# New Methods for Video-Based Tracking of Laparoscopic Tools

Alicia M. Cano, Pablo Lamata, Francisco Gayá, and Enrique J. Gómez

Grupo de Bioingeniería y Telemedicina (GBT),
ETSIT, Universidad Politécnica de Madrid
c/ Ciudad Universitaria s/n, 28040 Madrid, Spain
{acano, lamata, fgaya, fpozo, egomez}@gbt.tfo.upm.es
http://www.gbt.tfo.upm.es

**Abstract.** New methods for video-based tracking of laparoscopic instruments are presented. This aims to contribute for an objective evaluation of surgical skills, and for enabling augmented reality features in laparoscopic surgery. An optical geometrical model of the laparoscopic setting is developed, what is the basis for two methods proposed for assessing the 3D position of tools' tips. The first exploits the properties of the vanishing point of the tool in the image plane, and the second the apparent tool width with the distance to the camera. Ground-truth sequences are recorded and analysed, and preliminary results demonstrate the feasibility of these methods. Video-based approach constitutes a very promising alternative to mechanical, optical or electromagnetic tracking devices.

## 1 Introduction

Laparoscopic surgery is currently performed in clinical routine at hospitals. Nevertheless there is a crescent pressure to have transparent training programs, with objective metrics of surgical skill and alternatives that might be used at any time [1]. Laparoscopic training programs lack objective metrics for skills' assessment. Analysis of path and movements travelled with laparoscopic tools is an important source of objective parameters [2].

On the other hand laparoscopy offers a reduced workspace and a limited sensory interaction. One interesting improvement in this surgical technique would be the use of augmented reality techniques. This could provide useful information for guiding the surgical procedure or preventing delicate areas. A clear requisite for such applications is the knowledge of the position of laparoscopic tools in real time.

Therefore, the problem addressed is the tracking of laparoscopic instruments with only the information of the conventional video signal captured by an endoscope. This aims to contribute both for an objective evaluation of surgical skills, and for enabling the possibility of augmented reality features in laparoscopic surgery. Another important benefit is the improvement of studies for biomechanical characterization by reducing the complexity and cost of physical tracking instruments, which can be very bulky [3]. This is especially interesting for in-vivo studies in the laparoscopic operating theatre.

M. Harders and G. Székely (Eds.): ISBMS 2006, LNCS 4072, pp. 142 – 149, 2006.

Video-based tracking of laparoscopic tools can be decomposed in two main problems: (1) extraction of image parameters with segmentation techniques and (2) estimation of the 3D coordinates of the end of the tool with these parameters. There are some works already published in the literature about these two aspects.

Voros et al [4] have created an automatic and robust method to detect the edges of laparoscopic tools in real sequences. This method is based in a Hough transform fastened by the restriction of the localization of the trocar point. On the other hand, Tonet et al [5] have designed a method which is able to localize an instrument in 3D with regard to the position of the camera with an extra marker in the instrument. They propose an empirical estimation of the depth of the instrument based on the position and orientation of the instruments.

This work contributes with a geometrical optical model of the laparoscopic setting, which is the basis for two methods proposed for assessing the 3D position of tools' tips. This model is a conceptualization of the problem that encompasses former works [4,5].

## 2 Material and Methods

Two methods for assessing the 3D position of a tool tip are proposed. They depart from the knowledge of the edges of the instrument and its 2D localization of the tip, what would be the result of a video segmentation process. Moreover, the knowledge of the Field of View (FoV) of the camera permits to determine spatially the projective line (line from camera to the spatial point) for any image point.

### 2.1 Vanishing Point Based Method

This method gets the 3D coordinates of the tip of the tool through the information about the surgical setting (position of camera and trocar), the optical characteristic of the camera (field of view), and the 2D parameters extracted from the segmentation process of each frame (edges of tools and their tip point). It is developed from the geometrical model depicted in Fig. 1.

In this perspective, the point at infinity, also called vanishing point (V), is where the edge lines of the instrument meet in the image plane. The projective line of this point has the same direction vector of the tool. This is due to the fact that all 3D lines parallel to the tool ($\overline{TP}$) converges at this vanishing point in the image plane [6], and therefore the 3D line containing the camera point does ($\overline{CV}$, its projection is a single point). In this way, the instrument edges give information about the tool direction. ($\overline{CV}$ and $\overline{TP}$ are parallel vectors in Fig. 1)

Using the main idea, it is also built the optical ray to the tip of the tool, getting the director vector to the tip of the instrument from the camera. Along these lines, 2D localization of the tip and the FoV provide information about the direction vector camera-tip of the tool. (See $\overline{CP'}$ vector in Fig. 1)

Placing both vectors (director vector of the tool and the director vector camera-tip of the instrument) in the trocar and camera point (which are known), respectively, the position of the tip of the instrument is solved, like the intersection between both lines.

