

Using Shading to Register an Intraoperative CT Scan to a Laparoscopic Image

Sylvain Bernhardt^{1,2}, Stéphane A. Nicolau², Adrien Bartoli⁴, Vincent Agnus²,
Luc Soler^{1,3}, and Christophe Doignon²

¹ IHU, Institut Hospitalo-Universitaire, Strasbourg, France

² ICube, Université de Strasbourg, France

³ IRCAD, Virtual Surg, Strasbourg, France

⁴ ALCov-ISIT, Université d'Auvergne, Clermont-Ferrand, France

Abstract. *In abdominal surgery, augmented reality has been attempted by registering preoperative 3D data onto the intraoperative laparoscopic view. The registration may be aided by an interventional 3D imaging system such as a rotational C-arm. It has been shown that one can determine the transformation between an intraoperative 3D volume and the laparoscopic view by letting the laparoscope tip enter the C-arm acquisition field. However, the transformation estimation was up to a 1D rotation and a 2D translation. We propose to complete this registration by using local shading constraints with a piecewise constant albedo hypothesis on the surface of the surgical scene. Thus, the registration becomes fully automatic with no extra apparatus required. Results from experiments on in vivo data show a millimetric registration accuracy.*

Keywords: Registration; Abdominal imaging; Minimally invasive procedure; Intraoperative imaging; Endoscopic imaging

1 Introduction

With the advent of minimally invasive surgery and digital endoscopic cameras over the past few decades, intraoperative augmented reality has fostered much research in computer vision [1]. The general goal is to improve the surgeon's perceptions by augmenting the video feedback with a high definition 3D model provided by a preoperative CT or MRI [2–4]. Applications include revealing hidden vessels or tumors. Accurately performing this augmentation remains a challenge as the patient's anatomy may significantly change between the preoperative scanning and the intervention. Notably, in abdominal surgery, the cavity is insufflated with gas which applies pressure on the organs, thus hindering an accurate registration [5]. To compensate for this deformation, a solution may be to introduce a 3D rotational C-arm as an intermediary step in the augmentation process, as this type of apparatus is becoming increasingly popular. Given the non-rigid transformation of the organs of interest between the preoperative and intraoperative 3D scans ([6, 7]), all that is left is to determine the relationship between the intraoperative volume and the laparoscopic camera.

In [8], we presented a method to tackle this problem without external tracking. First, after a classic camera calibration using a checkerboard, the laparoscope is blocked so that it sees the organ of interest and also enters the acquisition field of the 3D rotational C-arm (Fig. 1). As shown in [8], the metallic presence of the laparoscope does not produce artifacts affecting the region of interest. Then, an intraoperative 3D scan is performed and the laparoscope’s body is extracted from the intraoperative volume. This allows us to estimate directly the rigid registration between the laparoscopic camera and the intraoperative 3D imaging system (Artis Zeego, Siemens). This relationship is valid only as long as the camera remains static, which already routinely occurs at several stages of an intervention like a liver segmentectomy.

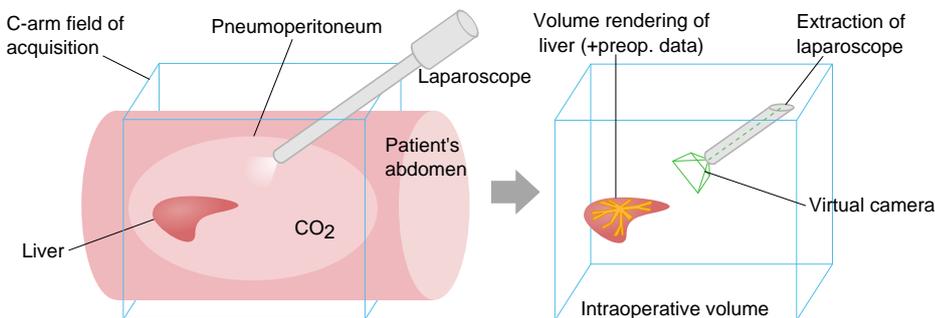


Fig. 1. Context of our method. As shown on the left, the laparoscope purposefully enters the acquisition field of the 3D rotational C-arm. Thus, as depicted on the right, our method allows us to extract the laparoscope from the resulting intraoperative volume, place a virtual camera accordingly and generate a virtual view of the organ of interest (usually with volume rendering) possibly including preoperative data. The laparoscopic image can then be augmented by superimposition with the virtual view.

