

Controlling virtual scenarios for minimally invasive surgery training using the EVA Tracking System

I. Oropesa^{1,2}, J.A. Sánchez-Margallo³, P. Sánchez-González^{1,2}, J.B. Pagador³, G. Tsolkas⁴,
N. Skarmas⁴, F.M. Sánchez-Margallo³, E.J. Gómez^{1,2}

¹ Grupo de Bioingeniería y Telemedicina, ETSI Telecomunicación, Universidad Politécnica de Madrid, Madrid, Spain; {ioropesa, psanchez, egomez}@gbt.tfo.upm.es

² Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Zaragoza, Spain

³ Centro de Cirugía de Mínima Invasión Jesús Usón, Cáceres, España; {jasanchez, jbpagador, msanchez}@ccmijesususon.com

⁴ Avaca Technologies, Athens, Greece; {gtsolkas, nskarmas}@avaca.gr

Abstract

Software-based solutions such as virtual reality simulators and serious games can be useful assets for training minimally invasive surgery technical skills. However, their high cost and lack of realism/fidelity can sometimes be a drawback for their incorporation in training facilities. In this sense, the hardware interface plays an important role as the physical connection between the learner and the virtual world. This article presents the first efforts in integrating the EVA Tracking System, based on computer vision, to control a virtual serious game with real laparoscopic instruments. Integration between the game engine and the tracking system is achieved using ZeroC Ice and Apache Thrift frameworks. Control of the game is viable in this manner, though there are several functional limitations (grasping, rotation), as well as a need to compromise between robustness and lag. Validation with real users will be the next step towards establishing the acceptance and viability of this new solution.

1. Introduction

Training of minimally invasive surgical (MIS) technical skills has undergone a profound transformation in recent decades. The long-standing mentor-apprentice model is being gradually transformed into structured, objective learning programmes. Hands-on performance in real surgeries is delayed until the resident masters technical skills in patient-free laboratory settings. These skills laboratories provide the means to train technical abilities required of MIS surgeons in box trainers, mannequins or cadavers without the stress and information overload of a real surgery, thus encouraging deliberate practice [1].

A common trend in recent decades has been the introduction of Technology Enhanced Learning (TEL) to support MIS learning. Software-based solutions, mainly represented by virtual reality (VR) simulators, have become consolidated companions to technical skills' learning. These simulators offer the chance to reproduce the technical and nontechnical challenges surgeons will face in the operating room (OR) with a high level of fidelity and reproducibility. Their use is widely accepted as a means to train technical skills in different validation studies in the literature, especially at basic level [2].

More recently, a new variant for software-based learning is being explored in the form of serious games. A serious

game can be defined as an 'interactive computer application, with or without significant hardware component, that has a challenging goal, is fun to play and engaging, incorporates some scoring mechanism, and supplies the user with skills, knowledge or attitudes useful in reality' [3]. Underlying beneath them is the notion that the competitiveness and engaging mechanisms of a game can be applied to induce a learning "stealth mode". This, however, requires that the learning objectives of the game are clearly defined and that the learner never loses perspective of the main goal when playing a serious game, which is to acquire or train an ability [3].

A common component both for VR simulators and serious games comes from the fidelity the system offers compared to a real OR situation [5]. In this sense, the hardware interface plays an important role as the physical connection between the learner and the virtual world [4]. This may be achieved either by affixing magnetic/mechanical/optical sensors to real laparoscopic instruments or, most typically, by using joystick-type constructs. Typical problems with these approaches are their cost (acquisition and maintenance), bulkiness and modification of instruments' ergonomics, which may have an impact on fidelity [2].

Video-based tracking of laparoscopic instruments can potentially solve the above mentioned problems. A common solution to the problem consists in using stereoscopic video-based approaches. This solution has been tested in box trainers, where dual-camera approaches are more feasible [6]. However, rarely do surgeons operate using stereoscopic laparoscopes, and thus research in this field has mostly been devoted to techniques for extracting the 3D location of instruments based on the monoscopic image of the endoscope. In this sense, in previous works we presented and validated the EVA Tracking System, which can track instruments inside a box trainer and perform motion analysis for the assessment of MIS technical skills [7].

