Computer-Assisted Laparoscopic Myomectomy by Augmenting the Uterus with Pre-operative MRI Data

Toby Collins   Daniel Pizarro   Adrien Bartoli   Michel Canis   Nicolas Bourdel

ALCoV-ISIT, UMR 6284 CNRS/Université d'Auvergne, Clermont-Ferrand, France
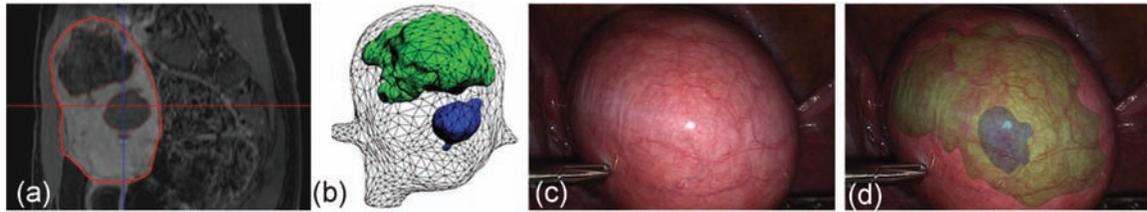
Figure 1: Proposed AR solution for computer-assisted myomectomy. (a) Sagittal MRI slice showing a uterus with two myomas in dark grey. (b) MRI segmentation of uterus and myomas. (c) Raw laparoscopic image and (d) translucent augmentation of the myomas.

ABSTRACT

An active research objective in Computer Assisted Intervention (CAI) is to develop guidance systems to aid surgical teams in laparoscopic Minimal Invasive Surgery (MIS) using Augmented Reality (AR). This involves registering and fusing additional data from other modalities and overlaying it onto the laparoscopic video in realtime. We present the first AR-based image guidance system for assisted myoma localisation in uterine laparosurgery. This involves a framework for semi-automatically registering a pre-operative Magnetic Resonance Image (MRI) to the laparoscopic video with a deformable model. Although there has been several previous works involving other organs, this is the first to tackle the uterus. Furthermore, whereas previous works perform registration between one or two laparoscopic images (which come from a stereo laparoscope) we show how to solve the problem using many images (e.g. 20 or more), and show that this can dramatically improve registration. Also unlike previous works, we show how to integrate occluding contours as registration cues. These cues provide powerful registration constraints and should be used wherever possible. We present retrospective qualitative results on a patient with two myomas and quantitative semi-synthetic results. Our multi-image framework is quite general and could be adapted to improve registration in other organs with other modalities such as CT.

1 INTRODUCTION AND BACKGROUND

1.1 Localising Myomas in Laparosurgery

Myomas, also known as uterine fibroids, are the most common tumors of the female genital tract. Myomas are nearly always benign but can lead to problems including prolonged menstrual periods (menorrhagia), a low blood count (anaemia), menstrual cramping (dysmenorrhoea), infertility or miscarriage. A common type of myoma are intramural myomas which grow within the uterine wall and are most successfully treated by laparoscopic myomectomy. When they are small to medium sized (up to 5cm in diameter) they often do not noticeably change the uterus' outer shape, and can be hard to localise. To assist localisation the surgeon usually consults pre-operative images (usually MRI or ultrasound (US)) during surgery. However it can be difficult even for experienced surgeons to predict a myoma's position accurately in this way. Augmented

Reality (AR) offers a possible solution to help. Foremost, this involves solving nonrigid pre-operative MRI/laparoscopic image registration of the uterus. Once achieved the position of myomas can be transferred via this registration and augmented onto the laparoscopic video. In this paper we propose the first method to solve this problem.

1.2 Prior work

There is much ongoing research to apply AR to help overcome some of the current limitations of MIS [4, 13]. The general objective is for realtime fusion between another modality that conveys additional information (such as the location of important subsurface structures) and the MIS video. Previous works can be broadly divided into two categories depending on whether the second modality is captured pre-operatively, or intra-operatively and synchronously with the MIS video. The second category involves determining a rigid transform between the sensors' 3D coordinate frames. Solutions to this in laparosurgery have been proposed by either externally tracking the laparoscope with optical or magnetic markers [18] or placing on tissue artificial markers that can be detected in both modalities [16].

If additional interventional modalities are not available or not sufficiently informative then AR can be performed with pre-operative modalities. In this case the registration problem can be much more challenging due to soft tissue deformation between acquisition times. Also once significant changes are made during surgery the pre-operative data becomes 'outdated', and less useful. Previous works can be divided into whether the MIS video is monocular [15] or stereo [2, 3, 5, 8, 9, 18]. In [15] an initial rigid registration is performed manually with an interactive Graphical User Interface (GUI), and then automatic feature-based tracking is performed. The registration problem is less difficult with stereo because one can reconstruct intra-operative 3D surface information by stereo reconstruction [17], and this provides greater registration constraints. Of these stereo methods, some perform the registration with a manual GUI [5, 9] and others perform it semi-automatically with manually located landmarks [2, 3, 8, 18]. In some works the registration is refined by minimising the distance between the pre and intra-operative surfaces [8, 18], and in other works tracking is performed with texture features [9]. Various deformation models have been used to constrain the registration problem, including the rigid model [18], 3D splines [3] and biomechanical models [8, 9].

The limitation of the previous works is that they use only one monocular or two stereo images to constrain the registration. For the stereo methods a partial 3D reconstruction is computed (known

as a *2.5D reconstruction* in computer vision), and all regions of the organ not visible in both images have no 3D information. Thus the registration is essentially guessed at these regions from the deformation model's prior. Secondly they do not use the organ's *occluding contours* as registration constraints. Typically a stereo reconstruction never computes 3D information well at the occluding contours. The occluding contours provide powerful boundary conditions and should be exploited whenever possible.

