p. 2014–9.
[4] Ginhoux R, Gangloff J, de Mathelin M, Soler L, Sanchez MMA, Marescaux J. Active filtering of physiological motion in robotized surgery using predictive control. IEEE Transactions on Robotics 2005;21(1):67–79.
[5] Sauvée M, Noce A, Poignet P, Triboulet J, Dombre E. Three-dimensional heart motion estimation using endoscopic monocular vision system: from artificial landmarks to texture analysis. Biomedical Signal Processing and Control 2007;2(3):199–207.
[6] Dey D, Gobbi DG, Slomka PJ, Surry KJM, Peters TM. Automatic fusion of freehand endoscopic brain images to three-dimensional surfaces: creating stereoscopic panoramas. IEEE Transactions on Medical Imaging 2002;21(1):23–30.
[7] Thoranaghatte R, Zheng G, Langlotz F, Nolte L. Endoscope-based hybrid navigation system for minimally invasive ventral spine surgeries. Computer Aided Surgery 2005;10(5–6):351–6.
[8] Siddique S, Jaffray D. Towards active image guidance: tracking of a fiducial in the thorax during respiration under X-ray fluoroscopy. In: Proceedings of the SPIE international symposium on medical imaging 2007: visualization and image-guided procedures, vol. 6509. 2007. p. 29–41.
[9] Giannarou S, Visentini-Scarzanella M, Yang G-Z. Affine-invariant anisotropic detector for soft tissue tracking in minimally invasive surgery. IEEE international symposium on biomedical imaging (ISBI) 2009; in press. Q2
[10] Mountney P, Lo B, Thiemjarus S, Stoyanov D, Yang G-Z. A probabilistic framework for tracking deformable soft tissue in minimally invasive surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007. LNCS 4792;2007:34–41.
[11] Stoyanov D, Mylonas GP, Deligianni F, Darzi A, Yang G-Z. Soft-tissue motion tracking and structure estimation for robotic assisted MIS procedures. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005. LNCS 2005;3750:139–46.
[12] Ortmaier T, Groger M, Boehm DH, Falk V, Hirzinger G. Motion estimation in beating heart surgery. IEEE Transactions on Biomedical Engineering 2005;52(10):1729–40.
[13] Groeger M, Ortmaier T, Sepp W, Hirzinger G. Tracking local motion on the beating heart. In: Proceedings of SPIE medical imaging 2002: visualization, image-guided procedures, and display, vol. 4681. 2002. p. 233–41.
[14] Burschka D, Li M, Ishii M, Taylor RH, Hager GD. Scale-invariant registration of monocular endoscopic images to CT-scans for sinus surgery. Medical Image Analysis 2005;9(5):413–26.
[15] Hager GD, Belhumeur PN. Efficient region tracking with parametric models of geometry and illumination. IEEE Transactions on Pattern Analysis and Machine Intelligence 1998;20(10):1025–39.
[16] Bourger F, Doignon C, Zanne P, de Mathelin M. A model-free vision-based robot control for minimally invasive surgery using ESM tracking and pixels color selection. IEEE International Conference on Robotics and Automation 2007:3579–84.
[17] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision 2002;47(1):7–42.

[18] Komodakis N, Tziritas G, Paragios N. Performance vs computational efficiency for optimizing single and dynamic MRFs: setting the state of the art with primal-dual strategies. Computer Vision and Image Understanding 2008;112(1):14–29.
[19] Tappen MF, Freeman WT. Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters. In: Proceedings of the ninth IEEE international conference on computer vision, vol. 2. 2003. p. 900–6.
[20] Lau WW, Ramey NA, Corso JJ, Thakor NV, Hager GD. Stereo-based endoscopic tracking of cardiac surface deformation. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004. LNCS 2004;3217:494–501.
[21] Richa R, Poignet P, Liu C. Efficient 3D tracking for motion compensation in beating heart surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:684–91.
[22] Stoyanov D, Darzi A, Yang G-Z. A practical approach towards accurate dense 3D depth recovery for robotic laparoscopic surgery. Computer Aided Surgery 2005;10(4):199–208.
[23] Horn BKP. Shape from shading: a method for obtaining the shape of a smooth opaque object from one view. Ph.D. MIT; 1970.
[24] Rashid H, Burger P. Differential algorithm for the determination of shape from shading using a point light source. Image and Vision Computing 1992;10(2):119–27.