The method [8] is very appealing but suffers from two main drawbacks. First, due to the tubular shape of the laparoscope, its roll angle cannot be determined from the intraoperative volume. This degree of freedom is estimated thanks to an accelerometer included in the camera, but this is *not* featured in most laparoscopes. Second, [8] assumes that the optical axis coincides with the revolution axis of the laparoscope, which may be violated depending on the model used, as illustrated in Fig. 2(a). This difference results in a 2D shift ϵ in the image plane. Though small at the scale of the device, it can yet result in up to several tens of pixels of registration error in the augmentation. Other parameters such as the zoom and focus also influence the position of the optical axis and thereby ϵ . A calibration dedicated to estimate ϵ is possible, but not relevant before the intervention. Indeed, many endoscopes are separable from the camera (Fig. 2(b)) and the surgeon may make it spin during the intervention to place the light cable upon desire (Fig. 2(c)), which changes ϵ . Likewise, the zoom and focus

of the endoscopic camera may also be changed intraoperatively and invalidate the preoperative estimation of ϵ . While performing a supplementary calibration intraoperatively might be feasible, the sake of preserving the workflow compels for a method purely based on image processing.

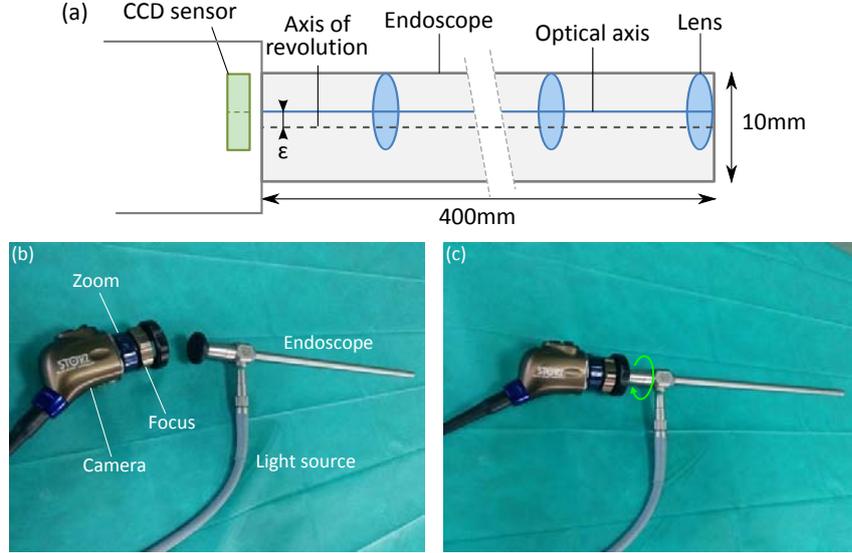


Fig. 2. Anatomy of a laparoscope. (a) A simplified illustration of the structure of a laparoscope presents an exaggerated misalignment ϵ between the optical axis and the axis of revolution. (b) Changing the zoom and focus influence ϵ . Also, the camera and the endoscope are separable. (c) As a result, both can rotate around each other once attached and ϵ varies.

In this paper, we present a novel method to complete [8]. It solves the previously mentioned registration issues using only information from the intraoperative volume and the laparoscopic image. As discussed, three degrees of freedom are to determine – the roll angle and the translation ϵ along the image axes. We propose to obtain these by optimizing a dissimilarity metric between the laparoscopic image and the view from the virtual camera upon the content of the intraoperative volume (Fig. 1). Given the relatively poor contrast between the different organs in an intraoperative CT image, the surface of the abdominal cavity is one of the most relevant information we could extract from the volume for the virtual camera. Since the cavity is insufflated with carbon dioxide, it presents a good contrast with the surrounding tissues and therefore extracting its surface is trivial, using for instance marching cubes.