The goal of this work is to test the feasibility of employing a video-based tracking system to control a virtual MIS training application. An adaptation of the EVA Tracking System will be integrated with a new

serious game for MIS technical skills' training, the Kheiron Training System (KTS). The game aims to provide a low cost, portable and innovative approach to surgical training to engage the learner while ensuring learning outcomes are accomplished. By integrating EVA, the game aims to incorporate a non-expensive controller scheme with a high level of fidelity, allowing the use of real laparoscopic instruments without significant ergonomic modifications.

2. KTS Serious game

The KTS serious game combines box trainer practice and video gaming to provide residents and medical students with an innovative and engaging approach to psychomotor skills training. Its goal is to encourage deliberate practice of basic tasks and manoeuvres through a virtual environment where the actual elements displayed and the goals to achieve are not MIS related, but rather a part of the plot of the game itself.

The game, called "The Alchemist", is set in the work bench of an apprentice alchemist in his/her process to become a master potion-maker. To reach this goal, the player must complete a series of tasks/challenges such as lighting candles, mixing ingredients, catching mice, etc. Tasks are presented in increasing difficulty, and each is designed to train one/several technical aspects of MIS. As such, explanatory audio-visual supporting material is provided with each one of them to correlate the skills just practiced with their application in real surgical scenarios.

All tasks are performed with two magic wands that stand for the laparoscopic instruments. As such, motion of the wands mimics the movement of two real laparoscopic instruments inside a box trainer, which will be used to control the game.

3. EVA Tracking System

The EVA Tracking System is employed to provide real time information on the position of the laparoscopic instruments and send said coordinates to the game in order to control movements of the wands. The EVA is a computer vision-based tracking system that obtains the instrument tip position based solely on the intrinsic camera parameters and the geometrical properties of the instruments, without the need of external sensors. A first MATLAB® version of EVA was previously validated for motion analysis of MIS skills in box trainer settings [7].

The EVA Tracking System is currently implemented in C++ using OpenCV libraries to handle all image and video processing functionalities. EVA can track two laparoscopic instruments on the screen, provided that they each incorporate two different colour markers near their tip. Colour-based semi-automatic segmentation in HSV space is used to identify each instrument in the video.

For each marker, first and second order moments are calculated. The first order moment defines the centroid of the marker, which is used as a reference tracking point. The second order moment provides the orientation of the marker. This information is used in the next stage, which computes a Hough transform of the image to identify the

instruments' borders. Orientation information is used to filter lines with a higher probability of belonging to either instrument. In order to perform Hough analysis, edge information of the image is previously extracted by means of a Canny filter.

The reference 2D tracking point and the information from the borders of an instrument can be used to obtain a 3D estimation of its spatial location. Two previously validated methods can be used indistinctly. The first uses information of the apparent diameter of the instrument to calculate the depth of the point being tracked [7]. The second uses information from the geometrical projection of the instrument (a cylinder in real life, a trapezoid in the video frame) to determine its orientation and position [8]. A final post-processing phase allows to remove noisy measurements and smooth the output motion signal. This is achieved by means of a Kalman filter for each of the laparoscopic instruments used within the game.

4. KTS - EVA communication

Communication between the KTS game and the EVA Tracking System is based on two elements: (1) A hardware platform; and (2) an API that sends motion data from the tracker to the game engine.

4.1. Hardware input requirements

The KTS serious game is designed to be played on any box trainer setting. To that end, the game comes with a hardware input device consisting of a white empty box, a LED ring that illuminates the interior of the box and a camera affixed to the box to show its working space.

The KTS box is designed to be placed inside a box trainer during gameplay. The essential purpose of the box is to provide homogeneous conditions for tracking regardless of the space, background and illumination conditions of the box trainer, effectively providing an environment for chroma keying of the instruments. The front and top of the KTS box are completely open, in order to ensure freedom of movement of the instruments.