We present the first framework which uses *several laparoscopic images simultaneously* to perform registration between a pre-operative 3D modality and laparoscopic images. In our current work the images are captured by a monocular laparoscope, but these could have come from a stereo device. We are also the first to integrate occluding contours as registration constraints from these images. Our framework has been tested on the uterus, but it is adaptable to other organs and other pre-operative 3D image modalities.

2 PROBLEM SETUP AND APPROACH OVERVIEW

2.1 Problem Setup

We use T2 weighted contrast-enhanced 1.2mm slice MR as pre-operative images. For the intra-operative images, we use images streamed from a standard High Definition handheld monocular laparoscope. We assume the outer surface of the uterus has been segmented in the MRI, which can be achieved reliably using standard interactive tools such as MITK [19]. Our registration problem is therefore a 3D surface-to-2D video registration problem. We use the function $f(\mathbf{p},t)$ to denote transforming a 3D point $\mathbf{p}$ located in the uterus in the MRI volume to its position in the laparoscopic coordinate frame at time t. Due to physical changes including abdominal insufflation, f is a nonrigid transform. We use $\mathcal{M}$ to denote a 3D mesh model constructed from the segmentation of the uterus' outer surface (Fig. 1-b). $\mathcal{M}$ is constructed using standard methods based on marching cubes followed by mesh smoothing and decimation. Usually the number of vertices in $\mathcal{M}$ will be in the order of several thousands. We restrict $\mathcal{M}$ to model the uterus up to the cervix (at which point the mesh is cut). Thus $\mathcal{M}$ is an open surface with disc topology.

2.2 The Reference-Update Formulation

We solve the registration problem by decomposing f into two transforms that can be optimised independently. We define $f(\mathbf{p},t_0)$ to be the *reference transform*. This transforms $\mathbf{p}$ to the 3D coordinate frame of the laparoscope at a reference time t_0. For all other times $t > t_0$ we write f as the composition of $f(\cdot,t_0)$ with an *update transform*: $g(\cdot,t)$, which gives: $f(\mathbf{p},t) = g(f(\mathbf{p},t_0),t)$ (Fig. 2). This decomposition splits f into a component $g(\cdot,t)$ that depends on the current time and a component $f(\cdot,t_0)$ which does not. With this decomposition the speed of registration at time t is determined entirely by the time taken to solve the update transform g. If g is solvable in realtime then full registration is solvable in realtime. With respect to AR, the time taken to determine $f(\cdot,t_0)$ manifests as a delay period before AR can run, and the time taken to solve $f(\cdot,t_0)$ does not affect the AR frame-rate. We call the registration problems corresponding to determining $f(\cdot,t_0)$ and $g(\cdot,t)$ the *reference problem* and the *update problem* respectively.

2.3 Solving the Reference Problem

The reference problem involves finding $f(\cdot,t_0)$, which is a 3D-to-2D registration problem. Although 3D-to-2D registration is well understood and well constrained for rigid objects, it is far more challenging for nonrigid objects. Our solution is to exploit the fact that the motion of the organ with respect to the laparoscope immediately after t_0 is approximately rigid (up to the time when the surgeon starts resecting). It has been shown recently that the rigid assumption holds well for describing the motion of the uterus during this period of surgery [6]. Thus we can extract *3D* information from the laparoscope using multiple images captured shortly after t_0 using rigid Structure-from-Motion (SfM). This 3D information converts
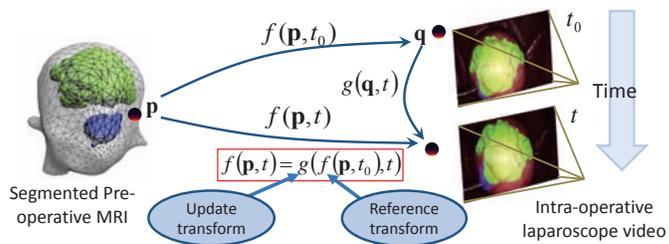
Figure 2: Breaking the registration problem into a nonrigid reference transform and a rigid update transform.

the 3D-to-2D registration problem to a 3D-to-3D registration problem. Specifically, the registration problem now involves registering $\mathcal{M}$ to a set of intrinsically and extrinsically calibrated laparoscopic images. This significantly improves the problem's tractability compared with registering $\mathcal{M}$ to a single laparoscopic image.

Therefore to solve the reference problem we perform two steps. In the first step we gather frames from a short interval $t_0 \leq t \leq t_1$. We call this interval the *exploratory phase*, and in our tests this lasts approximately 20 seconds. We then perform SfM to obtain the following information: *(i)* the laparoscope's 3D rigid transforms with respect to the uterus for times $t_0 \leq t \leq t_1$, *(ii)* 2D feature correspondences in the laparoscopic images for times $t_0 \leq t \leq t_1$, and *(iii)* an estimate of the feature correspondences' 3D positions. In the second step we use this information to apply 3D constraints for registering $\mathcal{M}$ to the laparoscopic images in the interval $t_0 \leq t \leq t_1$. To further constrain the problem, we introduce two other complementary cues. The first are natural landmarks corresponding to the junctions formed between the fallopian tubes and the uterus. The second is the boundary contour of the uterus.