Related work. There are three main ways to register an intraoperative surface to the laparoscopic image. One way is to use Shape-from-Shading (SfS), which reconstructs a surface from a single image based on the pixels' intensity and the

reflectance function [9, 10]. The reconstructed surface can then be registered to the surface extracted from the intraoperative volume using a method such as Iterative Closest Point (ICP). However, it has been established that SfS should not be used on its own in laparoscopic surface reconstruction [11], notably due to the falseness of the hypothesis of constant albedo throughout the scene. In our case, SfS would be overachieving since we do not need to reconstruct the surface from the laparoscopic image, but rather to design a dissimilarity metric between the intraoperative volume and the image. This enables us to use a local approach to shading and thus to alleviate the hypothesis of constant albedo (section 2.2). Another means to relate a surface with its image is simply to perform a correlation between their luminance using Mutual Information or an equivalent. However, the surface extracted from the intraoperative volume is textureless. There is thus no color information and approaches based purely on luminance are likely to fail (see Section 3 for experimental results supporting this assertion). We also cannot consider methods based on photo-consistency [12, 13] which has been successfully applied to endoscopic scenes [14], as two or more images are required.

Our proposed method to complete the registration uses a local formulation of the shading constraints. In the next section, we present the shading model and the formulation of the dissimilarity metric between the two inputs.

2 Methodology

This section describes the shading model used to determine the received light intensity. This model is simple because it is applied locally on the surface and uses piecewise constant albedo and piecewise constant light intensity hypotheses.

2.1 Shading Model

As illustrated by Fig. 3, the only light source inside the abdominal cavity is the one from the laparoscope, modeled as a point light source of position $S \in \mathbb{R}^3$ and intensity $l \in \mathbb{R}$ supposed constant locally. We consider Σ the surface extracted from the intraoperative volume and $\varphi \in C^2(\mathbb{R}^2, \mathbb{R}^3)$ the embedding of Σ which provides the surface point for each pixel $q \in \mathbb{R}^2$ in the laparoscopic image I . φ is known up to the sought pose of the virtual camera. The normal to Σ at φ is given by $\mathcal{N} \in C^2(\mathbb{R}^2, \mathbb{R}^3)$. In a typical laparoscopic image, there are often specularities and poorly lit areas. If we discard those (see section 2.2), it is reasonable to assume that the camera response is linear and therefore a quantity of light k is converted by the sensor into a pixel intensity given by $\tau(k) = ak, a > 0$. The albedo $\zeta \in C^0(\mathbb{R}^2, \mathbb{R})$, or surface reflection coefficient, is supposed constant on the surface locally for a same tissue and therefore $\zeta(q) = b, b > 0$. This is the classic limiting hypothesis in SfS, which we relax in section 2.2.

In a laparoscopic setting, the effect of illumination fall-off may be strong. We model this by dividing the amount of received light by the squared surface-to-light source distance. Assuming S and the origin O coincide, the illumination vector $\mathcal{L} \in C(\mathbb{R}^2, \mathbb{R}^3)$ at φ is thus given by:

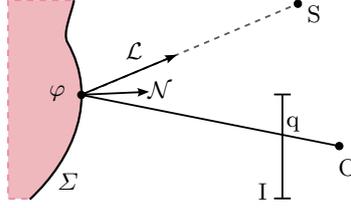


Fig. 3. Shading model. The point light source S emits a ray $-\mathcal{L}$ that hits the surface at φ . Assuming a Lambertian surface, the light is reflected with respect to the normal \mathcal{N} and the illumination vector \mathcal{L} . This reflection is projected onto the image plane I at q with O being the optical center and the origin of the world space.

$$\mathcal{L} = l \frac{\overrightarrow{\varphi S}}{\|\overrightarrow{\varphi S}\|^2} = l \frac{S - \varphi}{\|S - \varphi\|^2} = -l \frac{\varphi}{\|\varphi\|^2} \quad (1)$$