Any laparoscopic instrument can be tracked, as long as its shaft presents a black matte colour. As mentioned previously, the only modification required is the placement of a colour marker in their distal end, 2cm long. A calibration software allows beforehand to select and adjust the segmentation seed and thresholds for each marker (Fig. 1).

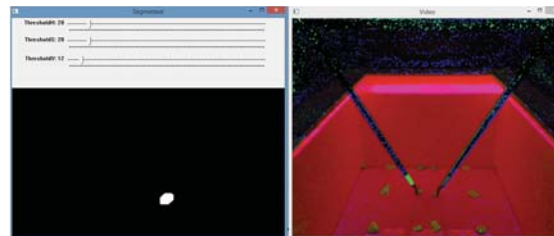


Figure 1. Calibration software. (Left) Segmented view and slide bars. (Right) HSV view inside the KTS box.

The software requires that the user places the instruments inside the KTS box and select a seed for each marker using the mouse. Around a neighbourhood of the seed, the

mean HSV values are computed. Three slide bars allow to control the tolerance/variance around the mean, to ensure a correct segmentation of the marker. The mean and the variance are then saved as the thresholds used for colour-based segmentation in the main tracking application. In this way, the possibility is opened for each user to try different combinations of marker colours.

4.2. Software API communication

Since EVA is developed in C++ and the game engine in C# and Unity 3D, an integration between the two platforms has to be performed on a low-level. ZeroC Ice and Apache Thrift frameworks are used. Although usually one of the above is enough, both of them are used to account for version differences of C# and C++ employed and incompatibilities between them.

The ZeroC/Thrift-based bridge acts as a mediator receiving calls from EVA and translating them to calls to the C# Unity engine. A message oriented approach is adopted to send motion data from the tracking system to the game (Fig. 2).

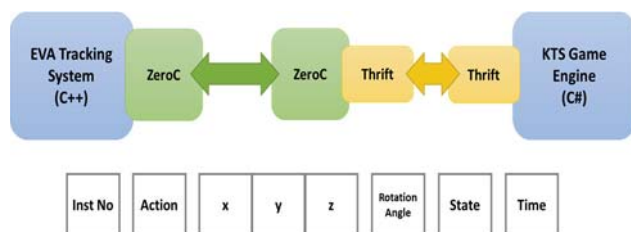


Figure 2. (Top) Communication architecture. (Bottom) Message configuration. Left to right: Instrument (left/right), Action (movement/grasping), Position (x,y,z), Rotation of instrument along its axis ($^{\circ}$, currently not used), State (open/closed), Time (s)

Movements are communicated to the game with respect to the optical centre of the tracking system and are resampled and transformed on a real time basis to X, Y, Z Cartesian coordinates (in cm) with reference point the lower left hand of the KTS box closer to the player. The transformation takes into assumption that the camera and the box edge coordinates are constant.

Currently, the EVA tracking system only supports motion information. In order to handle rotation and grasping actions within the game, software-based solutions have been implemented. Rotation of the in-game wands are achieved by placing them over two specially designated areas on the alchemist's workbench. On the other hand, grasping is achieved by keeping the wand still for two seconds over a virtual object. When this happens, the tracking system sends a flag on the *State* field to the game engine to open/ close the virtual tips of the wand. Objects are released when placed on a surface or when a collision with another object occurs.

5. Results and discussion

After integration of the different modules, the KTS game can be controlled using real laparoscopic instruments inside a box trainer. Figure 3 shows an image of a testing setting, installed inside the SIMULAP IC-05 box trainer, and running on a PC Intel® Core i7, 8GB DDR3 memory,

2.2GHz and NVIDIA® GeForce® 710M with 2Gb dedicated VRAM.

