Video-endoscopic image analysis for 3D reconstruction of the surgical scene

A.M. Cano[1,2], P. Sánchez-González[1,2], F.M. Sánchez-Margallo[3], I. Oropesa[2,1]F. del Pozo[1,2]
and E.J. Gómez[1,2]

[1] Bioengineering and Telemedicine Group, Universidad Politécnica de Madrid, Madrid, Spain
[2] Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Madrid, Spain
[3] Minimally Invasive Surgery Centre (CCMIJU), Cáceres, Spain

Abstract — Video-endoscopic image analysis is an important source of information for extracting relevant 3D information of the surgical scene in minimally invasive surgery. This work proposes a new approach for getting depth information and relative position between instruments and organs based on video image analysis. The proposed image analysis extracts information from: (1) the projective geometry of the surgical setting in the image to determine laparoscopic tool's pose, and (2) the scene illumination captured in the image for getting a depth-map of the surgical scene. Results have verified the potential of video-analysis based methods for obtaining 3D information of the whole surgical scene. This information offers new possibilities for developing new augmented reality surgical applications and for improving surgical training with objective metrics. The main benefit of this approach is the lack of extra elements which disturb surgeons' performance in the clinical routine.

Keywords — video-endoscopic analysis, minimally invasive surgery, 3D reconstruction, tool tracking, organs reconstruction.

I. INTRODUCTION

Laparoscopic surgery is a minimally invasive technique now firmly embedded in routine practice [1]. This kind of surgery presents a different interaction paradigm to the traditional open approach due to the distant manipulation of specialized surgical instruments, and the indirect visualization of the surgical scene captured by an endoscope. These techniques have involved an important challenge for surgeons in getting used to reduced workspace environment, and a limited sensory interaction, making their development a difficult task.

Video-endoscopic analysis appears as an important source of information to extract relevant 3D information of the surgical scene in minimally invasive surgery (MIS), such as depth information and relative position between instruments and organs. This information, which is lost due to the indirect visualization of the surgery through a monitor, is useful in multiple potential applications for overcoming current laparoscopic performance.

Augmented reality in surgical environments can provide useful information to overcome surgeon's location and orientation difficulties for intra-operative procedures in MIS such as guiding and advising about the proximity of delicate areas with respect to the surgical instruments, something essential [2], [3].

On the other hand, actual surgical training programs lack standardized exercises for making an objective evaluation. The analysis of surgical instruments' trajectory and movements respect to the scenario offers parameters suitable to define objective metrics [4].

This work proposes a new approach for obtaining 3D information in MIS environments based on video-endoscopic analysis. In particular, this approach focuses on the 3D reconstruction of the laparoscopic scenario from two different points: (1) laparoscopic tool's position manipulated by surgeon, and (2) reconstruction of the abdominal cavity, obtaining depth and relative position between organs, where the intervention is developed.

A previous approach in scientific literature focused on laparoscopic tools' tracking, uses a LED placed in the tool tip that project laser dots onto the surface of the organs [5]. The optical markers on the surface are then detected in the endoscope image to allow localizing the instrument with respect to the scene. Other more recent works have tackled the problem without extra markers. Tonet et al. [6] proposed a heuristic approach to estimate the depth of the instrument by means of the knowledge of the tool width and its orientation. Authors have previously published two methods for assessing the 3D position of the tools' tips [7]. The first exploits the properties of the vanishing point of the tool in the image, and the second one, the visualized tools width relative to the distance to the camera.

In addition, one of the major challenges for laparoscopic visualization is the acquisition of the depth map associated with the surgical theatre. Some researchers propose the use of video-endoscopic images to obtain organs' shapes and relative distance to endoscope. These reconstruction techniques are named shape-from-X approaches, and use cues like motion, stereo, shading, focus, defocus, contours, etc. for extracting the shape of the objects [8-11]. In spite of the numerous techniques, not one of them is able to solve all 3D reconstruction issues in laparoscopic surgery without drawbacks.

In this paper, an image analysis of the surgical scene is proposed in order to obtain a 3D reconstruction of the whole surgical scenario. More specifically, a laparoscopic tool's tracking method and an anatomic shape reconstruction algorithm are proposed.

II. METHODOLOGY

A. Surgical scene analysis

A previous study of the characteristics of the laparoscopic setting and the geometry of the surgical scenario, allows us to extract relevant information about the scene.

