EVA: Laparoscopic Instrument Tracking Based on Endoscopic Video Analysis for Psychomotor Skills Assessment

Ignacio Oropesa · Patricia Sánchez-González · Magdalena K. Chmarra · Pablo Lamata · Álvaro Fernández · Juan A. Sánchez-Margallo · Frank Willem Jansen · Jenny Dankelman · Francisco M. Sánchez-Margallo · Enrique J. Gómez

Abstract

Introduction The EVA (Endoscopic Video Analysis) tracking system is a new system for extracting motions of laparoscopic instruments based on nonobtrusive video tracking. The feasibility of using EVA in laparoscopic settings has been tested in a box trainer setup.

Methods EVA makes use of an algorithm that employs information of the laparoscopic instrument's shaft edges in the image, the instrument's insertion point, and the camera's optical center to track the three-dimensional position of the instrument tip. A validation study of EVA comprised a comparison of the measurements achieved with EVA and the TrEndo tracking system. To this end, 42 participants (16 novices, 22 residents, and 4 experts) were asked to perform a peg transfer task in a box trainer. Ten motion-based metrics were used to assess their performance.

Results Construct validation of the EVA has been obtained for seven motion-based metrics. Concurrent validation revealed that there is a strong correlation between the results obtained by EVA and the TrEndo for metrics, such as path length ($\rho = 0.97$), average speed ($\rho = 0.94$), or economy of volume ($\rho = 0.85$), proving the viability of EVA.

Conclusions EVA has been successfully validated in a box trainer setup, showing the potential of endoscopic video analysis to assess laparoscopic psychomotor skills. The results encourage further implementation of video tracking in training setups and image-guided surgery.

Keywords Minimally invasive surgery · EVA · TrEndo · Training · Motion analysis · Video

I. Oropesa (✉) · P. Sánchez-González · P. Lamata · Á. Fernández · E. J. Gómez
Bioengineering and Telemedicine Centre (GBT), ETSI Telecomunicación, Universidad Politécnica de Madrid (UPM), Avda Complutense, 30, 28040 Madrid, Spain
e-mail: ioropesa@gbt.tfo.upm.es

I. Oropesa · P. Sánchez-González · Á. Fernández · E. J. Gómez
Networking Research Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Poeta Mariano Esquillor s/n. 50018, Saragossa, Spain

M. K. Chmarra · F. W. Jansen · J. Dankelman
Department of BioMechanical Engineering, Faculty of Mechanical, Maritime and Materials Engineering (3mE), Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

M. K. Chmarra
Department of Circulation and Medical Imaging, Faculty of Medicine, Norwegian University of Science and Technology (NTNU), Trondheim, Norway

J. A. Sánchez-Margallo · F. M. Sánchez-Margallo
Jesús Usón Minimally Invasive Surgery Centre, Carretera N-521, km 41.8, 10071 Cáceres, Spain

F. W. Jansen
Department of Gynecology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Minimally invasive surgery (MIS, e.g., laparoscopy) techniques have been the focus of much attention in recent years [1–5]. The main motivation for MIS is to provide safer surgeries for the patient's benefit. This same incentive applies to the training and assessment of new professionals and thus needs to address society's demands for fully prepared and accredited surgeons [2]. Halsted-based training is gradually being replaced by structured training programs to optimize the learning curves, time, and efficiency of residents [3]. New learning models state that the direct involvement of residents to real surgeries should be postponed until they have acquired the necessary cognitive

and psychomotor skills [4, 6]. In this context, the first stages of basic psychomotor skills training take place in controlled laboratory settings by means of box trainers and virtual-reality simulators [2, 5, 7].

Part of the training and assessment processes in laboratory settings relies on the acquisition and analysis of efficiency metrics, which are based on the kinetics and dynamics of the trainees' actions performing a task [7, 8]. To this end, active tracking-based systems are used to measure accurately hand/instrument movements and interactions, based on optical [9], mechanical [10], ultrasonic [11], or electromagnetic [12] sensing. However, their use may modify the ergonomics and constrain movements of the surgical instruments, and thus alter the trainees' experience and performance. Moreover, transfer of these technologies into the operating room (OR) is compromised, because they often are bulky (impeding their portability and OR setting), are not easily sterilized, may require a clear line of sight (optical devices), or might be affected by ferromagnetic materials (electromagnetic devices) and hence cause errors in position tracking [13].