2.4 Solving the Update Problem and Image Fusion

We use our recently published Wide-Baseline Multi-Texturemap Registration (WBMTR) method for solving the update problem [6]. WBMTR registers the organ in realtime on standard workstation hardware, and is robust to occlusions from *e.g.* surgical tools or partial views, and can handle large viewpoint changes. It recovers automatically if the organ is not visible in sections of the video (*e.g.* if the surgeon removes and then reinserts the laparoscope), and unlike SLAM-based tracking methods does not require successful registration in previous frames. This means WBMTR can run indefinitely over long duration videos. Note that [6] does not solve the reference problem. Once registered we transform the myoma(s) to the laparoscopic coordinate frame using the full transform $f(\cdot,t) = g(f(\cdot,t_0),t)$. We then render the myomas using OpenGL and blend the render with the laparoscopic image at time t using 'x-ray' augmentation [11]. This gives the impression the uterus is translucent and the myomas can be seen inside (Fig. 1-d).

3 SOLVING THE REFERENCE PROBLEM

3.1 The Exploratory Protocol

The goal of the exploratory phase is to gather sufficient views of the uterus shortly after t_0 so that we can reconstruct with rigid SfM the uterus body, and to determine the 3D transforms of the laparoscope with respect to the reconstruction. To achieve this we use a simple protocol that our surgeons perform to gather these views in a repeatable manner. We call this the *exploratory protocol*. The exploratory protocol involves the laparoscope being kept in an approximately fixed position with respect to the patient, and the surgeon's assistant orbiting the uterus by moving the uterine cannula (Fig. 3). This gathers posterior, anterior and lateral views of the uterus body, and takes approximately 20 seconds to perform (giving approximately 500 video frames). The protocol is performed only once and workflow interruption is quite minimal.
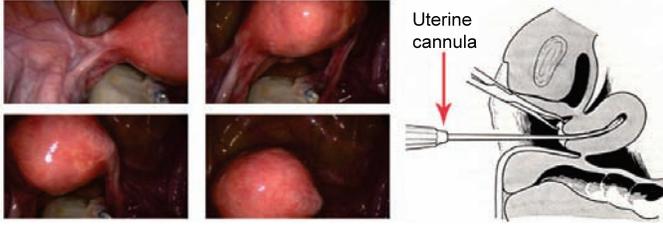
Figure 3: Frames from an exploratory video captured as the surgeon's assistant rotates the uterine cannula.

3.2 Partial Dense 3D Reconstruction of the Uterus

We first downsize the number of frames in the exploratory video to a smaller set of $K = 10$ frames (uniformly distributed in time), which we call the keyframe set. We manually mark in the keyframe set occluding contours of the uterus and the position (if visible) of the fallopian tube junctions (Fig. 4). This takes between 1-2 minutes to perform. We then perform dense SfM on the keyframe set, ensuring that the SfM engine only uses features and pixel intensities that lie within the boundary of the uterus in each keyframe (Fig. 5). This is to ensure that motion of background structures do not adversely affect reconstruction. We have found good success using reconstructions by Photoscan [1], which for 10 images takes approximately 1 minute. To improve reconstruction accuracy, we do not self-calibrate the laparoscope jointly with 3D reconstruction. Instead, before the laparoscope is inserted we perform intrinsic calibration using OpenCV's calibration library, and keep its intrinsics fixed during 3D reconstruction.

We define the 3D reconstruction in world coordinates, and define world coordinates to be the laparoscope's 3D coordinate frame at time t_0. We use $\mathscr{T}$ to denote all the information that we retain from the 3D reconstruction. This is given by $\mathscr{T} \stackrel{\text{def}}{=} (\mathbf{Q}_j, \mathbf{M}_i, I_i, \mathbf{q}_i^j, \mathscr{S})$. $i \in [1...K]$ denotes the index of the keyframe and $j \in [1...n]$ denotes the index of a reconstructed 3D feature. $\mathbf{Q}_j \in \mathbb{R}^3$ denotes the 3D position of the jth feature in world coordinates. $\mathbf{M}_i \in SE_3$ denotes the rigid 3D transform that maps world coordinates to the laparoscopic coordinate for the ith keyframe, with $\mathbf{M}_{t_0} = \mathbf{I}_4$. $\mathbf{q}_i^j \in \mathbb{R}^2$ denotes the 2D position of the jth feature in the ith keyframe. Usually $\mathbf{Q}_j$ will not be detected in all keyframes, so $\mathbf{q}_i^j$ may not be defined for all i. I_i denotes the ith keyframe. $\mathscr{S}$ denotes a 3D mesh model of the reconstructed uterus with vertices defined in world coordinates.
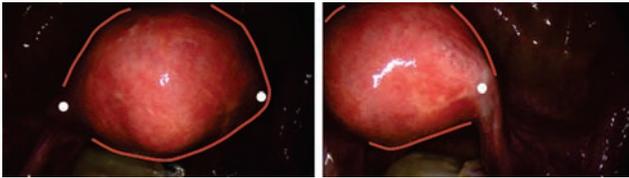
Figure 4: Keyframes marked-up with boundary contours and the 2D positions of the Fallopian tubes/uterus junctions.
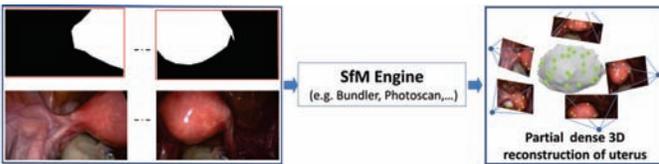
Figure 5: Partial dense 3D reconstruction of the uterus from segmented keyframes using SfM.