[25] Deguchi K, Okatani T. Shape reconstruction from an endoscope image by shape-from-shading technique for a point light source at the projection center. In: Proceedings of the workshop on mathematical methods in biomedical image analysis. 1996. p. 290–8.
[26] Rouy E, Tourin A. A viscosity solutions approach to shape-from-shading. SIAM Journal on Numerical Analysis 1992;29(3):867–84.
[27] Tankus A, Sochen N, Yeshurun Y. Reconstruction of medical images by perspective shape-from-shading. In: 17th international conference on pattern recognition (ICPR'04), vol. 3. 2004. p. 778–81.
[28] Kimmel R, Bruckstein AM. Global shape from shading. Computer Vision and Image Understanding 1995;62(3):360–9.
[29] Yuen SY, Tsui YY, Chow CK. A fast marching formulation of perspective shape from shading under frontal illumination. Pattern Recognition Letters 2007;28(7):806–24.
[30] Prados E, Camilli F, Faugeras O. A unifying and rigorous shape from shading method adapted to realistic data and applications. Journal of Mathematical Imaging and Vision 2006;25(3):307–28.
[31] Yuille A, Kersten D. Vision as Bayesian inference: analysis by synthesis? Trends in Cognitive Sciences 2006;10(7):301–8.
[32] Du W, Piater J. A probabilistic approach to integrating multiple cues in visual tracking. In: Computer Vision – ECCV 2008. LNCS 2008;5303:225–38.
[33] Saxena A, Chung SH, Ng AY. 3-D depth reconstruction from a single still image. International Journal of Computer Vision 2008;76(1):53–69.
[34] White R, Forsyth DA. Combining cues: shape from shading and texture. In: 2006 IEEE computer society conference on computer vision and pattern recognition, vol. 2. 2006. p. 1809–16.
[35] Lo BPL, Visentini-Scarzanella M, Stoyanov D, Yang G-Z. Belief propagation for depth cue fusion in minimally invasive surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:104–12.
[36] Stoyanov D, Mylonas G, Yang G-Z. Gaze-contingent 3D control for focused energy ablation in robotic assisted surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:347–55.
[37] Masood S, Yang G-Z, Pennell DJ, Firmin DN. Investigating intrinsic myocardial mechanics: the role of MR tagging, velocity phase mapping, and diffusion imaging. Journal of Magnetic Resonance Imaging 2000;12(6):873–83.
[38] Axel L, Dougherty L. MR imaging of motion with spatial modulation of magnetization. Radiology 1989;171(3):841–5.
[39] Osman NF, Sampath S, Atalar E, Prince JL. Imaging longitudinal cardiac strain on short-axis images using strain-encoded MRI. Magnetic Resonance in Medicine 2001;46(2):324–34.
[40] Aletras AH, Ding S, Balaban RS, Wen H. DENSE: displacement encoding with stimulated echoes in cardiac functional MRI. Journal of Magnetic Resonance 1999;137:247–52.
[41] Bryant DJ, Payne JA, Firmin DN, Longmore DB. Measurement of flow with NMR imaging using a gradient pulse and phase difference technique. Journal of Computer Assisted Tomography 1984;8(4):588–93.
[42] Nayler GL, Firmin DN, Longmore DB. Blood flow imaging by cine magnetic resonance. Journal of Computer Assisted Tomography 1986;10(5):715–22.
[43] Huntbatch A, Lee S-L, Firmin D, Yang G-Z. Bayesian motion recovery framework for myocardial phase-contrast velocity MRI. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:79–86.
[44] Peyrat JM, Sermesant M, Pennec X, Delingette H, Chenyang X, McVeigh ER, et al. A computational framework for the statistical analysis of cardiac diffusion tensors: application to a small database of canine hearts. IEEE Transactions on Medical Imaging 2007;26(11):1500–14.
[45] Tseng W-YI, Dou J, Reese TG, Wedeen VJ. Imaging myocardial fiber disarray and intramural strain hypokinesis in hypertrophic cardiomyopathy with MRI. Journal of Magnetic Resonance Imaging 2006;23(1):1–8.
[46] Cootes T, Taylor C, Cooper D, Graham J. Active shape models—their training and application. Computer Vision and Imaging Understanding 1995;61(1):38–59.
[47] Horkaew P, Yang G-Z. Construction of 3D dynamic statistical deformable models for complex topological shapes. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004. LNCS 2004;3216:217–24.