The laparoscopic scenario consists of the following elements:

- **Abdominal cavity:** it represents the framework where the surgery will be carried out, and is composed by the abdominal organs under inspection. For an easier surgery performance, an enlarged space inside is created thanks the *neumoperitoneum*.
- **Trocar:** it defines the access points for laparoscopic instruments and endoscope into the cavity, and it avoids air from seeping out of the internal cavity.
- **Endoscope:** it illuminates and captures the surgical scene through a light source and a small camera located in the same point (tip of the endoscope).
- **Laparoscopic instrument:** it consists of a long cylinder with the tong in one extreme and a handle in the other. These instruments are manipulated by surgeons to perform the intervention.

The cylindrical geometry of the instruments and their projective contours in the image give information about its orientation and spatial position.

Moreover, the study of the illumination model presented in the laparoscopic scene allows us to determine a depth-map of the scene. This method assumes the existence of a unique light source point and analyses how the light is reflected in the organ's surface, which are considered lambertian surfaces (reflect light equally in all directions).

B. Laparoscopic tool's tracking

Proposed tracking of laparoscopic tools consists of determining 3D pose of the tool tip and its orientation with respect to the endoscope. It exploits geometrical properties of the projection of a cylinder shaped instrument in the image to calculate the camera-tool distance, and tool spatial orientation.

This method employs the knowledge of the tool's edges and its tip's 2D localization in the image obtained from the

previous image processing. Moreover, it requires knowing the Field of View of the camera that allows to determine spatially the projective line of any point in the image.

Due to the projective geometry of the laparoscopic scene (see Fig. 1.a), the cylindrical-shaped tool is visualized as a trapezoid in the image. Its lateral edges are defined in the image (projective edges, E1 and E2) by the intersection of the projective plane and two planes of sight which are tangential to the tool ($\Omega1$, $\Omega2$).
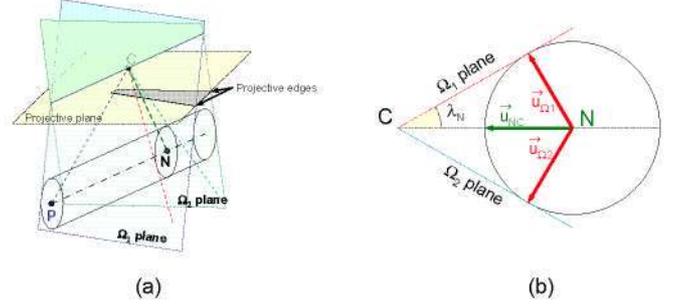

Fig. 1 (a) Projective model of the laparoscopic tool. C: optical center, P: tip of the laparoscopic tool, N: point of the cylinder axis whose projective line is perpendicular to this axis, Ω_1, Ω_2: planes of sight of the tool edges. (b) Transversal section of the laparoscopic tool at N point.

On the other hand, the perpendicular plane to the cylinder's axis that contains the optic centre (C) defines a circular section of the cylinder as shown in Fig. 1.b. Let N be the centre of this circle. Moreover, in this section tangential planes of sight define the $2\lambda_N$ angle.

The normal vector ($\vec{u}_{\Omega1}$, $\vec{u}_{\Omega2}$) to each tangential plane is obtained through cross product between border direction in the image ($\vec{u}_{E1}$, $\vec{u}_{E2}$) and the camera vector to any point of this border ($\vec{u}_{CE1}$, $\vec{u}_{CE2}$), because both vectors are included each Ω_i plane.

$$\vec{u}_{\Omega i} = \vec{u}_{Ei} \times \vec{u}_{CEi} , \; i = 1, 2 \qquad (1)$$

Then, as shown in Fig. 1.b the addition of vectors, $\vec{u}_{\Omega1}$ and $\vec{u}_{\Omega2}$, results in the direction of CN vector. Therefore angle λ_N is expressed like:

$$\tan \lambda_N = \frac{|\vec{u}_{\Omega1} + \vec{u}_{\Omega2}|}{|\vec{u}_{\Omega1} - \vec{u}_{\Omega2}|} \qquad (2)$$

According to Fig. 1.b, it is possible to determine the distance between the camera and the tool axis, |CN|, from the tool radius (r_{CYL}):

$$|CN| = r_{CYL} / \sin \lambda_N \qquad (3)$$

Finally, CP vector is determined like:

$$\overline{CP} = \frac{|CN| \cdot \vec{u}_{CP}}{\vec{u}_{CP} \cdot \vec{u}_{CN}} \qquad (4)$$

Note that P can be any point of the tool, it only requires to be identified in the image. And as $\vec{u}_{NP}$ is known, any other point of the tool P' (eg. the tip) can be obtained by knowing the physical distance between P and P'.