An alternative to active tracking is the analysis of the endoscopic images to determine the spatial position of the surgical instruments [14]. The noninvasive and unobtrusive nature of a tracking system based on computer vision techniques makes it suitable for training purposes (both in laboratory settings and the operating room) and for image-guided surgical navigation [14, 15]. An additional potential benefit of this approach is the possibility of an off-line analysis of recorded surgical procedures. In training, the concept has already been exploited in the CAE proMIS simulator (CAE Healthcare, Montreal, Canada), where the instruments position is triangulated by means of two cameras [16–18]. The main challenge in this research field, however, is tracking instruments' motion based on the *monoscopic* image of the endoscope, where depth information is not available.

In general, there are two main technological challenges that need to be addressed to achieve robust video-based tracking in MIS. The first challenge comprises the automatic identification and segmentation of the instrument in the image to perform a two-dimensional (2D) tracking on the video image. In the literature, segmentation has been obtained based on color [19–21], marker [22, 23], edge [24, 25], and/or geometric features [26]. The second challenge requires calculating the three-dimensional (3D) coordinates based on the geometrical features of the segmented marker [22, 27], instrument [28, 29], or from an estimation of the insertion point [30]. In previous works, we presented three different approaches to obtain the 3D coordinates: 1) based on the vanishing point of the instrument's borders [28]; 2) based on the instrument's apparent diameter [28]; and 3) based on the transversal section of the instrument [29].

In general, incorporating instrument tracking to the assessment of psychomotor MIS skills requires defining a series of evaluation metrics to allow discernment of the trainees' expertise based on their performance [31]. In this particular case, information provided by the video will be associated with motion parameters extracted from the instruments' movements. Applications in training and assessment of endoscopic video tracking have already been reported for physical simulators [15, 20]. Sánchez-Margallo et al. [15] employed pattern recognition techniques to track the tip of different laparoscopic instruments in the 2D image. Allen et al. [20] made use of the vanishing point approach and adapted it in the Fundamentals of Laparoscopic Surgery (FLS) simulator. In both cases, a complete clinical validation of the tracking systems is pending.

The purpose of this work is to introduce EVA—a software tracking system based on endoscopic video analysis for the assessment of laparoscopic skills. Construct and concurrent validation results are presented. A secondary objective is to test the hypothetical gain of 3D metrics compared with 2D evaluation metrics. The TrEndo tracking system was used for concurrent validation [9]. This optical sensor-based tracking system has been validated in a number of studies [32–35] and is currently employed for gynecologic resident training at Leiden University Medical Centre (LUMC, Leiden, The Netherlands).

Materials and Methods

System Description

The EVA tracking system employs the endoscopic image and the location of the insertion points of the camera and instruments to calculate the 3D position and orientation of the surgical tools. The edges and tip of each instrument are automatically extracted from each 2D image in the process. Detection is done offline (i.e., using recorded videos of the performances).

Let P be the tip of the surgical instrument, T the insertion point of the instrument, W_{INST} the instruments' diameter, and C and f the endoscope's optical center and focal length, respectively. The implemented solution includes the following steps:

