



HAL
open science

Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder

Sandrine Voros, Jean-Alexandre Long, Philippe Cinquin

► **To cite this version:**

Sandrine Voros, Jean-Alexandre Long, Philippe Cinquin. Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006, Oct 2006, Copenhagen, Denmark. pp.535-42. hal-00385051

HAL Id: hal-00385051

<https://hal.science/hal-00385051>

Submitted on 9 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder

Sandrine Voros¹, Jean-Alexandre Long², and Philippe Cinquin¹

¹ TIMC-IMAG, UMR CNRS 5525, Université Joseph Fourier, Grenoble
[Sandrine.Voros,Philippe.Cinquin]@imag.fr

² Department of Urology, University Hospital, Grenoble

Abstract. The use of a robotized camera holder in laparoscopic surgery allows a surgeon to control the endoscope without the intervention of an assistant. Today, the orders that the surgeon can give to robotized camera holders remain limited. In order to provide higher level interactions between the surgeon and a robotized camera holder, we have developed a new method for the automatic tracking of laparoscopic instruments which works in near real-time. The method is based on the measurement of the 3D positions of the insertion points of the instruments in the abdominal cavity and a simple shape model of the laparoscopic instruments. We present the results of our first experimentation on a cadaver.

1 Introduction

Laparoscopic surgery offers a lot of benefits for the patients compared to open surgery: less blood loss, less transfusions, a smaller consumption of analgesia and a shorter hospitalisation time [1]. However, because of the limited field of view, the specific instruments and the difficult coordination between the assistant manipulating the camera and the surgeon, laparoscopy is a more complex procedure for the surgeon than open surgery. Robotized systems were developed to assist the surgeon in performing laparoscopy, including robotized camera holders: they stabilize the image [2], reduce the staining of the endoscope and enable solo surgery. Several robotized endoscopic holders have been commercialised, including Aesop (Computer Motion) and EndoAssist (Armstrong Healthcare) [3].

Today, the applications allowing the surgeon to control robotized endoscopic holders only integrate simple commands. Providing higher level commands of the robot would be a major clinical added value: it would allow the surgeon to concentrate on his surgical procedure, rather than on the control of the robot. In this objective, the automatic tracking of the instruments is an intuitive approach, since the moving instruments must stay in the field of view of the camera for security reasons.

Several approaches were proposed to detect the instruments in laparoscopic images: [4] designed a robotized instrument holder equipped with a laser beam. The laser beam is detected in the images in order to track the tip of the instrument. This approach raises the questions of cost issues and sterilisation. [5], [6] and [7] taped a color mark on the instrument. The main limit of this approach

is the sterilisation of the mark. These four approaches put constraints and add complexity to the procedure. [8] used color classification to identify colors that are likely to correspond to instruments, but the computation time is not indicated. [9] proposed to detect the grey regions in the laparoscopic images, which are likely to correspond to the metallic tip of the instrument. A good computation time is achieved. However, some laparoscopic instruments do not have a metallic tip and the tip might be hidden behind an organ. Finally, [10] proposed an approach based on the detection of the lines in the image, but some parameters have to be tuned by the user. None of these approaches give solutions to identify the instrument that has to be tracked among the other instruments.

Our approach is based on the measurement of the 3D positions of the insertion points of the instruments. With a calibrated robot and a calibrated camera, we can project the insertion point of the instrument we wish to track on the image plane of the camera (even if the insertion point is not visible on the image). This information provides an important constraint on the position of the instrument on the image. It is used to automatically detect the axis, edges and tip of the instrument. The knowledge of the 2D position of the tip allows to perform a visual servoing of the orientation of the camera. For this purpose, we do not require subpixel precision but rather need to roughly determine the position of the tip. As suggested in [7], the orientations of the edges can be used to find the depth of the instrument and include the zoom in the visual servoing loop. [11] showed encouraging results on laparoscopic images and presented our visual servoing on a test bench. In this paper, we focus on the results of our first experimentation on a cadaver.