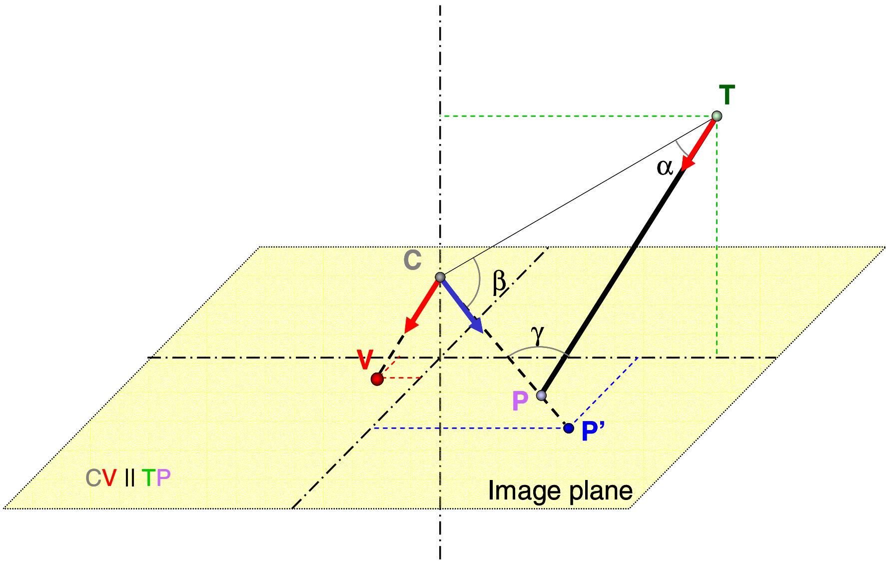


**Fig. 1.** Geometrical optical model of the working scene. C: optical center, TP vector: surgical tool, P: tool tip. P': projection of the tool tip. V: vanishing point.

Finally the vector $\overline{CP}$ can be expressed:

$$\overline{CP} = \left|\overline{CT}\right| \cdot \frac{\overline{CP'}}{\left|\overline{CP'}\right|} \cdot \frac{\sin(\alpha)}{\sin(\gamma)} = \left|\overline{CT}\right| \cdot \vec{u}_{CP'} \cdot \frac{\sin(\alpha)}{\sin(\gamma)} \tag{1}$$

## 2.2 Apparent Tool Size-Based Method

This method determines the tools' location according to the diameter size of the instrument in the image and the angle defined by the projective lines to two opposite edge points.

Knowing the real distance between these points (diameter of the instruments) and the angle defined by the camera as vertex and the two opposite edge points placed in the chosen position in the tool (α in the Fig. 2), it is possible to calculate the distance from this position to the camera (height of a isosceles triangle once known the basis). This is the module of the vector that joins the tip of the instrument (P) with the camera (C). Moreover, the projective line to this position is known, so it is possible to determine the coordinates of the tip of the vector.

Consequently the $\overline{CP}$ vector is calculated:

$$\overline{CP} = \left|\overline{CP}\right| \cdot \frac{\overline{CP'}}{\left|\overline{CP'}\right|} = \frac{Physical\,Tool\,Width\,(mm)}{Image\,Tool\,Width\,(pix)} \cdot \overline{CP'} \tag{2}$$

In conclusion, we can apply these methods thanks to the geometry of the endoscopic instruments. These instruments present a cylindrical shape which allows

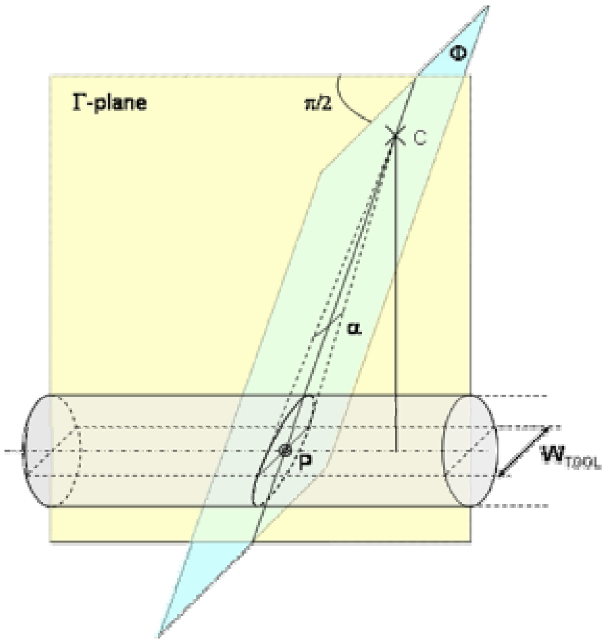


**Fig. 2.** Section of the instrument according to the projection line to the point P

knowing its diameter in the image plane in all situations and the perspective of its edge gives information about its direction.