[48] Lee S-L, Horkaew P, Caspersz W, Darzi A, Yang G-Z. Assessment of shape variation of the levator ani with optimal scan planning and statistical shape modeling. Journal of Computer Assisted Tomography 2005;29(2):154–62.
[49] Lee S-L, Tan E, Khullar V, Gedroyc W, Darzi A, Yang G-Z. Physical-based statistical shape modeling of the levator ani. IEEE Transactions on Medical Imaging; in press. Q3
[50] Lorenzo-Valdés M, Sanchez-Ortiz GI, Elkington AG, Mohiaddin RH, Rueckert D. Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm. Medical Image Analysis 2004;8(3):255–65.
[51] Lekadir K, Yang G-Z. Optimal feature point selection and automatic initialization in active shape model search. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5241:434–41.
[52] Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, Walker MJ, et al. Automatic model-based segmentation of the heart in CT images. IEEE Transactions on Medical Imaging 2008;27(9):1189–201.
[53] Masood S, Gao J, Yang G-Z. Virtual tagging: numerical considerations and phantom validation. IEEE Transactions on Medical Imaging 2002;21(9):1123–31.
[54] Lee S-L, Wu Q, Huntbatch A, Yang G-Z. Predictive K-PLSR myocardial contractility modeling with phase contrast MR velocity mapping. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2007;4792:866–73.
[55] Lee S-L, Huntbatch A, Yang G-Z. Contractile analysis with Kriging based on MR myocardial velocity imaging. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5241:892–9.
[56] Taylor RH, Stoianovici D. Medical robotics in computer-integrated surgery. IEEE Transactions on Robotics and Automation 2003;19(5):765–81.
[57] Peters TM. Image-guidance for surgical procedures. Physics in Medicine and Biology 2006;51:R505–40.
[58] Bajura M, Fuchs H, Ohbuchi R. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. In: 19th annual conference on computer graphics and interactive techniques. 1992. p. 203–10.
[59] Abolmaesumi P, Salcudean SE, Zhu WH. Visual servoing for robot-assisted diagnostic ultrasound. In: Proceedings of the 22nd annual international conference of the IEEE, vol. 4. 2000. p. 2532–5.
[60] Bainbridge D, Jones DL, Guiraudon GM, Terrence M Peters. Ultrasound image and augmented reality guidance for off-pump, closed, beating, intracardiac surgery. Artificial Organs 2008;32(11):840–5.
[61] Rickers C, Jerosch-Herold M, Hu X, Murthy N, Wang X, Kong H, et al. Magnetic resonance image-guided transcatheter closure of atrial septal defects. Circulation 2003;107(1):132–8.
[62] Johnson LG, Edwards P, Hawkes D. Surface transparency makes stereo overlays unpredictable: the implications for augmented reality. Studies in Health Technology and Informatics 2003;94:131–6.
[63] Sielhorst T, Bichlmeier C, Heining SM, Navab N. Depth perception—a major issue in medical AR: evaluation study by twenty surgeons. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006. LNCS 2006;4190:364–72.
[64] Bichlmeier C, Heining M, Feuerstein M, Navab N. The virtual mirror: a new interaction paradigm for augmented reality environments. IEEE Transactions on Medical Imaging; in press. Q4
[65] Lerotic M, Chung AJ, Mylonas G, Yang G-Z. pq-space based non-photorealistic rendering for augmented reality. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007. LNCS 2007;4792:102–9.
[66] Horn BKP. Height and gradient from shading. International Journal of Computer Vision 1990;5(1):37–75.
[67] Chiu AM, Dey D, Drangova M, Boyd WD, Peters TM. 3-D image guidance for minimally invasive robotic coronary artery bypass. The Heart Surgery Forum 2000;3(3):224–31.
[68] Coste-Manière È, Adhami L, Mourgues F, Carpentier A. Planning, simulation, and augmented reality for robotic cardiac procedures: the STARS system of the ChIR team. Seminars in Thoracic and Cardiovascular Surgery 2003;15(2):141–56.
[69] Dombre E, Michelin M, Pierrot F, Poignet P, Bidaud P, Morel G, et al. MARGE project: design, modeling and control of assistive devices for minimally invasive surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004. LNCS 2004;3217:1–8.
[70] Michelin M, Poignet P, Dombre E. Dynamic task/posture decoupling for minimally invasive surgery motions: simulation results. In: Proceedings of the 2004 IEEE/RSJ international conference on intelligent robots and systems (IROS 2004), vol. 4. 2004. p. 3625–30.