2 Method

2.1 Calibration of the robotized laparoscopic holder and the camera

Our robotized camera holder [12] was calibrated, as well as the endoscopic camera [13]: the 3D position of the camera is known in a fixed frame linked to the robot (fig. 1).

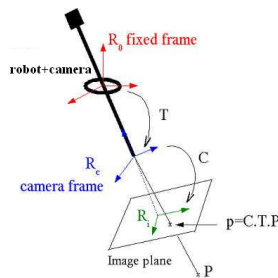


Fig. 1. Calibrated system: if the coordinates of P are known in R_0 , its projection on the image plane can be computed thanks to the calibration matrixes T (geometric model of the robot) and C (pinhole model of the camera).

The 3D coordinates of a point can be measured by stereovision: two views of this point for different positions of the robot are sufficient.

2.2 Measurement of the 3D positions of the insertion points of the instruments

The 3D positions of the insertion points of the instrument are measured at the beginning of the intervention. Since the surgeon gives orders to the robot with a vocal command, we have developed a ‘vocal mouse’ which allows to move a cursor on the endoscopic image. For two different positions of the camera in which the insertion point is visible, the surgeon selects the position of the insertion point in the image with the vocal mouse.

This initialisation step could be easily integrated to the surgical protocol since the surgeon creates the insertion points at the beginning of the intervention under visual control. We show in section 3 our measurements of the 3D position of an insertion point: the hypothesis that they are relatively fixed during an intervention is validated.

2.3 Model of the laparoscopic instruments in the images

We define $P = proj(T)$ as the projection of an insertion point T in the image plane of the camera. In a laparoscopic image, an instrument can be represented as (fig. 2) :

- three lines: a symmetry axis Δ , and two edges which are symmetrical compared to Δ . The positions of these three lines are constrained by the position of P : C is a circle centered on P of radius R such that the symmetry axis and the two edges intersect C ;
- a point S , which represents the tip of the instrument. S must belong to Δ .

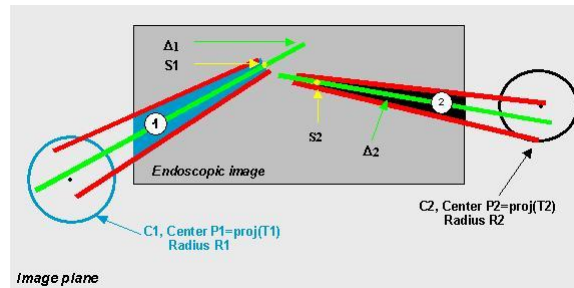


Fig. 2. Representation of the instruments in a laparoscopic image. The position of each instrument in the image is constrained by the position of the projection of its insertion point in the abdominal cavity.

The value of R is determined by a simulation: since the diameter d of a laparoscopic instrument is known (usually 6 mm), we bounded the value of R

from above by the diameter d_0 in pixels the tool would have in the image if it was parallel to the image plane. Our simple model of the abdominal cavity, the instrument and the camera is illustrated by fig. 3, as well as the simulation results.

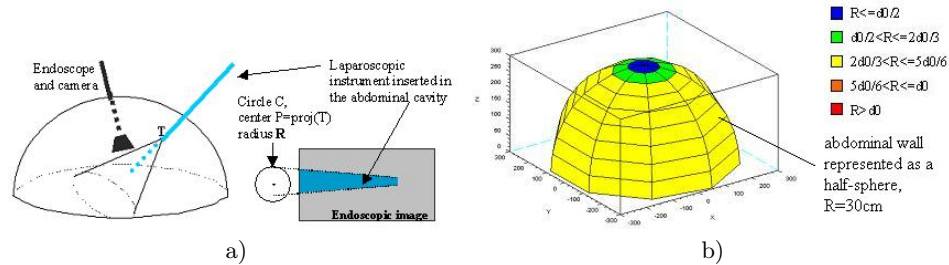


Fig. 3. a) Model of the abdominal cavity, camera and instrument to validate our choice for the value of R , b) result of the simulation: false color representation of R depending on the position of the instrument on the abdominal wall

To detect the instrument inserted through T , we search for a symmetry axis and two edges which intersect C of radius R , and we search for the tip S of the instrument on the symmetry axis.