### 2.3 Validation Process

The validation of both methods is done using two sequences of movements acquired in laboratory. This is necessary for the definition of a ground truth of the positions of the tools' tip in each frame without an external tracking device.

A laparoscopic setting is built, see Fig. 3, emulating the same conditions than a real surgery: relative positions between camera and trocar points are carefully respected. The key idea is to introduce a physical guide for the movements that relates the 2D coordinates with the depth. This physical guide is an inclined board (30º) placed at a given distance from the optical focus point (C point in the model). 3D trajectory of the tool keeping the tip alongside this board is defined with only the 2D information gathered from the segmentation process. Therefore a ground-truth of 3D coordinates from 2D information is defined and used to validate proposed methods. Two sequences are recorded (see Fig. 3): one with a constant depth (Y movement) and other with a variable depth (XZ movement). Each sequence has about 275 frames, around 10 seconds of recording.

Segmentation of each frame of the sequence is necessary for arriving into the estimation of the 3D coordinates. This is roughly done with digital derivate masks and temporal information gathered from former frames.

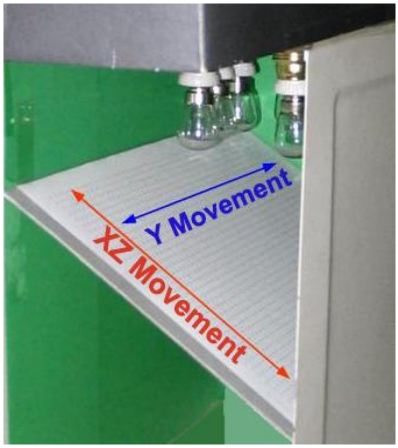


**Fig. 3.** Setting for recording lab sequences valid for a ground-truth definition

The recording was carried out with a home-videocamera and the sequence was analysed with Matlab program on a Pentium IV (3.00 GHz, 512 Mb RAM) computer.

# 3 Results

The two lab sequences have been analysed with the two proposed methods. Estimated 3D coordinates have been estimated in a total of 550 frames. Accuracy has been assessed comparing these estimations with the ground-truth trajectory by two means, regression lines for the variable-depth sequence (see Fig. 4) and error characterization for both sequences (mean absolute error and standard deviation, see Table 1 and Table 2).

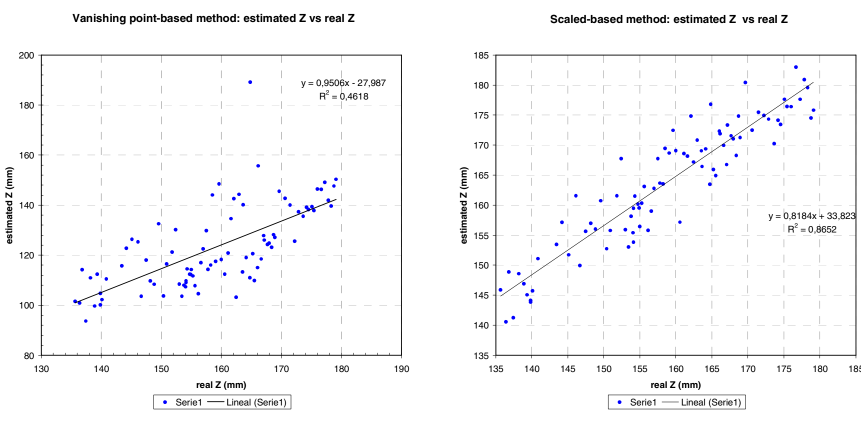


**Fig. 4.** Estimated Z coordinate versus ground-truth Z coordinate with both methods in variable-depth sequence

**Table 1.** Error characterization for the variable-depth sequence (SD: standard deviation)

| | Vanishing point | | | Apparent size | | |
|---|---|---|---|---|---|---|
| | \|X error\| | \|Y error\| | \|Z error\| | \|X error\| | \|Y error\| | \|Z error\| |
| mean (mm) | 5.59 | 0.84 | 33.60 | 0.67 | 1.59 | 6.65 |
| SD (mm) | 3.55 | 0.97 | 9.26 | 0.56 | 0.52 | 3.85 |

**Table 2.** Error characterization for the constant-depth sequence (SD: standard deviation)

| | Vanishing point | | | Apparent size | | |
|---|---|---|---|---|---|---|
| | \|X error\| | \|Y error\| | \|Z error\| | \|X error\| | \|Y error\| | \|Z error\| |
| mean (mm) | 2.20 | 7.26 | 43.69 | 0.68 | 1.90 | 3.00 |
| SD (mm) | 1.12 | 7.28 | 16.30 | 0,13 | 0.76 | 2.19 |

The "apparent tool size-based" method has shown a better behaviour than vanishing point-based one in all comparison metrics: R^2 regression coefficient, mean absolute error and standard deviation of absolute error (see Table 1 and Table 2). This method also shows a better behaviour in the constant-depth sequence compared to the variable-depth sequence. The vanishing point-based shows similar features in both sequences.