3.3 Registration Cost

We use three complementary cues for registering the MRI uterus surface $\mathscr{M}$ to the set of laparoscopic keyframes from the exploratory phase. These are as follows. *Anatomical landmarks*: We have identified two anatomical that can usually be located well in laparoscopic images and on $\mathscr{M}$. These are the junctions formed between the uterus body and the Fallopian tubes (Fig. 4). If $\mathscr{M}$ has been registered well the distance between the landmarks on $\mathscr{M}$ projected into a laparoscopic image should be close to their measured positions. *Contours*: If $\mathscr{M}$ has been registered well the boundary contours of the uterus in a laparoscopic image (Fig. 4) should lie close to the boundary of the silhouette of $\mathscr{M}$. *Feature correspondence*: If $\mathscr{M}$ has been registered well then the correspondences in $\mathscr{T}$ should be consistent with the registration. We now give the formal definitions of these constraints as cost functions. We represent the reference transform $f(\cdot, t_0)$ with an (unknown) parameter vector θ by $f(\cdot, t_0; \theta)$, and our goal is to recover θ.

3.4 Anatomical Landmark Constraints

We use $\mathbf{v}_i^L \in \mathbb{R}^2$ and $\mathbf{v}_i^R \in \mathbb{R}^2$ to denote the 2D position of the left and right Fallopian tube/uterus junctions in the ith keyframe. We do not assume that a junction will be visible in all keyframes. We use $\mathbf{U}^L \in \mathbb{R}^3$ and $\mathbf{U}^R \in \mathbb{R}^3$ to denote the 3D position of the left and right Fallopian tube/uterus junctions on $\mathscr{M}$. The cost is as follows:

$$C_J(\theta) \stackrel{\text{def}}{=} \frac{1}{2K} \sum_{i=1}^{K} \left\| \pi \left(\mathbf{M}_i f(\mathbf{U}^L, t_0; \theta), \mathbf{K} \right) - \mathbf{v}_i^L \right\|_2^2 + \left\| \pi \left(\mathbf{M}_i f(\mathbf{U}^R, t_0; \theta), \mathbf{K} \right) - \mathbf{v}_i^R \right\|_2^2 \quad (1)$$

The perspective projection of a 3D point in the laparoscopic coordinate frame onto its image is given by $\pi(\mathbf{x}, \mathbf{K})$, where $\mathbf{K}$ is the intrinsic calibration matrix. Because the laparoscope is calibrated, we can undo the effects of lens distortion, and π is then given by $\pi(\mathbf{x}, \mathbf{K}) = \frac{1}{\mathbf{K}_{3}\mathbf{x}}[\mathbf{K}_1\mathbf{x}, \mathbf{K}_2\mathbf{x}]^\top$, where $\mathbf{K}_k$ denote the kth row of $\mathbf{K}$. Notice that information in the 3D reconstruction $\mathscr{T}$ is used in this cost. Take the left junction $\mathbf{U}^L$ as an example. First $\mathbf{U}^L$ is transformed according to a nonrigid deformation $f(\mathbf{U}^L, t_0; \theta)$ to transform it into world coordinates. Then $\mathbf{M}_i$ is applied to transform it into the laparoscopic coordinate frame. Thus for any θ we have the predicted 2D positions of $\mathbf{U}^L$ in *each* keyframe.

3.5 Contour Constraints

We assume that the user has been able to partially mark the boundary contours of the uterus in the keyframes. We do not assume the whole boundary has been provided by the user. This cannot be achieved if for example the uterus is occluded by other organs such as the bladder or intestines. We rasterise the user's contours in each keyframe to obtain a collection of M_i pixels on the contour. We store this as $\mathbf{C}^i \in \mathbb{R}^{M_i \times 2}$, where the mth row of $\mathbf{C}^i$ (which we denote by $\mathbf{C}_m^i$) holds the 2D pixel position of the mth point on a contour in the ith keyframe. We use the function $s(\mathscr{M}, t_0; \theta, \mathbf{M}_i, \mathbf{K})$ to denote the silhouette generator function. This function first deforms $\mathscr{M}$ according to θ and then transforms it to the ith coordinate frame via $\mathbf{M}_i$. Next the silhouette of $\mathscr{M}$ is computed on the 2D image plane, and 2D contours are generated on the boundary of the silhouette. The contour cost is then defined as:

$$C_S(\theta) \stackrel{\text{def}}{=} \frac{1}{K} \sum_{i=1}^{K} \frac{1}{M_i} \sum_{m=1}^{M_i} d\left(s(\mathscr{M}, t_0; \theta, \mathbf{M}_i, \mathbf{K}), \mathbf{C}_m^i \right)^2 \quad (2)$$

where $d(\mathscr{C}, \mathbf{c})$ denotes the minimum Euclidean distance from a 2D point $\mathbf{c} \in \mathbb{R}^2$ to the 2D contours $\mathscr{C}$. In practice $s(\mathscr{M}, t_0; \theta, \mathbf{M}_i, \mathbf{K})$ can be approximated by rendering the deformed mesh using z-buffering. This makes computing the silhouette contours fast, however it causes difficulty for gradient-based optimisation because C_S is hard to differentiate due to quantisation induced by z-buffering. Our solution is to approximate C_S using *virtual contour correspondences*, in a similar fashion to [10], but here we work directly with the 3D mesh. The idea is that at a given iteration during optimisation we can construct for each $\mathbf{C}_m^i$ a virtual correspondence on

