Video-Based Tracking of Laparoscopic Instruments Using an Orthogonal Webcams System

Fernando Pérez, Humberto Sossa, Rigoberto Martínez, Daniel Lorias, and Arturo Minor

Abstract—This paper presents a system for tracking the movement of laparoscopic instruments which is based on an orthogonal system of webcams and video image processing. The movements are captured with two webcams placed orthogonally inside of the physical trainer. On the image, the instruments were detected by using color markers placed on the distal tip of each instrument. The 3D position of the tip of the instrument within the work space was obtained by linear triangulation method. Preliminary results showed linearity and repeatability in the motion tracking with a resolution of 0.616mm in each axis; the accuracy of the system showed a 3D instrument positioning error of 1.009 ± 0.101mm. This tool is a portable and low-cost alternative to traditional tracking devices and a trustable method for the objective evaluation of the surgeon's surgical skills.

Keywords—Laparoscopic Surgery, Orthogonal Vision, Tracking Instruments, Triangulation.

I. INTRODUCTION

LAPAROSCOPIC surgery is the minimally invasive surgical technique used more frequently in recent years. In this procedure, the long surgical instruments and the endoscope are inserted in the patient's abdominal cavity through small incisions. The surgeon performs the surgery while viewing the image captured by the endoscope on a monitor.

This technique offers many benefits to the patient, e.g., less organ damage, shorter periods of hospitalization and recovery and therefore, an earlier return to the patient's regular activities. However, laparoscopic procedures involve significant technical challenges different to those related to conventional open surgery.

Laparoscopic surgery requires high skill in the use of surgical instruments, adaptation to limited workspace, and adaptation to the loss of information due to the depth. Therefore, the training of the surgeon's skills [1], [2] and the development of new training systems which contribute significantly to the surgeon's improvement of surgical and technical skills are very important. Physical training systems [3], [4] and virtual reality and augmented reality simulators [5], [6] have shown their efficiency for learning and continuous improvement of the surgeon's psychomotor skills.

F. Pérez, D. Lorias, and A. Minor are with the Depto. de Ingeniería Eléctrica, Sección de Bioelectrónica, Cinvestav del IPN, México D.F (e-mail: fperez@cinvestav.mx dlorias@cinvestav.mx aminor@cinvestav.mx).

H. Sossa is with the Centro de Investigación en Computación, Instituto Politécnico Nacional, México D.F (e-mail: hsossa@cic.ipn.mx).

R. Martínez is with the Facultad de Ingeniería, UAEM, Toluca, Edo. de México (e-mail: rigo@ieee.org).

However, these systems, especially those for simulators, use devices based on sensors and/or mechanical joints for the localization of the position of the surgical instruments; these devices are very bulky and restrict the freedom of movement of the instruments.

Currently, there are alternatives to capture and track the movement of laparoscopic instruments with computer vision techniques in training systems. This approach uses the simulator ProMIS [6] in which three cameras capture the movement of the instruments from three different angles in Cartesian coordinates (*x, y, z*) within a dummy. Other training systems as in [7], [8], use the endoscope 2D image of a training box (FLS) and the spatial information of the points of incision, the field of view of the endoscope, and the shape and tool vanishing point in the image to track the movement of the laparoscopic instruments.

In this work, a system for recording the movement of laparoscopic instruments based on an orthogonal system of webcams installed in a physical trainer and video image processing is being presented. This system allows obtaining the 3D movements of the tip of the laparoscopic instruments by triangulating their position in the workspace. This recording device is simple, easy to install, and offers a low-cost and portable alternative for the traditional tracking systems (mechanical, electromechanical or optical sensors) of the movement of laparoscopic instruments.

II. MATERIALS AND METHODS

A. Description of the Tracking System

The physical trainer consists of a semi-cylindrical cavity which simulates the abdominal area of the human body; it has uniform illumination and several ports of entry. The port through which the zero degrees optical is introduced is in the center of the semi-cylinder; the surgical instruments are inserted through two other ports located on the sides. Training models are positioned distally from the optics within the workspace. This setting recreates a working diamond similar to that of regular laparoscopic surgery [3], [4].

Two webcams are installed in the same trainer in order to track the movement of the instruments in the working space. The cameras are placed in orthogonal configuration at a distance of 13.5cm. With this arrangement, the image plane in camera 1 and camera 2 are parallel to the planes inside of the trainer *x–y* and *x–z*, respectively (Fig. 1).