[71] Taylor R, Jensen P, Whitcomb L, Barnes A, Kumar R, Stoianovici D, et al. A steady-hand robotic system for microsurgical augmentation. The International Journal of Robotics Research 1999;18(12):1201–10.
[72] Ang WT, Riviere CN, Khosla PK. An active hand-held instrument for enhanced microsurgical accuracy. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000. LNCS 2000;1935:878–86.
[73] Patronik NA, Zenati MA, Riviere CN. Crawling on the heart: a mobile robotic device for minimally invasive cardiac interventions. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004. LNCS 2004;3217:9–16.
[74] D'Attanasio S, Tonet O, Megali G, Carrozza M, Dario P. A semi-automatic handheld mechatronic endoscope with collision-avoidance capabilities. IEEE International Conference on Robotics and Automation 2000;2:1586–91.
[75] Swain P. A justification for NOTES-natural orifice translumenal endosurgery. Gastrointestinal Endoscopy 2007;65(3):514–6.
[76] Lerotic M, Chung AJ, Clark J, Valibeik S, Yang G-Z. Dynamic view expansion for enhanced navigation in natural orifice transluminal endoscopic surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:467–75.

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007

[77] Craige JC. Tactile pattern perception and its perturbations. The Journal of the Acoustical Society of America 1985;77(1):238–46.
[78] Okamura AM. Methods for haptic feedback in teleoperated robot-assisted surgery. Industrial Robot: An International Journal 2004;31(5):499–508.
[79] Pezzementi Z, Ursu D, Misra S, Okamura AM. Modeling realistic tool–tissue interactions with haptic feedback: a learning-based method. In: Symposium on haptic interfaces for virtual environment and teleoperator systems (Haptics 2008). 2008. p. 209–15.
[80] Mahvash M, Okamura AM. Friction compensation for enhancing transparency of a teleoperator with compliant transmission. IEEE Transactions on Robotics 2007;23(6):1240–6.
[81] Basdogan C, De S, Kim J, Muniyandi M, Kim H, Srinivasan MA. Haptics in minimally invasive surgical simulation and training. IEEE Computer Graphics and Applications 2004;24(2):56–64.
[82] Davies B, Jakopec M, Harris SJ, Baena FRY, Barrett A, Evangelidis A, et al. Active-constraint robotics for surgery. Proceedings of the IEEE 2006;94(9):1696–704.
[83] Jakopec M, Baena FRY, Harris SJ, Gomes P, Cobb J, Davies BL. The hands-on orthopaedic robot "Acrobot": early clinical trials of total knee replacement surgery. IEEE Transactions on Robotics and Automation 2003;19(5):902–11.
[84] Rosenberg LB. Virtual fixtures: perceptual tools for telerobotic manipulation. In: Proceedings of the IEEE annual international symposium on virtual reality. 1993. p. 76–82.
[85] Burschka D, Corso JJ, Dewan M, Lau W, Li M, Lin H, et al. Navigating inner space: 3-D assistance for minimally invasive surgery. Robotics and Autonomous Systems 2005;52(1):5–26.
[86] Ren J, Patel RV, McIsaac KA, Guiraudon G, Peters TM. Dynamic 3-D virtual fixtures for minimally invasive beating heart procedures. IEEE Transactions on Medical Imaging 2008;27(8):1061–70.
[87] Mylonas GP, Darzi A, Yang GZ. Gaze-contingent control for minimally invasive robotic surgery. Computer Aided Surgery 2006;11(5):256–66.
[88] Mylonas G, Kwok KW, Darzi A, Yang G-Z. Gaze-contingent motor channelling and haptic constraints for minimally invasive robotic surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008. LNCS 2008;5242:676–83.
[89] Fofi D, Sliwa T, Voisin Y. A comparative survey on invisible structured light. Proceedings of SPIE 2004;5303:90–8.
[90] Mylonas GP, Darzi A, Yang G-Z. Gaze contingent depth recovery and motion stabilisation for minimally invasive robotic surgery. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004. LNCS 2004;3150:311–9.

Please cite this article in press as: Lee S-L, et al. From medical images to minimally invasive intervention: Computer assistance for robotic surgery. Comput Med Imaging Graph (2009), doi:10.1016/j.compmedimag.2009.07.007