Assuming that the surface is Lambertian, the reflectance $\mathcal{R} \in C^2(\mathbb{R}^2, \mathbb{R})$ is given by $\mathcal{R} = \mathcal{L} \cdot \mathcal{N}$. Finally, using the camera response function τ , the intensity I of a pixel q is predicted by:

$$I = \tau \circ (\zeta \mathcal{R}) = ab(\mathcal{L} \cdot \mathcal{N}) = -c \frac{\varphi^\top \mathcal{N}}{\|\varphi\|^2} \quad \text{with } c = abl \quad (2)$$

Thus, based on reasonable assumptions about shading in the abdominal cavity, Eq. (2) is a simple solution to relating the surface to the luminance in the laparoscopic image. The coefficient c would ideally be a function of space as both albedo and light intensity vary in the scene. Therefore, we assume c to be constant only locally. The next section explains how this piecewise relationship between the surface and the laparoscopic image can be used in order to determine the three unknown registration degrees of freedom.

2.2 Shading-based Surface-Image Dissimilarity

Eq. (2) is valid for areas in the scene that are not extremely lit (specularities), unlit and for which the albedo is approximately constant. Therefore, we first apply a simple large median filter (23×23) on the 1080p laparoscopic image in order to robustly remove high frequencies (texture and specularities) while preserving the edges. Dark areas are discarded with a simple threshold on luminance. Satisfying the locally constant c requirement is equivalent to locally enforcing constancy for both albedo and intensity.

Therefore, we divide the image into a set \mathcal{P} of homogeneous patches using the watershed algorithm (Fig. 4). The distance between the watershed seeds is related to the size of the image and the kind of its content. In a typical laparoscopic scene filmed at 1080p, the size of the different organs is commonly above 100 pixels, due to the close-up view. Setting the seeds too coarsely would

result in missing small structures, while patches not large enough would not contain enough shading information and thus would fail at constraining the dissimilarity measurement. From our experience, a distance between the seeds of 150-200 pixels is ideal for 1080p laparoscopic images.

For each patch $p \in \mathcal{P}$, we use Eq. (2) at each pixel $q \in p$ to estimate c by linear regression. The resulting residuals constitute a least-squares cost function f_p that measures how well the laparoscopic image and the virtual view of the cavity surface concur for a patch p . The variable is the camera pose ω , which affects both φ and \mathcal{N} through the location of the coinciding points O and S.

$$f_p(\omega) = \arg \min_{c \in \mathbb{R}} \sum_{q \in p} \left\| I(q) + c \frac{\varphi_\omega(q)^\top \mathcal{N}_\omega(q)}{\|\varphi_\omega(q)\|^2} \right\|^2 \quad (3)$$



Fig. 4. Image processing and clusterization. The input laparoscopic image is undistorted (left), applied a median filter (middle) and divided into homogeneous patches by watershed (right). Dark areas are discarded (middle right in image).

Finally, we obtain the transformation $\hat{\omega}$ composed of the three sought degrees of freedom by minimizing the residuals for each patch $p \in \mathcal{P}$ in the global cost function F :

$$F(\omega) = \sum_{p \in \mathcal{P}} f_p(\omega) = \sum_{p \in \mathcal{P}} \left(\arg \min_{c \in \mathbb{R}} \sum_{q \in p} \left\| I(q) + c \frac{\varphi_\omega(q)^\top \mathcal{N}_\omega(q)}{\|\varphi_\omega(q)\|^2} \right\|^2 \right) \quad (4)$$

We solve $\arg \min_{\omega \in \mathbb{R}^3} F(\omega)$ by using a continuous numerical optimization algorithm (Powell's conjugate direction search in our case). The registration between the laparoscopic image and its virtual equivalent can thus be completed in rotation and translation, allowing an accurate augmentation of the surgical scene.

3 Experiments and Results

In the previous section, we proposed to minimize the cost function (4) in order to accurately register the laparoscopic image and the intraoperative volume.