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
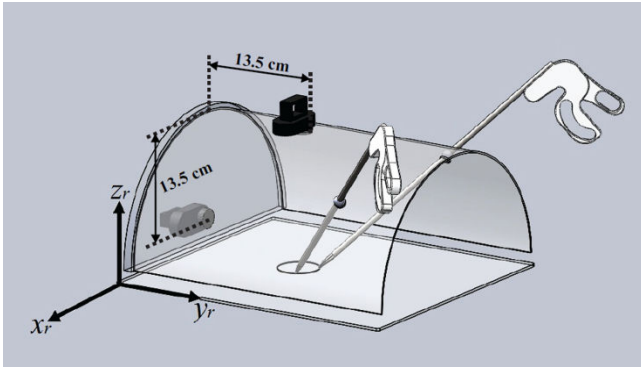Vol:7, No:8, 2013

Fig. 1 Placement of the webcams system with respect to the physical trainer

B. Image Correction

Camera images are affected by a phenomenon known as image distortion. This distortion occurs due to the camera lens and deforms the images in a nonlinear way, which affects this registration system. In order to determine the distortion parameters and correct the alteration, image calibration, by means of a chessboard, was performed. Then, with this reference, the library OpenCV v2.1 and algorithms [9], software in C language was developed to make the correction on the images coming from the two cameras.

C. Detection of Instruments

In order to facilitate the detection of laparoscopic instruments in the image, green and blue markers were placed near the tip of each instrument. The orthogonal webcams system uses a color-segmentation algorithm that isolates instrumental markers within the image. This algorithm was developed in C language and the OpenCV library and is similar to that described by [10].

Once the objects of interest were detected, first-order moments were calculated [11] and each marker centroid was determined in image coordinates (u, v). The noise in the image, which did not belong to the object of interest, was removed using two morphological operations of erosion and dilation [12]. Fig. 2 shows the final detection of the surgical instruments on the images.
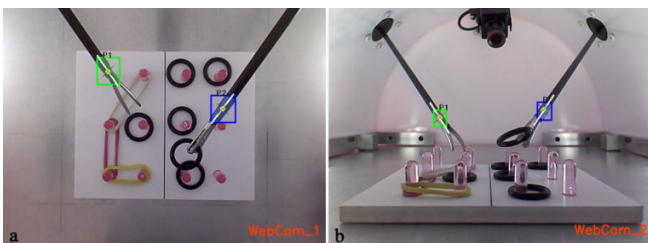

Fig. 2 Detection of surgical instruments with color markers in the image (a) View in the *x-y* plane, and (b) View in the *x-z* plane

D. Determining the Position of the Instruments

In order to find the 3D position of surgical instruments within the trainer workspace, an algorithm based on linear triangulation method was implemented. This method requires the intersection of two lines given the position of an object in two images taken with a camera system [13].

Suppose that $\mathbf{x} = (x, y, z, 1)^T$ is a point in real-world coordinates and $\mathbf{u} = w\,(u, v, 1)^T$ are its coordinates in the image written in homogeneous coordinates. Projective transformation is defined by a matrix P of 3x4 known as *camera matrix* where 11 of the 12 coefficients of this matrix allow the modeling of parameters such as the location and orientation of each camera by the equation:

$$\boldsymbol{u} = P\boldsymbol{x} \tag{1}$$

Rewriting (1) as:

$$wu = \mathbf{p}_1{}^T\mathbf{x}, \quad wv = \mathbf{p}_2{}^T\mathbf{x}, \quad w = \mathbf{p}_3{}^T\mathbf{x} \tag{2}$$

where 1, 2 and 3 is the row of the matrix P. Substituting the w term, it is obtained:

$$\begin{aligned} v\mathbf{p}_3{}^T\mathbf{x} &= \mathbf{p}_2{}^T\mathbf{x} \\ u\mathbf{p}_3{}^T\mathbf{x} &= \mathbf{p}_1{}^T\mathbf{x} \end{aligned} \tag{3}$$

E. Calibration of the Webcams System

For the calibration of the orthogonal camera system, a set of 6 control points in image coordinates (u, v) with its corresponding real coordinates (x, y, z) within the working space of the cameras was taken. These control points are required to calculate the coefficients of the matrix P for each camera using (3). This procedure resolves the ambiguity arising from the orthogonal collocation of the webcams.