C. Anatomic shape reconstruction

The method proposed in this section addresses the anatomic shape reconstruction problem based on the illumination information. Thus, it can be classified as a Shape from Shading (SfS) technique and pursues the aim of obtaining a depth-map of the surgical scene from a monoscopic image.

The reconstruction process analyzes light reflected by organs' surface. Considering negligible the indirect light intensity, it can be proved that the simplified illumination equation of the surgical scene is [11]:

$$I = k_d \cdot L \cdot \cos \alpha \qquad (5)$$

where k_d is a surface intrinsic constant, L is the emitted power per unit solid angle, and α is the angle between the light source and surface normal vectors.

The reconstruction approach is shown in Fig. 2. As exposed there, depth variation (Δz) through pixel p is proportional to the tangent of the angle between the light source vector (l) per solid angle (θ,φ) and the surface normal vector (N) (angle α).
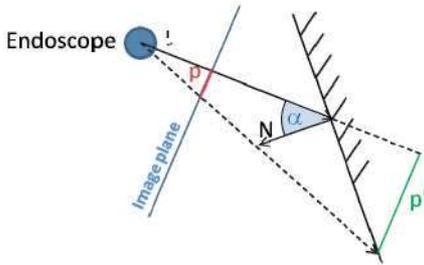

Fig. 2 Variation of the distance to the endoscope according to the tangent of the angle (α) between the visual (θ,φ) and the surface normal (N).

The first step in the process is to compute the tangents of angle α, the intensity slopes. These slopes can be obtained calculating the intensity gradient of each pixel in a given direction of the image and then dividing this gradient by the intensity in the same direction.

The next stage in the algorithm is the surfaces' shape reconstruction from the intensity slopes, which can be achieved adding all the slopes in the gradient direction, considering an orthographic camera system.

III. RESULTS AND DISCUSSION

A. Laparoscopic tool's tracking

Two validation sequences are used to evaluate the performance of the laparoscopic tool's tracking method. They consist on two different movements, "Constant depth movement" and "Variable depth movement", acquired in a laboratory surgical setting. This scenery allows us to determine the ground truth of the tool's tip position thanks to the introduction of auxiliary elements, like a background plane (a grid of points) and an inclined board (30º). Each sequence has about 275 frames.

Both sequences have been analyzed with proposed method, and accuracy of assessed 3D coordinates is given by comparing them to the ground truth (see table 1).

Results obtained show a good behaviour of proposed method, whose standard deviation is lower than 2.09 and 3.55 in constant depth movements and variable depth ones. However, it presents a constant error lower than 9 mm. (variable depth movements), whose origin is probably some inaccuracy in the tool edges detection in the image. The presence of blurred tool's edges caused by its movements limits this edge detection, adding an error in the later 3D pose estimation. A design of a robust edge detection method could improve current accuracy.

Table 1 Error characterization of the method for validation sequences. Mean error and standard deviation (SD), in mm.

Axis	Constant depth		Variable depth	
	Mean	SD	Mean	SD
X	0.08	0.08	0.03	1.40
Y	0.35	1.25	1.54	3.66
Z	2.03	2.09	8.65	3.55

In our previous works [7], two methods were developed exploiting the properties of the vanishing point of the tool in the image, and the visualized tools width relative to the distance to the camera. In particular, the second approach presents a similar stability to the one presented in this paper (SD lower than 3 mm.), but it requires visualization without obstacles of the tool tip, limiting its potential use in respect to the proposed method. A combination of both methods could improve the behaviour of each other, since the 3D localization is determined through different features of the tool.

B. Anatomic shape reconstruction

The original laparoscopic image and the obtained depth-map of the surgical scene are shown in Fig. 3. Simplicity and low computational cost (lower than 1 sec.) are two important characteristics of the reconstruction method proposed, taking into account the required real-time conditions.
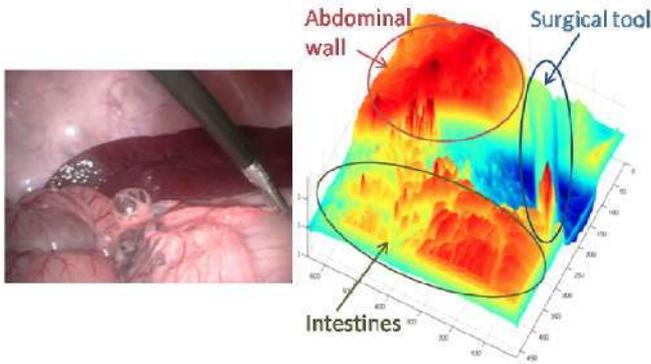

Fig. 3 Original image and its depth map.