$\mathcal{M}$. This virtual correspondence is the point on $\mathcal{M}$ that projects closest to $\mathbf{C}_m^i$ in the ith keyframe. By keeping this virtual correspondence fixed for a single optimisation iteration, we can obtain a differentiable cost function. We achieve this as follows. First $\mathcal{M}$ is deformed according to θ and then transformed to the ith coordinate frame via $\mathbf{M}_i$. Next $\mathcal{M}$ is rendered with z-buffering, and from this we compute the silhouette of $\mathcal{M}$. For each point $\mathbf{C}_m^i$ we then compute the closest pixel on the silhouette's contour. If this pixel is at a point $\mathbf{d}_m^i \in \mathbb{R}^2$, we then cast an optical ray through $\mathbf{d}_m^i$ and intersect it with $\mathcal{M}$. This intersection point gives us a virtual correspondence on $\mathcal{M}$. We define $\mathbf{D}_m^i \in \mathbb{R}^3$ to be the virtual correspondence in the MRI's 3D coordinate frame. Our approximation to C_S is then given by:

$$C_S'(\theta) \stackrel{\text{def}}{=} \frac{1}{K} \sum_{i=1}^{K} \frac{1}{M_i} \sum_{m=1}^{M_i} \left\| \pi \left(\mathbf{M}_i f(\mathbf{D}_m^i, t_0; \theta), \mathbf{K} \right) - \mathbf{C}_m^i \right\|_2^2 \quad (3)$$

C_S' is useful for gradient-based optimisation because if f is differentiable up to the cth order, then so is C_S'. Note that at each iteration $\mathbf{D}_m^i$ may change, and so $\mathbf{D}_m^i$ is like an ICP correspondence.

3.6 Correspondence Constraints

So far we have used only the rigid transforms $\mathbf{M}_i$ from the 3D reconstruction information in $\mathcal{T}$. Our third cost function enforces that the correspondence information in $\mathcal{T}$ should be consistent with $f(\cdot, t_0; \theta)$. Specifically, we enforce that $\mathcal{M}$ should be deformed by θ such that the 2D correspondences $\mathbf{q}_i^k$ are predicted well. We do this by introducing N point correspondences $\mathcal{Y} = \{\mathbf{y}_1, ..., \mathbf{y}_N\}$ that we define on the surface of $\mathcal{M}$. A cost is then given to θ if these points do not project close to the correspondence positions $\mathbf{q}_i^j$ in the images. Because $\mathcal{M}$ is a 2D manifold embedded in 3D space, it is easier to define $\mathcal{Y}$ in 2D using a 2D parameterisation of $\mathcal{M}$. Because $\mathcal{M}$ has disc topology a natural 2D parameterisation can be made by conformally flattening $\mathcal{M}$. We denote the domain of this flattening by $\Omega \in \mathbb{R}^2$, and the flattening gives us the function $h(\mathbf{y}_j) : \mathbb{R}^2 \to \mathbb{R}^3$ which maps the 2D point $\mathbf{y}_j \in \Omega$ in Ω to its 3D position in the MRI. The feature correspondence cost then writes as:

$$C_F(\theta, \mathcal{Y}) \stackrel{\text{def}}{=} \frac{1}{KN} \sum_{i=1}^{K} \sum_{j=1}^{N} \rho \left(\pi \left(\mathbf{M}_i f(h(\mathbf{y}_j), t_0; \theta), \mathbf{K} \right) - \mathbf{q}_i^j \right) \quad (4)$$

$$\rho(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{i=1}^{2} 2 \left(\sqrt{1 + \mathbf{x}_i^2 / 2} - 1 \right)$$

Eq. (4) enforces that the positions of $\mathbf{y}_j$ projected into the ith keyframe should be close to $\mathbf{q}_i^j$. Because $\mathcal{Y}$ is unknown, it must be optimised together with θ. One can think of Eq. (4) as a type of constrained multiview triangulation where the feature positions are optimised on the surface of the uterus jointly with estimating the nonrigid deformation of the uterus. The function ρ is an M-estimator which we use to handle the fact that the output from a SfM engine is never perfect and some correspondences will be incorrect. The M-estimator provides robustness to such errors. In our current system we use the $\mathcal{L}_1 - \mathcal{L}_2$ M-estimator.

3.7 Deformable Model, Initialisation and Optimisation

The total registration cost is given by:

$$C(\theta, \mathcal{Y}) = C_J(\theta) + \gamma C_S'(\theta) + \lambda C_F(\theta, \mathcal{Y}), \quad (5)$$

Where γ and λ are weights associated with C_S' and C_F respectively. In the framework outlined so far we have kept things independent of the deformation model. We now explain the model we currently use for the uterus. The biomechanical properties of the uterus have been studied in *e.g.* [14], which showed considerable intra and inter-subject variability (its Young's modulus ranged from 20.3 kPa to 1379 kPa). Thus using a precise biomechanical model is challenging. We propose using a simpler deformable model, since the deformation between the uterus in the MRI and during laparosurgery is

simple and mostly caused by compression from the abdominal wall. To model this we use an affine transform, which has been used successfully for registering other soft organs [12]. Thus $f(\cdot, t_0)$ transforms a 3D point $\mathbf{p}$ in the MRI according to $f(\mathbf{p}, t_0) = \mathbf{A}\mathbf{p} + \mathbf{t}$, with $(\mathbf{A} \in \mathbb{R}^{3\times3}, \mathbf{t} \in \mathbb{R}^{3\times1})$ being a 3D affine transform. Therefore θ holds the 12 coefficients of $(\mathbf{A}, \mathbf{t})$.

If ground-truth registrations were available γ and λ can be learned from training examples. At present we do not have ground-truth registrations (since these are very difficult to obtain), so instead we manually set γ and λ. A nice property of our cost functions is that all costs are defined in pixels, which makes setting γ and λ easier. In our experiments we use $\gamma = 1\mathrm{e}^{-1}$ and $\lambda = 1\mathrm{e}^{-1}$.