A total of 4 linear equations with three unknown factors in real-world coordinates $\mathbf{x}= (x, y, z)$ were obtained. The solution for (3) in real-world coordinates was obtained by the technique of Single Value Decomposition (SVD). As a result of this algorithm, the coordinates of the surgical instruments in the images are needed to know their location in the 3D real-world scenario. These image coordinates are introduced into the equation system and are rewritten in the form $A\mathbf{x} = 0$ in order to obtain the corresponding values in the real-world scenario (Fig. 3).
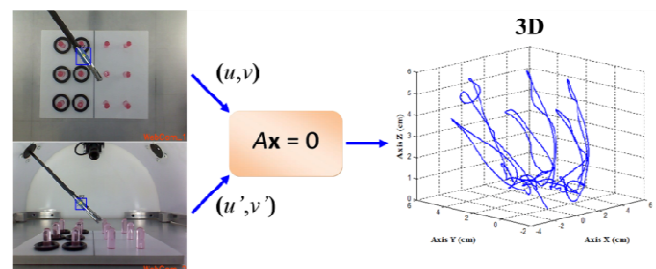

Fig. 3 Example of determination of the 3D position of surgical instrument (cm) by triangulation the position of the colored marker in the images (pixel)

F. Validation Process

A universal milling machine was used to validate the tracking system. This machine has a resolution of 25.4μm on each axis, and was used to accurately control the position of

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:7, No:8, 2013

the tip of the instrument in the directions X, Y and Z. This milling machine also allows making controlled movements at different speeds.

The two webcams were installed in a steel sheet with the same distance and orthogonal disposition as they were placed inside the trainer and the sheet was mounted on the platform of the milling machine. The laparoscopic instrument was simulated by means of a pattern with a green marker attached to the machine's chuck (Fig. 4).
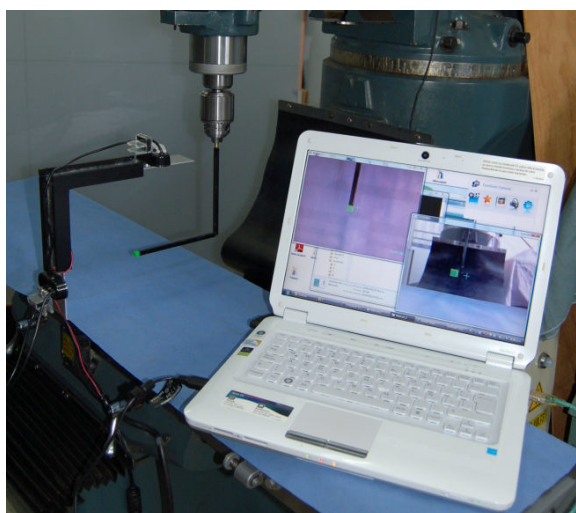

Fig. 4 Placing of the orthogonal webcams system on the milling machine platform for validation process

With this experimental configuration, validation consisted in shifting the tip of the simulated instrument within a working space with the dimensions of an imaginary cube of 50mm per side at a controlled speed of 180mm/min. The movement within this cube was recorded 10 times. These records allowed verifying the linearity and repeatability of the movements as well as the display of random and systematic errors. It also allowed obtaining the resolution on each axis of the system.

The accuracy of the 3D position of instrumental was evaluated by taking 9 points of reference in the 50mm cube. These points were the center and the 8 vertices of the imaginary cube. Reference points were assessed by using the average error expressed in the following equation:

$$Error = \frac{1}{N}\sum_{i=1}^{N}[\,(X_r - X_i)^2 + (Y_r - Y_i)^2 + (Z_r - Z_i)^2]^{1/2} \quad (4)$$

where (X_r, Y_r, Z_r) are the actual coordinates obtained with the milling machine and (X_i, Y_i, Z_i) are the coordinates measured in millimeters by the system.

III. RESULTS

A. Preliminary Results

The data of the simulated instrument movement following the imaginary cube edges were plotted in Matlab R2008a (MathWorks Headquarters, Natick, MA). The resolution of the system was 0.616mm in the three axes of work. Fig. 5 shows the result of the monitoring of the imaginary 50mm cube traced 3 times at a speed of 180mm/min within the workspace using the tracking system.