1. Distortion correction: Camera systems tend to suffer from a phenomenon known as radial or barrel distortion, by which images tend to be deformed nonlinearly radially from the optical center [36]. Because the tracking system relies on the detection of the instrument's straight edges, a correction algorithm based on Brown's distortion model [37] is applied to ensure recognition [38] (Fig. 1A). This same algorithm provides the camera parameters C and f.
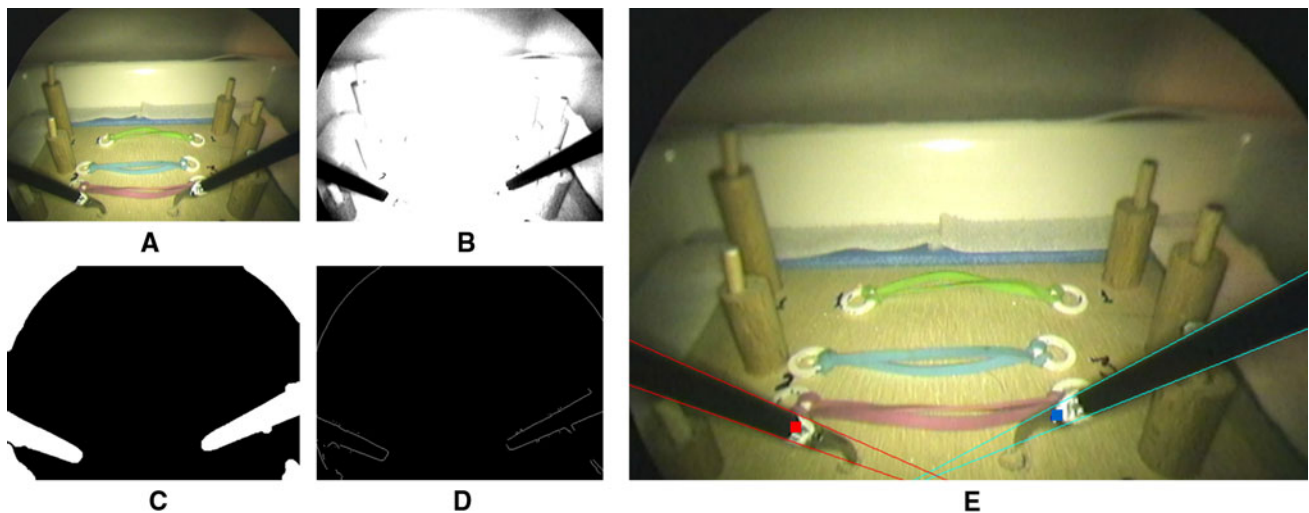
Fig. 1 2D tracking. **A** Original image, distortion corrected, **B** Contrast-enhanced image, **C** Segmentation mask, **D** Sobel filtered edge image, **E** Final detection of instruments and P_2. Lines over the instruments edges represent their detection in Hough space. P_2 is detected in the area of the instruments white marker

2. Image processing: The instrument's edges are detected on the image, by means of image contrast enhancement (Fig. 1B), color-based segmentation (Fig. 1C) to isolate the instruments and Sobel filtering (Fig. 1D) [36, 39]. Hough transform is then used to isolate and characterize its borders [39]. Up to two laparoscopic instruments may be detected this way, considering their insertion angle into the box trainer [20].

3. Determination of a tracking point in the image space (P_2): We define P_2 as the 2D point in the image at the shaft-metal interface of the instrument's middle line. A directional gradient employing the instrument's orientation is used for detection of P_2, as described by Allen et al. [20]. To guarantee a more robust detection, a small white marker is placed on the interface. Figure 1E shows an example of P_2 detection in the box trainer.

4. Determination of tracking window: to ease the computational load, tracking in successive frames is performed in an adaptive window, the margins of which are determined as a neighborhood dependent of: 1) the speed and direction of the instrument; and 2) the apparent width of the tool in the image at the shaft-metal interface (W_{IMG}). If P_2 cannot be determined within this window, tracking is performed once more in the whole image.