We initialise θ and $\mathcal{Y}$ as follows. We know in advance the approximate pose of the uterus with respect to the laparoscope at time t_0. This is because the laparoscope is inserted through the abdomen at a standard location above the navel, its tip is approximately 10cm from the uterus, and the laparoscope is rotated to view the uterus in a canonical 'upright' position (Fig. 4-left). This information provides *a priori* a rough estimate of the rigid transform $\mathbf{M}_0 \in SE_3$ that maps the MRI to world coordinates. We thus initialise $\mathbf{A}$ with $\mathbf{M}_0$. We initialise $\mathbf{y}_j \in \mathcal{Y}$ as follows. We first transform $\mathcal{M}$ to world coordinates using $\mathbf{M}_0$. Then we take the 3D estimate $\mathbf{Q}_j$ of the correspondences in world coordinates (from $\mathcal{T}$) and find the closest 3D points on $\mathcal{M}$. These closest points are then mapped to the 2D parameterisation of $\mathcal{M}$ via h (§3.6) to give us an initial estimate of $\mathbf{y}_j$. Having initialised θ and $\mathcal{Y}$, these are then refined by minimising Eq. (5) using gradient-based optimisation. Specifically we use Gauss-Netwon iterations with backtracking line-search. In practice we have found convergence to be fast, and typically taking fewer than 10 iterations.

4 EVALUATION

4.1 Evaluation with a Synthetic Phantom

We present quantitative results with a synthetic phantom that we constructed by 3D printing a uterus CAD model in a rigid material. The phantom simulates the intra-operative state of the uterus, where we assume rigid motion. The phantom body (shown in Fig. 6(a)) was $90 \times 60 \times 60$ mm^3 and it was printed in color with a realistic texture obtained from images of a human uterus.
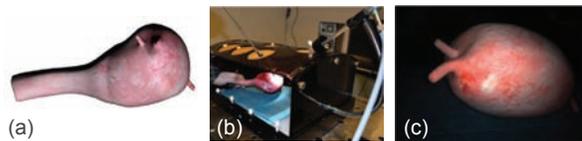
Figure 6: (a) CAD model of the rigid phantom. (b) 3D printed phantom inside a pelvic trainer with a 10mm Karl Storz HD laparoscope. (c) Laparoscopic images of the phantom.

We introduce the rigid phantom into a pelvic trainer (see Fig. 6(b)) to simulate the laparoscopic conditions. We use a 10mm Karl Storz HD laparoscope that is fixed to the pelvic trainer trough a flexible arm. We perform the exploratory video, keeping $K = 20$ keyframes (see Fig. 7) from where we have manually marked the boundary contours and the fallopian tube junctions. Fig. 7 shows the reconstructed sparse point cloud from PhotoScan.

We artificially deform the intra-operative CAD model to simulate the shape of the uterus in its pre-operative state. To this end we use a volumetric deformation model represented with the Thin Plate Spline (TPS). Our TPS is defined by a set of $2 \times 3 \times 3$ control points that uniformly cover the intra-operative uterus volume. By perturbing the reference grid of control points we create non-linear volumetric deformations (see Fig 7, right). We generate many pre-operative test models by first applying a random global affine transform to the reference control points followed by a random displacement of standard deviation $\sigma_p = 10mm$. We write the
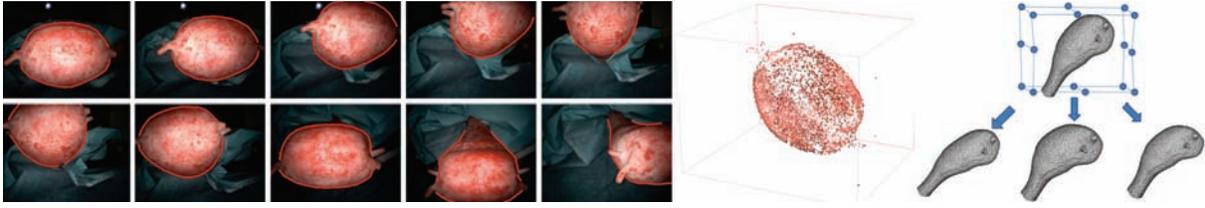
Figure 7: Left: Some of the $K = 20$ exploratory video keyframes with manually marked contours. Middle: Sparse 3D cloud from SfM computed by PhotoScan. Right, up: Pre-operative model overlaid with the TPS grid of control points. Right, down: Several synthetic deformations obtained with the TPS.

TPS as $w_{TPS}(\cdot, \theta_{TPS})$, where θ_{TPS} holds the control centers (*i.e.* 18×3 parameters in total). We highlight that the synthetic deformation we induce with the TPS model is not exactly approximated by our affine model. We use this fact to test our system under non-modelled deformations which can likely occur when dealing with a real uterus. Our quantitative experiment tests the influence of the number of views used during the registration. We randomly select sets of $\{2, 3, 4, 5, 6, 7, 10, 15\}$ images from the $K = 20$ keyframes. For each set of views we test 50 randomly-generated pre-operative models. Fig. 8 shows the initial shapes of three pre-operative models overlaid with the marked contours from three different views.
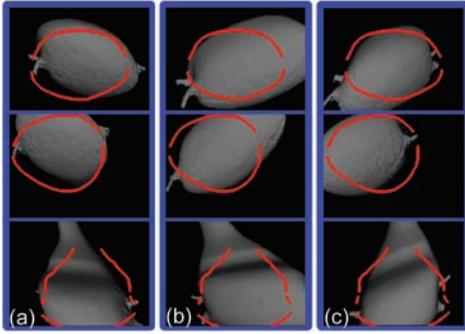
Figure 8: (a), (b) and (c) show the initial projection of three different pre-operative models (one model per column) in three different views (one view per row) using the z-buffer. We overlay on top the contours from each view given by the user.