The figure makes evident the repeatability of the controlled movements it does not present nonlinearities in any of the three working axes (X_i, Y_i, Z_i) of the registration system. The same result was observed in all 10 records taken from the imaginary cube.
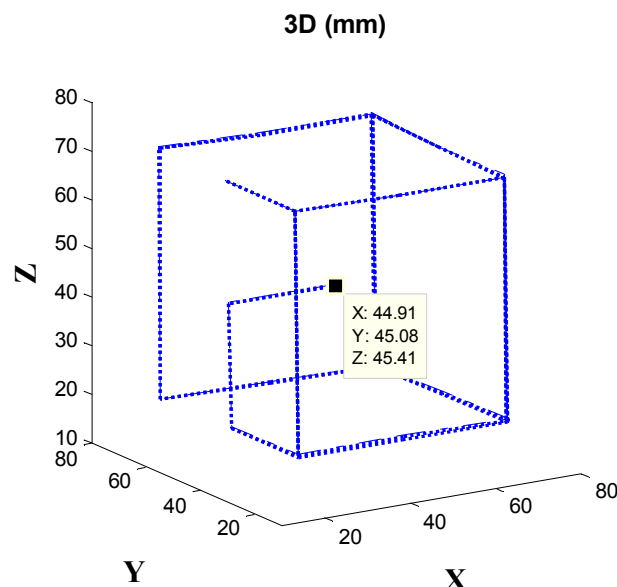

Fig. 5 Traces of the imaginary cube edges using the milling machine and the orthogonal webcams system. In the figure, the cube was traced 3 times at a speed of 180mm/min

The results of the evaluation of the 3D accuracy of the instrument using the 9 reference points of the cube are presented in Table I.

TABLE I
ERROR IN THE ESTIMATION OF REAL-WORLD COORDINATES

Points	Real-world coordinates [a]	Result of estimated coordinates [a]	Error ± S.D.
Center	(45.0, 45.0, 45.0)	(44.907, 45.077, 45.412)	0.429 ± 0.251
Vertex 1	(20.0, 20.0, 20.0)	(20.077, 19.803, 20.406)	0.491 ± 0.140
Vertex 2	(70.0, 20.0, 20.0)	(69.601, 20.330, 19.945)	0.568 ± 0.081
Vertex 3	(70.0, 70.0, 20.0)	(70.186, 69.959, 19.050)	0.974 ± 0.089
Vertex 4	(20.0, 70.0, 20.0)	(20.604, 70.232, 19.710)	0.722 ± 0.056
Vertex 5	(20.0, 70.0, 70.0)	(19.327, 70.842, 71.487)	1.859 ± 0.191
Vertex 6	(70.0, 70.0, 70.0)	(69.335, 70.752, 70.856)	1.337 ± 0.051
Vertex 7	(70.0, 20.0, 70.0)	(68.847, 20.622, 70.151)	1.377 ± 0.033
Vertex 8	(20.0, 20.0, 70.0)	(18.996, 20.295, 70.821)	1.330 ± 0.023
		Average Error	1.009 ± 0.101

[a] The coordinates are expressed in millimeters.

IV. DISCUSSION

This work presents a new system for the tracking of the movement of laparoscopic instruments based on an orthogonal webcam system and video image processing. This recording system allows the tracking of the tip of the instrument and the recording of the movement within the workspace inside the physical trainer. Moreover, this system does not restrict the natural movement of laparoscopic instruments because it does

World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:7, No:8, 2013

not use elements designed with mechanical joints or sensors coupled to the input ports in order to track the surgical maneuvers, and it can be easily installed on other physical simulators without making any changes to their structure.

The preliminary results demonstrate the potential of the tracking system based on an orthogonal webcam system. In the validation test, the system presented a resolution of 0.616 mm in the three-axis of work, linearity and repeatability in the recording of the movements. It also provided good accuracy (average error was 1.009 ± 0.101 mm) in the estimation of the 3D position of the instrument evaluated with the 9 reference points taken within the cube.

The biggest coordinate error estimation came when Z = 70. This could be due to two factors: the image correction was insufficient to reduce the distortion of the cameras, and the control points taken in the workspace for calibrating the orthogonal webcams system were not optimal. Considering this, we will conduct studies to improve the calibration of the image as well as search and analysis of the best set of benchmarks for more accurate results and an estimation of the instrumental position optimal. However, an accuracy of about 3 or 4 mm is enough for the surgical instrument movement analysis and the objective evaluation of the surgeon's psychomotor skills using metrics described in the literature [14].

The tracking system allows the partial occlusion of the laparoscopic instruments without suffering considerable loss of precision of the movement. However, in the future we will work on a solution to total occlusion of the markers of the instruments, performing experiments in most realistic scenarios, which involve the collision of instruments inherent in the manipulation of objects during the training in the simulator.