2.4 Main steps of the detection algorithm

We suppose that the surgeon has selected an instrument to track (by selecting an insertion point T). The projection P of the insertion point on the image plane is computed.

Step 1: segmentation:

In this step, we compute the edge points in the image by a simple gradient method. The orientation of the edges corresponding to the tool are constrained by the position of P . After this step, we obtain a cloud of edge points, most of which correspond to the instrument.

Step 2: search for the symmetry axis of the instrument:

We use the edge points computed in the previous step to find the symmetry axis of the instrument. To do this, we use a Hough method [14] centered on P . This allows to consider only the lines that have a satisfying orientation, and the computation time is very reasonable compared to a classical Hough method.

Step 3: search for the edges of the instrument:

All the edge points which voted for the symmetry axis were stored in the previous step. The symmetric of the cloud of edge points compared to the symmetry axis is computed to obtain more points, and linear regressions are performed for all the points above the axis and below the axis.

Step 4: search for tip of the instrument:

The tip of the instrument is searched on the symmetry axis. The points of the axis are ordered, according to the position of the insertion point. An Otsu threshold [15] is used to separate the points of the axis in two classes. The pixels of the axis belonging to the same class are grouped into zones. The difficulty of this step is to determine which of the two classes corresponds to the instrument, and to group the zones correctly in order to find the tip. We developed two approach, one for the first image of the tracking, the other for the next images.

For the first image, we search for the longest zone of each class and select the class corresponding to the instrument according to its position with regard to the tip. The average color of the zone corresponding to the tool and the average color of the zone corresponding to the instrument are stored for the next images. In the next images, we use this color information to determine which zone corresponds to the instrument and to the background.

Once the zone corresponding to the instrument has been found, we look if it can be concatenated with other zones corresponding to the tool according to color and length considerations.

3 Results

The method was tested on anatomical pieces. Our first step was to validate our hypothesis that the insertion points are fixed during the intervention. We measured along time the position of the insertion point of an instrument (section 2.2). To check that the first computation of the insertion point was satisfying, we projected the insertion point on the image plane (fig. 4). The distance between the measured points and the reference was inferior to 5mm except for measure $n^{\circ}3$, but this was due to leak in of gas in the abdominal cavity. As soon as the leak was stopped the results were satisfying.

n°	time (min)	X (mm)	Y (mm)	Z (mm)	distance (mm)
1	0	78.9	148.4	-39.3	reference
2	11	79.2	152.3	-42.3	1.53
3	14	91.6	207	-39.8	57.3
4	25	71.3	143.5	-52.3	4.0
5	43	71	148	-52	1.2
6	205	73.4	147	-53	1.3
7	234	77.14	152.2	-47.6	4.6

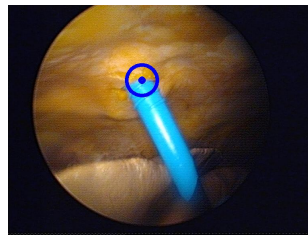


Fig. 4. Variation of the position of the insertion point along time. On the right image, the insertion point was projected on the image plane.

We were able to perform the tracking of an instrument when the specular reflexions were limited. The computation time is around 100 ms for a 200x100 image. It allows a smooth tracking when the instrument is moved at a normal speed. It could still be reduced by optimizing the code allowing to take into

account abrupt and large movements. A resolution of 200x100 pixels is enough for the tracking of a laparoscopic instrument, since we are not looking for subpixel precision to control the camera. We considered that the tool was centered as long as the distance between the detected tip and the previous tip was less than 11 pixels. Fig. 5 shows the distance in pixels between the tip found by the method and the tip manually selected. In 70% of the images, the error is inferior to 5 pixels, and in 87 % the error is less than 11 pixels.