## 4 Discussion

Two methods have been proposed for assessing the 3D coordinates of the tip of laparoscopic tools. They are simply based in an analysis of the video sequences, and constitute a very interesting alternative to mechanical, optical or electromagnetic tracking devices.

Results have demonstrated the feasibility of these methods and of the video-based alternative, as recently concluded in a previous work [5]. Moreover, the developed geometrical optical model has provided an insight of the problem. This has enabled the development of the vanishing point-based method, which uses the information of the orientation of the lateral edges of the tool. Previous empirical results and formulas [5] are now understood from the right perspective provided by this model. Moreover, projection of the insertion point in this model is a crucial piece of information for restricting the Hough transform for segmentation purposes, as already addressed in [4].

The methods have been validated in two lab sequences and not in real laparoscopic ones. This has been required for the definition of a ground truth of the positions of the tools' tip in each frame of the sequences. This is a right alternative because the aim was to assess the feasibility of assessing the 3D coordinate with the 2D information extracted from a frame, and not to develop a robust segmentation technique for obtaining this 2D information from a laparoscopic frame.

The accuracy of proposed methods, which has been assessed in the two analysed sequences, could be qualified as moderate. Standard deviation of absolute errors has been lower than 16.3 mm and 3,85 mm in respectively the “vanishing-point-based” and “apparent tool size-based” methods. It has to be regarded that this accuracy depends on the result of the segmentation process, which could be a little rough in the implementation that has been used.

Both methods have showed a mean absolute error in the estimation of the 3D coordinates. The origin and behaviour of this bias is probably some uncertainty in the parameters of the models, like the relative position between of the camera focus point and the trocar point. There are also some corrections that need to be done in order to prevent some optical distortions.

The more noisy estimation of the Z coordinate of the “vanishing-point-based” method has been explained with its sensibility to errors in the slope of detected tools’ edges that determine the vanishing point. This is more critic in those frames in which there is only a small portion of the tool depicted. However, the “apparent tool size-based” method is quite noisy in those situations in which the tool is far from the camera, when its size is small and the pixel quantification error is bigger. Therefore a combination of both methods, together with a robust image segmentation process, seems to be a good alternative to be studied in future work in order to get a good accuracy.

According to the results, both methods present a low accuracy in the present state of development. However, most of the laparoscopic surgery applications could need accuracy around 2-3 mm., for example in laparoscopic training programs. Nevertheless the required accuracy depends on the desired objective of the particular application, it rarely needs accuracy better than 0,5 mm, due to factors as tool movement magnification because of pivoting on the insertion point.

Other interesting issues to be addressed are the automatic determination of the 3D coordinates of the trocar and camera points, the tracking of camera movements, and the optimization of all processes in order to reach a real-time application. This will eventually lead to a complete system for tools tracking.

## 5 Conclusion

Laparoscopic tools’ tracking is possible with an analysis of conventional video sequences. Developed geometrical optical model provides an insight of this problem. Preliminary results of the two proposed alternatives for assessing the 3D coordinates with 2D information are very promising.

## Acknowledgement

This research work has been partially funded by the SINERGIA Thematic Network (G03/135) of Spanish Ministry of Health.

Authors would like to thank the valuable contributions of Patricia Sánchez and Samuel Rodríguez.

## References

1. Kneebone,R.: "Simulation in surgical training: educational issues and practical implications". Medical Education, 37(3):267-277, 2003.
2. Rosen,J., J.D.Brown, L.Chang, M.N.Sinanan, B.Hannaford: "Generalized Approach for Modeling Minimally Invasive Surgery as a Stochastic Process Using a Discrete Markov Model". Biomedical Engineering, IEEE Transactions on, 53(3):399-413, 2006.
3. Immersion: Digiziting. www.immersion.com/digitizer/
4. Voros S., Orvain E., Cinquin P., 'Automatic detection of instruments in laparoscopic images: a first step towards high level command of robotized endoscopic holders'. BIOROB'06 1st IEEE/RAS-EMBS Int. Conf. On Biomedical Robotics and Biomechatronics, Pisa, Italy, February 2006.
5. Tonet O., Armes T.U., Magali G., Dario P. 'Tracking Endoscopic Instruments Without Localizer: Image Análisis-Based Approach' Stud. Health Technol. Inform., 119:544-549, 2005.
6. Casas J.'Óptica'. University of Zaragoza. Pons, 1994.