This tracking system, based on video image processing, provides a low-cost alternative to traditional recording devices, a non-obstructive solution for capturing and analyzing the movement of the laparoscopic instruments without altering the performance of the surgeon, and a portable tool for the learning and objective assessment of the skills acquired by the surgeon in training.

The information of the instrument trajectories obtained with this motion capture tool could lead to a better understanding and modeling of the surgeons' intentional and unintentional movements made during their training in this surgical technique. This data would allow determining the surgeon's needs of improvement and posterior analysis of surgical movements.

V. CONCLUSION

In this article we present an orthogonal webcams system installed in a physical trainer and video image processing for recording the movement of laparoscopic instruments. Preliminary results showed a resolution of 0.616 mm in each axis of work, linearity and repeatability in the motion tracking, as well as the automatic detection of the 3D position of the tip of the surgical instruments with enough accuracy. This tracking system provides a low-cost and portable alternative to

the traditional systems of instrumental movement registration (mechanical, electromechanical or optical sensors). The system aims to help in training, learning and objective assessment of the surgical skills of laparoscopic surgeons.

REFERENCES

[1] E. M. Ritter, D. J. Scott, "Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery,". *Surg Innov.*, vol. 14 pp. 107–112, 2007.

[2] M. C. Vassiliou, G. A. Ghitulescu, L. S. Feldman, D. Stanbridge, K. Lef-fondré, H. H. Sigman, G. M. Fried, "The MISTELS program to measure technical skills in laparoscopic surgery," *Surg Endosc*, vol. 20, pp. 744–747, 2006.

[3] A. M. Martinez, D. L. Espinoza, "Novel laparoscopic home trainer," *Surg Laparosc Endosc Percutan Tech,* vol. 17, pp. 300–302, 2007.

[4] A. M. Martinez, A. C. Kalach, D. L. Espinoza, "Millimetric laparoscopic surgery training on a physical trainer using rats," *Surg Endosc.*, vol. 22, pp. 246–249, 2008.

[5] A. Maciel, Y. Liu, W. Ahn, T. P. Singh, W De S. Dunnican, "Development of the VBLaST: a virtual basic laparoscopic skill trainer," *Int J Med Robot.* vol. 4, pp. 131–8, 2008.

[6] K. R. Van Sickle, D. A. McClusky, A. G. Gallagher, C. D. Smith, "Construct validation of the ProMIS simulator using a novel laparoscopic suturing task," *Surg Endosc.* vol. 19, pp. 1227–1231, 2005.

[7] B. F. Allen, F. Kasper, G. Nataneli, E. Dutson, P. Faloutsos, "Visual tracking of laparoscopic instruments in standard training environments," *Stud Health Technol Inform*, vol. 163, pp. 11–17, 2011.

[8] I. Oropesa, P. Sánchez-González, M. K. Chmarra, P. Lamata, A. Fernández, J. A. Sánchez-Margallo, F. W. Jansen, J. Dankelman, F. M. Sánchez-Margallo, E. J. Gómez, "EVA: Laparoscopic Instrument Tracking Based on Endoscopic Video Analysis for Psychomotor Skills Assessment," *Surg Endosc*, vol. 27, pp. 1029–39, 2013.

[9] Z. Y. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions On Pattern Analysis and Machine Intelligence.* vol. 22 Ed-11, pp. 1330–1334, 2000.

[10] G. Q. Wei, K. Arbter, G. Hirzinger, "Automatic tracking of laparoscopic instruments by color coding," *Proceedings First Int. Joint Conf. CVRMed-MRCAS'97,* Grenoble, France, Mrch 20–22, 1997, Springer Verlag: Lecture Notes in Computer Science, 1997.

[11] R. J. Prokop, A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 438–460, 1992.

[12] R. Van den Boomgard, R. Van Balen, "Methods for Fast Morphological Image Transforms Using Bitmapped Images," *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 254–258, 1992.

[13] R. Hartley, R. Gupta, T. Chang, "Stereo from uncalibrated cameras," in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, pp. 761–764, 1992.

[14] S. Cotin, N. Stylopoulos, M. P. Ottensmeyer, P. Neumann, D. W. Rattner, S. L. Dawson, "Metrics for laparoscopic skills trainers: The weakest link!, " In: T. Dohi and R. Kikinis (Eds.): *MICCAI 2002.* LNCS, vol. 2488, pp. 35–43. Springer, Heidelberg, 2002.