Therefore, the success of our method also depends on the difficulty that optimization algorithms may have to find the global minimum in the search space. A couple of considerations ensure that an initialization at $(0,0,0)$ is close to the global optimum. First, the surgeon is very unlikely to rotate the laparoscope so much that the scene would be upside down. Second, the sensor cannot diverge too much from the laparoscope axis without hindering the completeness of the image captured out of the optics. An example of a clear global optimum in such a 3-dimensional search space around the initialization is illustrated by Fig. 5. These data originate from an *in vivo* acquisition of a pig’s liver, for which we applied our method. A total of three different acquisitions on three different pigs were performed. Each time, the intraoperative images were taken during breathhold. Results are displayed in Fig. 6.

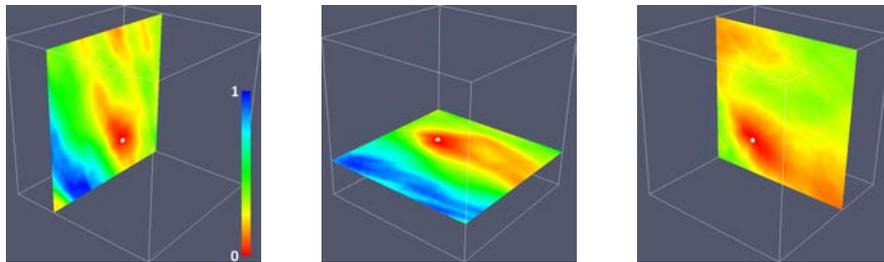


Fig. 5. Example of search space typically ranging $\pm 30^\circ$ and $(\pm 150)^2$ pixels. The cost issued by $F(\omega)$ is here normalized and colored from blue (high) to red (low). Sections are displayed along each of the three dimensions and passing by the global optimum (white dot).

For these experiments, one can notice the very good accuracy in registration achieved by our method. Over the three data sets, we performed manual measurements of the Target Visualization Error (TVE) by pointing 15 visual cues such as edges or corners in both images (Tab. 1). Our method proved to be more than twice as accurate than [8], with an average TVE of 11.3 ± 4.7 pixels in the image. This corresponds to less than a millimeter in the scene at nominal distance (around 70 mm). Thus, the remaining three degrees of freedom are accurately determined and so is the complete relationship between the laparoscopic image and the intraoperative 3D data, without additional apparatus or calibration. Typical optimization computation times range from 15 to 30s on a standard PC. Added to the initialization, the complete augmentation process takes between 25 to 55s.

Finally, in the introduction we asserted that classic 2D image-to-image registration methods such as Mutual Information would fail with such data. For the sake of verification, we calculated for each case the Normalized Mutual Information (NMI) between the endoscopic image and the surface view, while setting the translation to its correct value and varying only the angle. Similarly, to