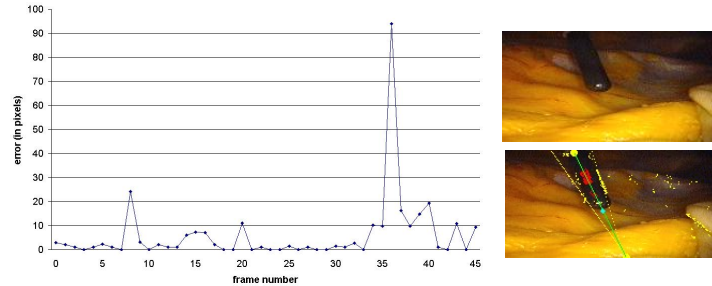


Fig. 5. Error in pixels between the tip found by the method and the tip selected manually and an example of correct detection (yellow dots: candidate points, red points: points corresponding to the axis, blue dot on the axis: tip of the instrument).

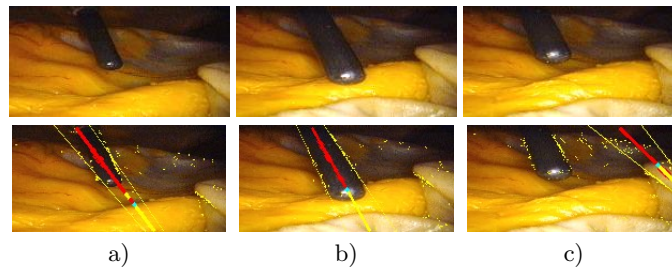


Fig. 6. a) wrong detection due to a lack of contrast b) wrong detection due to specular reflections, c) wrong detection of the axis.

We analysed the images for which the detection was wrong (fig. 6): fig. 6 a) corresponds to the first error peak (frame 9). The detection of the axis is correct, but the tip is wrongly positioned: the red segments correspond to the pixels labelled as pixels belonging to the instrument and the yellow segments as pixels labelled as the background. Since the background close to the tip is very dark, it is wrongly labelled by the Otsu threshold as a region belonging to the instrument. This is also the cause of the error for frames 35, 36, 61, 63 and 64.

Fig 6 b) corresponds to the second error peak (frame 21). Again, the axis is correctly detected, but not the tip. This time, the error is due to specular reflexions, which are labelled as background. This is also causing errors for frames 67 and 69. The important error for frame 37 is caused by a false detection of the axis of the tool (fig. 6 c)): the edges of the instrument are barely visible in the

image because of the lack of contrast, and the organ on the far hand side of the image provides a lot of edges, which deteriorates the axis detection.

When the specular reflections are important, they do not deteriorate the detection of the axis, since they are along the axis of the instrument. However, they cause the failure of tip detection, because the wrong classes are attributed to the instrument and background.

4 Discussion

This experiment allowed us to validate our approach by showing that the insertion point of the instruments are relatively fixed during the intervention. We were able to robustly detect the axis of an instrument, except in cases for which the tool is in a border of the image with a low contrast, or very close to the insertion point. To enhance further this detection, we could give more credit to radial lines than lines with other orientations, since the instruments are usually directed towards the center of the image, as suggested by [10]. The detection of the edges is also quite robust, and we plan to integrate the zoom in the visual servoing of the robot.

The detection of the tip works well on images with little specular reflections. It must be improved to deal with them. We are currently investigating a solution in which the specular reflections along the axis of the instrument would be detected and removed. Since we detect correctly the axis, we could perform this detection only on a region of interest around the axis, with a reasonable computation time.

To further enhance the speed of the detection, we also plan to take into account the position of the tip of the instrument in the previous images. However, considering the very strong magnification of an endoscope, the instrument might very well move a lot in the image between two frames, so the region of interest (ROI) used must be carefully chosen. Moreover, if the ROI is too small, only a tiny portion of the instrument will be visible, and our shape model will not be accurate, but we consider making a more complete model of the tool.