For each trial we measure registration error in 3D by discretising the intra-operative volume with a grid of $100 \times 100 \times 100$ voxels. We denote as $\mathcal{V}$, the set of voxels in the interior of the uterus. We define the registration error ε_r as:

$$\varepsilon_r^2 = \frac{1}{N_\mathcal{V}} \sum_{\mathbf{q} \in \mathcal{V}} \|f(w_{TPS}(\mathbf{q}; \theta_{TPS}), t_0; \theta) - \mathbf{q}\|^2. \tag{6}$$

Fig. 9(a) shows the mean registration error we achieve as a function of the number of keyframes available, whereas Fig. 9(b) shows the mean number of iterations it takes our method to reach convergence. Fig. 9(c) shows the registration accuracy by texturing the surface with the mean registration error at each surface's point. What we see is a clear reduction in registration error with increased numbers of views, and saturating at approximately 8 views beyond which the registration error is approximately 1.4mm. The distribution of errors at 2 views in Fig. 9(c) shows a tendency for greater error away from the uterus head towards the cervix. The reason for this is that with only a few views the uterus head is quite well constrained by the feature correspondences from the SfM reconstruction. However, further into the uterus the registration becomes less well constrained which leads to greater error. By contrast with more views we have more constraints from contour information and also more from feature correspondences, and this significantly im-

proves the registration. We also see a significant reduction in the number of iterations for convergence with more views.

4.2 Evaluation with a Patient

We present preliminary qualitative results on a patient with two myomas (Fig. 1). One myoma was large, with a diameter of 121mm, and the other medium sized with a diameter of 52mm. $\mathcal{M}$ was constructed by performing an interactive segmentation of the MRI using MITK [19], followed by meshing with marching cubes, two iterations of Laplacian mesh smoothing and then mesh decimation with quadratic edge collapse [7] to give a mesh of 6210 vertices. The laparoscopic video was recorded during the patient's myomectomy and we performed the registration and augmentation offline after surgery on a 50 second video clip before resection began (including the 20 second exploratory phase). We used Photoscan to perform SfM on 10 keyframes from the exploratory phase, and this computed correspondences for 321 image features located on the uterus. The cost function in Eq. (5) was optimised in 6 Gauss-Newton iterations. This took approximately 15 seconds to compute using an unoptimised Matlab implementation with a C++ z-buffer implementation. The registration for the frames after the exploratory video was achieved using the method presented in [6], which processed each frame in approximately 37.0ms (or 27fps). We show four frames from the clip in Fig. 10 (top row), together with an overlay of $\mathcal{M}$ registered to each frame. We show with a red contour the silhouette boundary of $\mathcal{M}$ in each frame. Qualitatively we see that the red contours align well to the real image contours, and the supplied video shows clearly that the registration track the uterus well over the sequence. This indicates that the affine deformable model we use is sufficiently complex to approximate well the nonrigid deformation of the uterus. In Fig. 10 (bottom row) we show the visual augmentation of the myomas in each frame (the larger myoma is in green and the smaller is in blue). The surgeon who conducted the myomectomy reported that localising the small myoma was difficult, and took them approximately 15 minutes. After surgery they inspected our augmentations and confirmed that both myomas appeared to be localised well by our system. This promising result indicates the potential benefit for using our AR system during surgery.

5 CONCLUSION

We have presented an Augmented Reality solution for assisting the localisation of myomas during laparoscopic surgery. This has been achieved by nonrigidly registering the uterus and myomas from the MRI to the laparoscopic video. This is a challenging problem which we tackle by factoring the registration into two components. The first component is the reference registration, which registers the uterus at a single reference time during surgery. The second component is the update registration, which can be achieved quickly and fully automatically by rigid registration. This paper has focused on solving the reference problem. The main contribution is to show that this can be solved by registering the MRI of the uterus to *multiple* laparoscopic images (which we call keyframes) that are intrinsically and extrinsically calibrated using rigid SfM. Therefore we solve the reference problem by registering the MRI to multiple keyframes *simultaneously*. Registration is performed using three
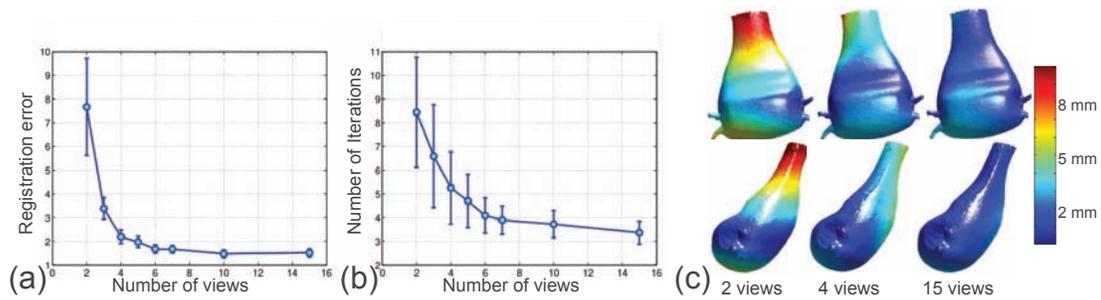
Figure 9: (a) Registration error vs. number of views. (b) Number of iterations vs. number of views. (c) Mean surface registration error for 2,4 and 15 views.
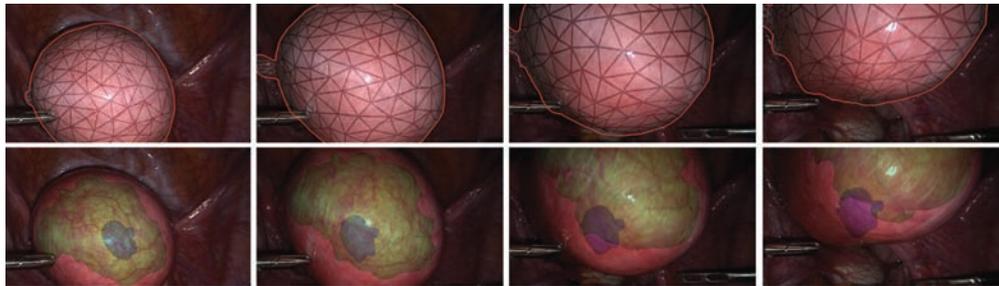
Figure 10: Registration between the uterus in a pre-operative MRI to intra-operative laparoscopic images, and visual augmentation of two hidden myomas. The top row shows $\mathcal{M}$ (the mesh model of the uterus from the MRI) overlaid in each image. The red contours denote the silhouette boundaries of $\mathcal{M}$ in each image. The bottom row shows the augmented myomas. Best Viewed in colour.