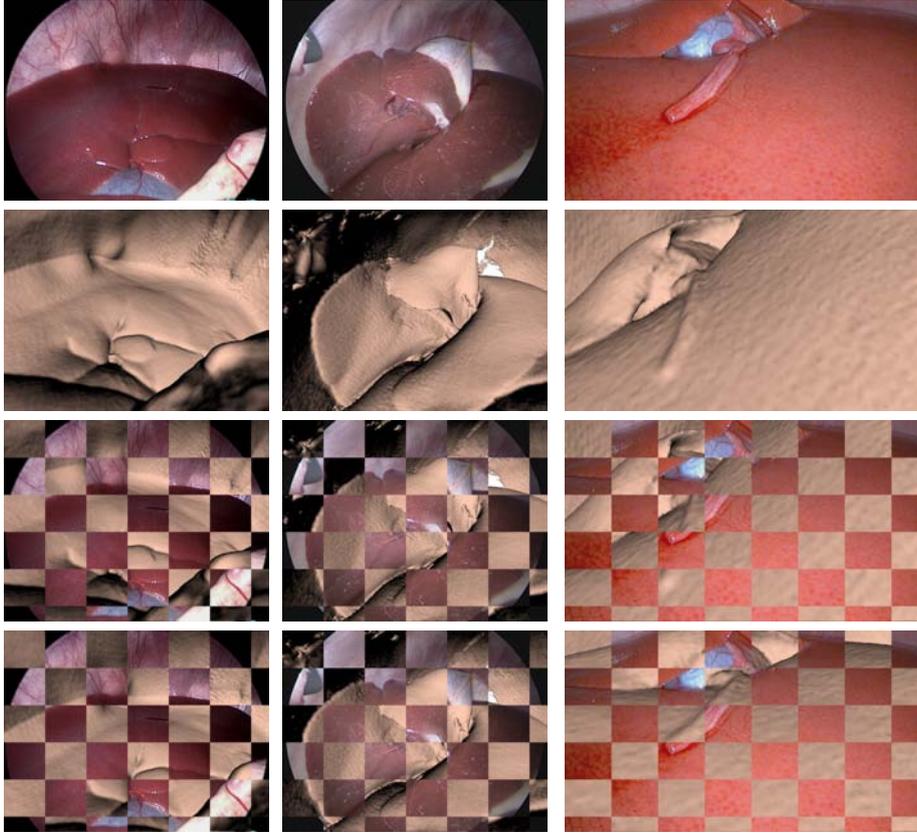


Fig. 6. The laparoscopic image (top) is registered with the view from the virtual camera upon the surface extracted from the intraoperative 3D data and rendered in VTK (middle top). A mosaic of the two shows the alignment before the proposed optimization (middle bottom) and after (bottom).

demonstrate the importance of a piecewise approach to shading, we calculated the proposed cost function $F(\omega)$ with globally constant c and piecewise constant c . These three cost functions are compared against each other in Fig. 7. One can notice that NMI does not show a global optimum for any of the three *in vivo* data sets. Moreover, our method with a globally constant c performs well only in Case 3, for which most of the laparoscopic image displays mostly only one organ and thus a same albedo.

4 Conclusion and Discussion

We have presented a novel method to complete a partial registration between a laparoscopic image and a surface extracted from intraoperative 3D data. When

	Initialization	Method from [8]	Proposed method
Case 1	123	13	6
Case 2	59	21	13
Case 3	>300	44	15
Average	>161 ± 124	26 ± 16.1	11.3 ± 4.7

Table 1. TVE (in pixels) manually measured across the three datasets at initialization at (0,0,0), after performing [8] and after the proposed method.

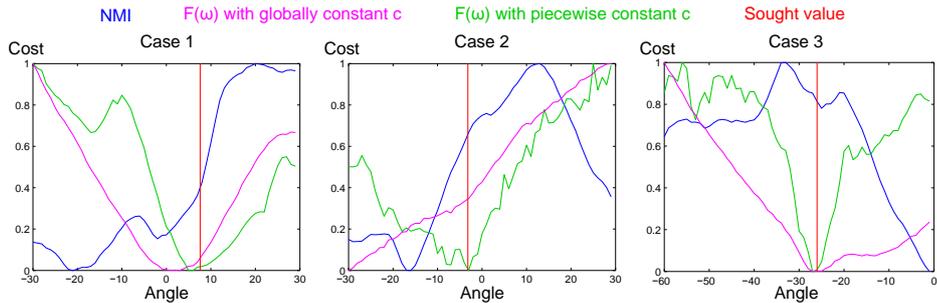


Fig. 7. Display of the normalized cost function in rotation only for NMI (blue), $F(\omega)$ with globally constant c (pink) and $F(\omega)$ with piecewise constant c (green). The graphs show that only our piecewise approach clearly displays a global optimum at the correct angle value (red) in all three cases.

combined with [8], we can provide a millimetric registration between the laparoscopic view and the intraoperative referential frame, using only standard hybrid operating room equipment and requiring no extra calibration process. This facilitates a fast and reliable augmentation of the scene with relevant information coming either from the intraoperative or the preoperative acquisitions.

So, while most shading methods aim at recovering the structure of the scene, we seek the camera pose. Thus, we do Pose-From-Shading rather than Shape-From-Shading. The concept of using shading to estimate the camera pose with respect to a known model is new. Moreover, most existing work on shading assumes a constant albedo over the whole image. It is obviously wrong in a typical intra-abdominal scene where different organs and tissues have different albedo and reflectance. This is why we propose this novel piecewise approach to shading, making it compatible with such scenes.