5. Determination of point P in real space: We define P_3 as an auxiliary point, the 3D correspondence in space of P_2. By means of the apparent diameter method described in [28], P_3 is obtained using C, f and W_{INST} by triangle similarity (Fig. 2 left). Finally, knowing P_3 and T, P (instrument tip in real space measured with respect to C) is obtained by point translation along the instrument axis (Fig. 2 right).
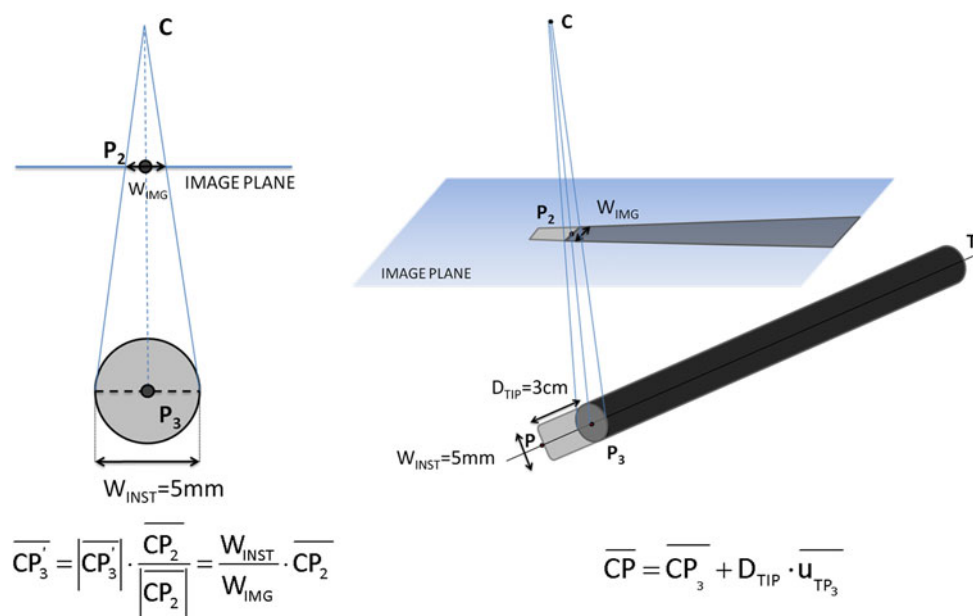
6. Data post processing: to remove noisy data caused by incorrect line and/or incorrect detections of P, we implement a filter based on the k-medians algorithm [40]. In this way, point P at a determined instant (P_t) is reevaluated as the median of a cluster formed with its immediate 30 successors ($[P_t, P_{t+30}]$). This results in an increased robust detection in detriment of a negligible displacement of the trajectory in real space. Additonally, outlier points outside the ranges of the setting are discarded.

Step 2 of the algorithm is of great importance in the tracking process, because it copes with the detection of the whole instrument from the image and is thus setting dependent. For this reason, a calibration of processing parameters must be performed according to the setting's contour conditions: image quality, illumination, background colors, etc. In this study, the system's image processing parameters were empirically determined to enable detection on one of the physical simulators used for learning at the skills training laboratory of LUMC. The determination of the parameters has been done based on the validation setting's characteristics and illumination conditions.

Experimental Validation

All validation trials were carried out at the skills training laboratory at LUMC. Forty-two participants were asked to participate in this study: 16 novices (N), 22 residents (R), and 4 experts (E). All participants performed a one-instrument peg grasping task, where a series of spherical objects (chickpeas) had to be picked, transported, and placed on corresponding sequential holes using a laparoscopic

Fig. 2 3D tracking. *Left*: determination of P_3 (shaft–metal interface). *Right*: determination of P (instrument tip)

$$\overline{CP_3'} = \left| \overline{CP_3'} \right| \cdot \frac{\overline{CP_2}}{\left| \overline{CP_2} \right|} = \frac{W_{INST}}{W_{IMG}} \cdot \overline{CP_2}$$

$$\overline{CP} = \overline{CP_3} + D_{TIP} \cdot \overline{u_{TP_3}}$$

grasper. The scenario included different blocks providing different heights, orientations and occlusion zones (Fig. 3 left).

All participants had no prior training with the task proposed, and only a brief explanation of its goals was given. To ensure the same conditions for all participants, the position of the task within the box trainer, and the insertion points for the camera and instrument were identical for each participant. The endoscope (ACMI CIRCON Microdigital II MV-9695) provided a 0° view on a monitor, and was fixed at the same angle for all participants. Video recordings of performance were stored for posterior offline processing. Simultaneously, the instrument's movements were tracked by means of the TrEndo tracking system. The starting and ending position of the tip of the instrument and the order of target's placement at predefined positions (indicated by the numbers located next to them) were fixed and were the same for all the participants. No time restrictions were imposed, nor completion forced-upon. The test setup is represented in Fig. 3 (right).