Finally, we must avoid as much as possible false detections. Our method deals partly with them: if the amount of edge points which voted for the axis is too low, or if the length of the instrument is too small an error is raised. If the tool was correctly detected in the first image, its color and the background color are stored for the next images. Thus, if the detection algorithm finds a tool which color is very close to the color of the background, an error is raised. We must now find a method to reject a false detection when the image is the first image of the tracking. One solution could be to gather a few images of the abdominal cavity without an instrument at the beginning of the intervention to obtain information about the color of the background.

5 Conclusion

We developed a novel method to detect the instruments in laparoscopic images. The method is automatic and works almost in real-time. Our first cadaver test

showed encouraging results and allowed us to find the difficult cases still to solve. Several high level control modes based on this method could be implemented to supervise surgical tasks such as the insertion/removal of an instrument or suturing tasks.

Acknowledgement: the authors would like to thank Bruno Thibaut from the anatomy laboratory for his help during these experiments.

References

1. B. Makhoul, A. De La Taille, D. Vordos, L. Salomon, P. Sebe *et al.*, Laparoscopic radical nephrectomy for T1 renal cancer: the gold standard? A comparison of laparoscopic vs open nephrectomy. *BJU International* 2004, Vol. 93, pp. 67-70, 2004.
2. L. R. Kavoussi, R. G. Moore, J. B. Adams *et al.*, Comparison of robotic versus human laparoscopic camera control, *J Urol*, 154:2134, 1995.
3. P. Ballester, Y. Jain, K. R. Haylett and R. F. McCloy, Comparison of task performance of robotic camera holders EndoAssist and Aesop. *International Congress Series*, Vol. 1230, pp. 1100-1103, June 2001.
4. A. Krupa, J. Gangloff, C. Doignon, M. de Mathelin, G. Morel *et al.*, Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual servoing. *IEEE Trans. on Robotics and Automation*, vol. 19(5), pp. 842-853, 2003.
5. G. Wei, K. Arbter, G. Hirzinger, Real-Time Visual Servoing for Laparoscopic Surgery. *Controlling Robot Motion with Color Image Segmentation*. *IEEE Engineering in Medicine and Biology*, pp. 40-45, 1997.
6. X. Zhang, S. Payandeh, Application of Visual Tracking for Robotic-Assisted Laparoscopic Surgery, *Journal of Robotics Systems* (2002) 19(7) : 315-28.
7. O. Tonet, T.U. Ramesh, G. Megali, P. Dario, Image analysis-based approach for localization of endoscopic tools, *Proc. of Surgetica'05*, pp. 221-228, 2005.
8. Yuang Wang, D. R. Uecker and Yulun Wang, A new framework for vision enabled and robotically assisted minimally invasive surgery. *Computerized Medical Imaging and Graphics*, Vol. 22, pp. 429-437, 1998.
9. C. Doignon, F. Nageotte, M. De Mathelin, Detection of grey regions in color images : application to the segmentation of a surgical instrument in robotized laparoscopy, *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp 3394-3399, 2004.
10. J. Climent, P. Marés, Automatic instrument localization in laparoscopic surgery, *Electronic Letters on Computer Vision and Image Analysis*, 4(1), pp. 21-31, 2004
11. S. Voros, E. Orvain, J-A. Long, P. Cinquin, Automatic detection of instruments in laparoscopic images: a first step towards high level command of robotized endoscopic holders, *proceedings of the IEEE / RAS-EMBS International Conference on Biomedical Robotics and Biomechatronics*, 2006.
12. P. J. Berkelman, Ph. Cinquin, J. Troccaz, J-M. Ayoubi, C. Létoublon, Development of a Compact Cable-Driven Laparoscopic Endoscope Manipulator. *MICCAI 2002*, Vol. 2488, pp. 17-24, 2002.
13. Z. Zhang, A Flexible New Technique for Camera Calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(11) pp. 1330-1334, 2000.
14. Richard O. Duda, Peter E. Hart, Use of the Hough Transformation To Detect Lines and Curves in Pictures. *Communications of the ACM*, Vol. 15(1), pp. 11-15, 1972.
15. N. Otsu, A threshold selection method from gray level histograms. *IEEE Trans. Systems, Man and Cybernetics*, Vol. 9, pp. 62-66, 1979.