Although the proposed method only extracts relative distances from surfaces to the endoscope, this limitation can be solved knowing one real depth value per continuous region (e.g. applying the method exposed in section II-B when tool tip is in contact with an organ) and propagating it.

The SfS algorithm proposed is also vulnerable to a sign ambiguity (concave/convex ambiguity). This ambiguity leads to the same shading pattern results if the surface is reversed in depth and illuminated from a mirror-symmetric direction. In SfS techniques, it occurs when a lambertian surface is illuminated under a light source with a camera system located on the same spatial point.

Both problems can be addressed mixing information provided by other shape-from-X techniques: mainly motion and texture, which can deal with absolute distance and sign ambiguity problems, respectively.

IV. CONCLUSIONS

This work presents a new approach for obtaining 3D information in MIS based on video-endoscopic analysis. On one hand, analysis of projective geometry of the laparoscopic setting allows us to define a laparoscopic tool's tracking method. On the other hand, an anatomic shape reconstruction algorithm based on the scene illumination is proposed.

Results have verified the potential of video-analysis based methods for obtaining 3D information of the whole surgical scene. The main benefit of this approach is the lack of extra elements which disturb surgeons' performance in the clinical routine. This work is a first step for 3D reconstruction of the surgical scene.

REFERENCES

1. Cuschieri A, (2005) 'Laparoscopic surgery: current status, issues and future developments', Surgeon, 3(3):125-138.
2. Blum T, Sielhorst T, and Navab N, (2007) 'Advanced augmented reality feedback for teaching 3D tool manipulation'. New Technology frontiers in minimally invasive therapies, : 223-236.
3. Fuchs KH, Livingston MA, Raskar R, Colucci D'nardo, Keller K, State A, Crawford JR, Rademacher P, Drake SH, Meyer A (1998), 'Augmented Reality Visualization for laparoscopic surgery', Lecture Notes in Computer Science, 1496: 934-943.
4. Rosen J, Brown JD, Chang L et al. (2006) 'Generalized Approach for Modeling Minimally Invasive Surgery as a Stochastic Process Using a Discrete Markov Model'. IEEE Transactions on Biomedical Engineering, 53(3):399-413.
5. Krupa A, Gangloff J, Doignon C, de Mathelin M, Morel G, Leroy J, Soler L, Marescaux J, (2003), 'Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing'. IEEE Transactions on Robotics and Automation, 19(5): 842-853.
6. Tonet O, Armes TU, Magali G, Dario P (2005), 'Tracking endoscopic instruments without localizer: image analysis-based approach' Stud. Health Technol. Inform., 119:544-549.
7. Cano AM, Lamata P, Gayá F, del Pozo F, Gómez EJ (2006) 'New Methods for Video-Based Tracking of Laparoscopic Tools', Lecture Notes in Computer Science, 4072: 142-149.
8. Durou J-D, Falcone M, Sagona M (2008) 'Numerical Methods for Shape-from-shading: A New Survey with Benchmarks'. Computer Vision and Image Understanding, 109(1): 22-43.
9. Gurnsey R, Poirier FJ, Bluett P, Leibov L. (2006) 'Identification of 3D shape from texture and motion across the visual field', Journal of Vision, 6(5): 543-553.
10. Pradeep KS, Rajagopalan AN. (2007) 'Improving shape from focus using defocus cue'. IEEE Trans Image Processing, 16(7):1920-1925.
11. Sánchez-González P, Gayá F, Cano AM, Gómez EJ, (2008), 'Segmentation and 3D reconstruction approaches for the design of laparoscopic augmented reality environments', Lecture Notes in Computer Science, 5104: 127-134.

Author:	Alicia M. Cano González
Institute:	Grupo de Bioingeniería y Telemedicina, Universidad Politécnica de Madrid
Street:	Ciudad Universitaria s/n
City:	Madrid
Country:	Spain
Email:	acano@gbt.tfo.upm.es