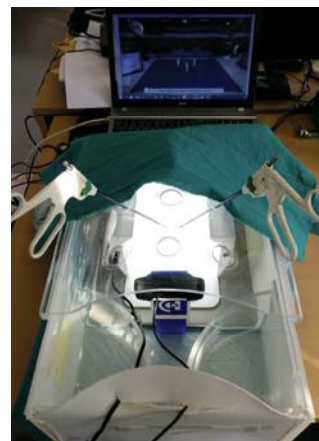


Figure 3. Testing setting. The KTS box is placed inside the SIMULAP IC-05 box trainer and covered with a surgical drape. The camera is used to track laparoscopic instruments, and occupies a fixed position

Figure 4 shows a capture of the game running parallel with an augmented view of the tracking system. Movements of the laparoscopic instruments inside the KTS box are tracked at ~ 10 fps and sent to the KTS game engine, which in turn has a refresh rate of ~ 15 fps. When the laparoscopic instruments are kept still for two seconds, the "Grasp" signal is sent and interpreted by the game. Finally, rotation of instruments works as expected when placed in the designated points inside the game.

Test runs with technical experts and surgeons show that in general, the tracking system performs best capturing long range motion (movements across the box) than precision movements around a target. This is reflected inside the game by a jerkiness in the movements of the instruments that may affect precise manoeuvres, such as lighting a candle or grasping a flask. This jerkiness is mainly caused by small variations in the detection of instrument borders in EVA, which in turn affect the depth estimation. Modification of Kalman parameters can increase the precision of the system, at the cost of introducing a slight lag to the controller. This delay in turn may affect the user experience and impact on the fidelity of the game. A compromise must thus be reached between speed and robustness when controlling the game.

Part of these limitations can also be explained away by the current hardware configuration of the game. Video-based tracking is usually sensitive to lighting conditions. In this sense, EVA could be negatively influenced by light variances inside the KTS box. The SIMULAP box trainer offers a transparent casing, which needs to be covered by surgical drapes to create the necessary lighting conditions, but even then, small variances in illumination may occur. Moreover, the glossy surface of the box can create reflections under the glow of the LEDs, leading to tracking errors. Although the KTS box was specifically designed to facilitate working conditions of EVA together with the KTS serious game, illumination could be further analysed to determine what changes should be introduced.

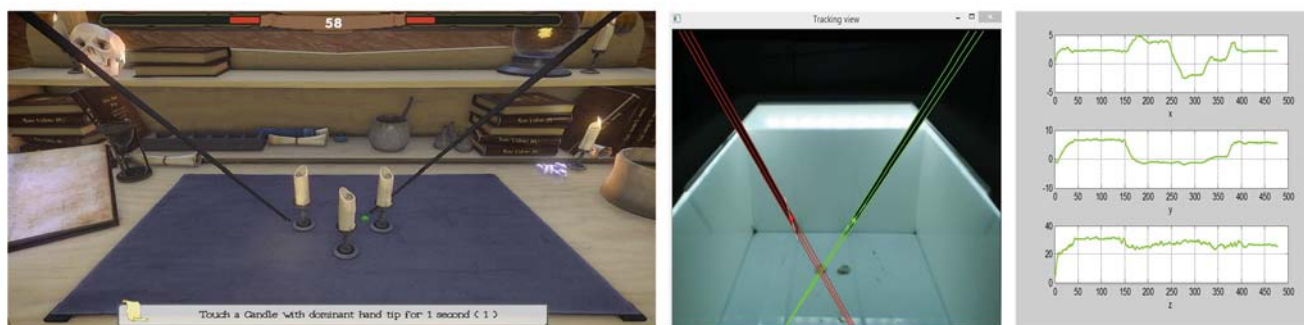


Figura 4. KTS game and EVA Tracking system. (Left) In-game capture of the KTS Candles (Dominant) Task. Goal of the task is to light the three candles shown hovering the right hand wand over them. (Middle) Tracking view inside the box trainer, corresponding with the in-game situation (Right) Final path covered with the dominant instrument performing the task (retrieved with Matlab®).