complementary cues involving natural anatomical landmarks, the uterus' boundary contour and feature correspondences.

This is the first work which registers a pre-operative MRI of the uterus to laparoscopic images. In our ongoing research we will perform a deeper quantitative evaluation, test other deformable models, and test the system's performance on live videos. Our solution is not limited to visualising myomas. Any anatomical structure that is visible in the MRI can be visualised in the laparoscopic images, such as the cervix and uterine cavity. We expect that augmenting this information may also help the surgeon perform myomectomies and other uterine surgery, improve safety and reduce time in the operating room.

REFERENCES

[1] AgiSoft LLC. PhotoScan, http://www.agisoft.ru/products/photoscan.

[2] H. O. Altamar, R. E. Ong, C. L. Glisson, D. P. Viprakasit, M. I. Miga, S. D. Herrell, and R. L. Galloway. Kidney deformation and intraprocedural registration: A study of elements of image-guided kidney surgery. *Endourology*, 2010.

[3] A. Amir-Khalili, M. S. Nosrati, J.-M. Peyrat, G. Hamarneh, and R. Abugharbieh. Uncertainty-encoded augmented reality for robot-assisted partial nephrectomy: A phantom study. In *AECAI@MICCAI*, 2013.

[4] A. Bartoli, T. Collins, N. Bourdel, and M. Canis. Computer assisted minimally invasive surgery: Is medical computer vision the answer to improving laparosurgery? *Medical Hypotheses*, 2012.

[5] D. Cohen, E. Mayer, D. Chen, A. Anstee, J. Vale, G.-Z. Yang, A. Darzi, and P. J. Edwards. Augmented reality image guidance in minimally invasive prostatectomy. In *Prostate Cancer Imaging*, pages 101–110, 2010.

[6] T. Collins, D. Pizarro, A. Bartoli, M. Canis, and N. Bourdel. Realtime wide-baseline registration of the uterus in laparoscopic videos using multiple texture maps. In *MIAR@MICCAI*. 2013.

[7] M. Garland and P. S. Heckbert. Surface simplification using quadric error metrics. In *SIGGRAPH*, 1997.

[8] G. Hamarneh, A. Amir-Khalili, M. Nosrati, I. Figueroa, J. Kawahara, O. Al-Alao, J.-M. Peyrat, J. Abi-Nahed, A. Al-Ansari, and R. Abugharbieh. Towards multi-modal image-guided tumour identification in robot-assisted partial nephrectomy. In *MECBME*, 2014.

[9] N. Haouchine, J. Dequidt, I. Peterlík, E. Kerrien, M.-O. Berger, and S. Cotin. Image-guided simulation of heterogeneous tissue deformation for augmented reality during hepatic surgery. In *ISMAR*, 2013.

[10] S. Ilić, M. Salzmann, and P. Fua. Implicit meshes for effective silhouette handling. *International Journal of Computer Vision*, 2006.

[11] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *ISMAR*, 2010.

[12] T. Mertzanidou, J. H. Hipwell, M. J. Cardoso, X. Zhang, C. Tanner, S. Ourselin, U. Bick, H. J. Huisman, N. Karssemeijer, and D. J. Hawkes. MRI to X-ray mammography registration using a volume-preserving affine transformation. *Medical Image Analysis*, 2012.

[13] D. J. Mirota, M. Ishii, and G. D. Hager. Vision-based navigation in image-guided interventions. *Annual Review of Biomedical Engineering*, 13, 2011.

[14] G. Pearsall and V. Roberts. Passive mechanical properties of uterine muscle (myometrium) tested in vitro. *Journal of Biomechanics*, 1978.

[15] G. Puerto-Souza, J. A. Cadeddu, and G. Mariottini. Toward long-term and accurate augmented-reality for monocular endoscopic videos. *Biomedical Engineering*, 2014.

[16] T. Simpfendorfer, M. Baumhauer, M. Müller, C. N. Gutt, H.-P. Meinzer, J. J. Rassweiler, S. Guven, and D. Teber. Augmented reality visualization during laparoscopic radical prostatectomy. *Endourology*, 25(12):1841–1845, 2011.

[17] D. Stoyanov, M. V. Scarzanella, P. Pratt, and G.-Z. Yang. Real-time stereo reconstruction in robotically assisted minimally invasive surgery. In *MICCAI*, pages 275–282, 2010.

[18] L.-M. Su, B. P. Vagvolgyi, R. Agarwal, C. E. Reiley, R. H. Taylor, and G. D. Hager. Augmented reality during robot-assisted laparoscopic partial nephrectomy: toward real-time 3D-CT to stereoscopic video registration. *Urology*, 73(4):896–900, 2009.

[19] I. Wolf, M. Vetter, I. Wegner, M. Nolden, T. Böttger, M. Hastenteufel, M. Schöbinger, T. Kunert, and H.-P. Meinzer. The medical imaging interaction toolkit (MITK), http://www.mitk.org/.