However, there is still room for improvement. First, the piecewise approach of our method makes it highly parallelizable and a GPU implementation would allow it to reach a shorter processing time. This would make our application more suitable for clinical applications, but also could compensate for breathing if real-time processing is achieved. Second, our approach obviously requires that the laparoscope tip has to show in the intraoperative scan. Although various experiments with surgeons have proved that doing so is not problematic for them,

we plan to investigate the possibility of extrapolating our work and determining all the six registration degrees of freedom only from the shading constraints. If not feasible in real time, and for the sake of providing a dynamic augmented reality solution in the hybrid operating rooms, we could also look into updating the augmentation with laparoscope tracking techniques such as SLAM or a robotic arm.

References

1. T. Sielhorst, M. Feuerstein, and N. Navab. Advanced medical displays : A literature review of augmented reality. *Journal of Display Technology*, 4(4):451–467, 2008.
2. M. Baumhauer, M. Feuerstein, H.-P. Meinzer, and J. Rassweiler. Navigation in endoscopic soft tissue surgery: Perspectives and limitations. *Journal of endourology / Endourological Society*, 22(4):751–766, 2008.
3. S. A. Nicolau, L. Soler, D. Mutter, and J. Marescaux. Augmented reality in laparoscopic surgical oncology. *Surgical oncology*, 20(3):189–201, 2011.
4. Peter Mountney, Johannes Fallert, Stephane Nicolau, Luc Soler, and Philip W Mewes. An augmented reality framework for soft tissue surgery. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, pages 423–431. Springer, 2014.
5. F. M. Sánchez-Margallo, J. L. Moyano-Cuevas, R. Latorre, J. Maestre, et al. Anatomical changes due to pneumoperitoneum analyzed by MRI: an experimental study in pigs. *Surgical and radiologic anatomy*, 33(5):389–396, 2011.
6. J. Bano, S. A. Nicolau, A. Hostettler, C. Doignon, J. Marescaux, and L. Soler. Registration of preoperative liver model for laparoscopic surgery from intraoperative 3D acquisition. In *Proc. of Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pages 201–210, 2013.
7. O. Oktay, L. Zhang, T. Mansi, P. Mountney, P. Mewes, S. Nicolau, L. Soler, and C. Chefhdotel. Biomechanically driven registration of pre- to intra- operative 3D images for laparoscopic surgery. In *Proc. of MICCAI*, pages 1–9, 2013.
8. Sylvain Bernhardt, Stéphane A Nicolau, Vincent Agnus, Luc Soler, Christophe Doignon, and Jacques Marescaux. Automatic detection of endoscope in intraoperative ct image: Application to AR guidance in laparoscopic surgery. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, pages 563–567. IEEE, 2014.
9. J.-D. Durou, M. Falcone, and M. Sagona. Numerical methods for shape-from-shading: A new survey with benchmarks. *Computer Vision and Image Understanding*, 109(1):22–43, 2008.
10. L. Maier-Hein, P. Mountney, A. Bartoli, H. Elhawary, D. Elson, A. Groch, et al. Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. *Medical image analysis*, 17(8):974–996, 2013.
11. T. Collins and A. Bartoli. Towards live monocular 3D laparoscopy using shading and specular information. In *Information Processing in Computer-Assisted Interventions*, pages 11–21. Springer, 2012.
12. M. J. Clarkson, D. Rueckert, D. L. G. Hill, and D. J. Hawkes. Using photo-consistency to register 2D optical images of the human face to a 3D surface model. *Trans. on Pattern Analysis and Machine Intelligence*, 23(11):1266–1280, 2001.
13. Z. Jankó and D. Chetverikov. Photo-consistency based registration of an uncalibrated image pair to a 3D surface model using genetic algorithm. In *Proc. of 3D Data Processing, Visualization and Transmission*, pages 616–622, 2004.

14. M. Figl, D. Rueckert, D. Hawkes, R. Casula, M. Hu, O. Pedro, D. P. Zhang, et al. Image guidance for robotic minimally invasive coronary artery bypass. *Computerized Medical Imaging and Graphics*, 34(1):61–68, 2010.