Additionally, new alternatives for detection of grasping and rotation are required to augment the fidelity of the game. The current solutions are thought of as provisional, and given the aforementioned precision issues, need to be re-evaluated. An obvious alternative comes from the use of sensing technology to detect the instrument state or the rotation. This, however, would again imply modifying the instruments and adding extra hardware. Instead, new alternatives are being considered based on computer vision combined with machine learning to, for example, detect if an instrument is open or closed [9].

A last limitation is the lack of haptic feedback for the trainee, since instruments are manipulated inside an empty box. However, it has been shown that for tasks where force application is necessary, artificially-generated feedback does not necessarily improve technical skills, but natural force feedback does. Thus, a possible improvement could come from exploring hybrid solutions in which the game requires manipulation of physical objects with a software counterpart inside the game.

6. Conclusions

We have presented the first efforts to include a computer vision-based tracking system, the EVA, in a virtual training scenario. This configuration is meant to increase the fidelity of virtual environments while decreasing their costs. Improvements still need to be carried out to fully ascertain the viability of the solution, but this first proof of concept reflects its potential.

The game is currently entering the last stages of development. Validation will take place on four European sites (Germany, Hungary, Romania and Spain) with end-users (residents and medical students). Trials will start in October 2015. Results will not only provide feedback on the pedagogical value of the game, but on the fidelity that a video-based tracking system such as EVA can provide.

Acknowledgements

This work has been carried out under project KTS (543202-LLP-1-2013-1-ES-KA3-KA3MP). This project has been funded with support from the European Commission. This communication reflects the views only of the author, and the Commission cannot be held

responsible for any use which may be made of the information contained therein.

References

- [1] Hiemstra E. Acquiring Minimally Invasive Surgical Skills. Leiden: Department of Minimally Invasive Surgery in Gynaecology, Faculty of Medicine/Leiden University Medical Center (LUMC), Leiden University, 2012
- [2] Oropesa I, Sanchez-Gonzalez P, Lamata P, Chmarra MK, Pagador JB, Sanchez-Margallo JA, Sanchez-Margallo FM Gomez EJ. Methods and Tools for Objective Assessment of Psychomotor Skills in Laparoscopic Surgery. *Journal of Surgical Research* 171(1): E81-E95, 2011.
- [3] Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. *British Journal of Surgery*, 99(10), 1322–30, 2012.
- [4] Våpenstad C, Buzink SN. Procedural virtual reality simulation in minimally invasive surgery. *Surgical Endoscopy*, 27(2), 364–77, 2013
- [5] Lamata P, Gomez EJ, Bello F, Kneebone RL, Aggarwal R, Lamata F. Conceptual framework for laparoscopic VR simulators. *Computer Graphics and Applications, IEEE* 26(6): 69-79, 2006
- [6] Pérez Escamirosa F, Ordorica Flores RM, Oropesa I, Zalles Vidal CR, Minor Martínez A. Face, content, and construct validity of the EndoViS training system for objective assessment of psychomotor skills of laparoscopic surgeons. *Surgical Endoscopy*; doi: 10.1007/s00464-014-4032-6, 2014.
- [7] Oropesa I, Sánchez-González P, Chmarra MK, Lamata P, Fernández A, Sánchez-Margallo JA, Jansen FW, Dankelman J, Sánchez-Margallo FM, Gómez EJ. EVA: Laparoscopic instrument tracking based on endoscopic video analysis for psychomotor skills assessment, *Surgical Endoscopy*; 27(3): 1029-1039, 2013.
- [8] Cano AM, Sánchez-González P, Sánchez-Margallo F, Oropesa I, Pozo F Gómez E. Video-endoscopic image analysis for 3D reconstruction of the surgical scene. *4th European Conference of the International Federation for Medical and Biological Engineering, IFMBE Proc.*22(8): 923-926.
- [9] Sánchez-Margallo JA, Sánchez-Margallo FM, Pagador JB, Gómez EJ, Sánchez-Gonzalez P, Uson J and Moreno J Video-based assistance system for training in minimally invasive surgery. *Minimally Invasive Therapy & Allied Technologies*, 20(4): 197-205, 2011.