All motion metrics employed in the study were derived from the position $r(t)$ of the instrument on every frame. Metrics were computed using Matlab Release 2009b (Mathworks, Natick, MA). Two sets of metrics were
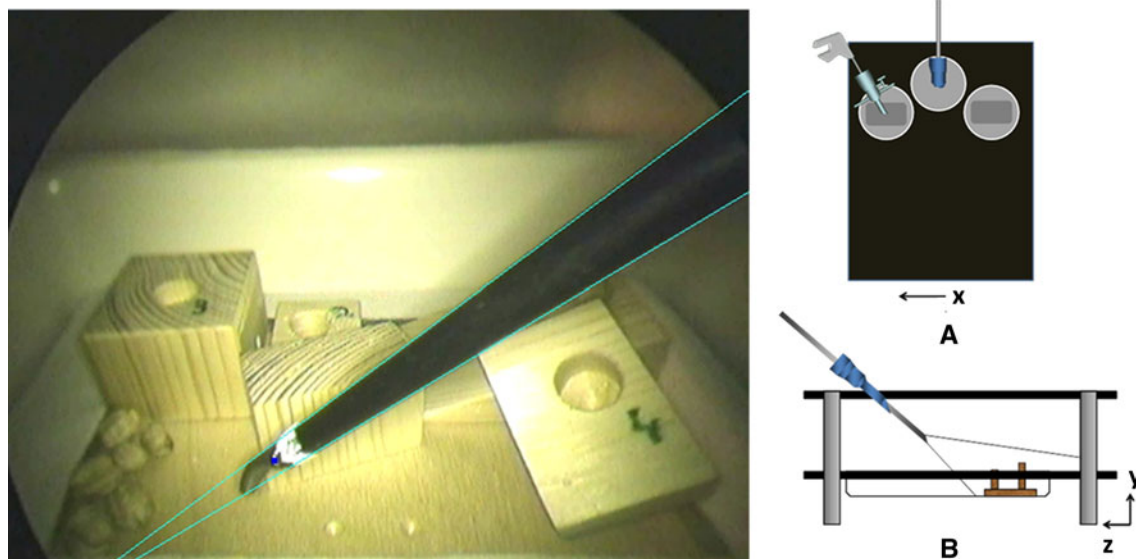

Fig. 3 Experimental setting. *Left*: assessment task employed in validation. *Right*: box trainer schematic disposition and coordinate system: **A** *top view*, **B** *side view*

obtained: one from the 2D image space ($r(t) = [x(t), y(t)]_{t=0}^{T}$), and one from the 3D real space ($r(t) = [x(t), y(t), z(t)]_{t=0}^{T}$). The purpose of this distinction was to determine whether validation may be achieved from studying solely the 2D movements on the monitor. Table 1 sums up the metrics explored, as well as the mathematical formulations that define them.

Statistical Analysis

Significant differences were considered at $p < 0.05$ both for construct and concurrent validation. SPSS v. 17 (SPSS Inc., Chicago, IL) was used for all statistical analysis. Construct validation was performed along the three groups (Kruskal–Wallis analysis) and for each pair of groups (Mann-Whitney U test). Additionally, the variations in p value (the probability given by the statistical tests, ΔP) were determined, where a decrement in the p value from the 2D case to the 3D case reflects an increment of the significant differences

between metrics. Additionally, Pearson's nonparametric correlation (ρ) was used to measure differences between metrics featured both in the 2D image and 3D space.

Concurrent validation was performed for 3D motion metrics between EVA and TrEndo. Their coordinate systems were previously aligned as shown in Fig. 3 for accurate correspondence of measurements. Pearson's correlation was once more employed to measure concurrence. Values between 0.4–0.7 were considered as medium correlating values, whereas values greater than 0.7 show a strong correlation between metrics [41].

Results

All participants were able to complete the task. Three of 42 participants were left-handed, and thus performed the task with the opposite instrument. Typical results obtained by a novice and an expert are presented in Fig. 4, both for 2D and 3D tracking.

Table 1 Motion metrics. Application reflects whether a metric was measured for the 2D image or the 3D real setting. Units for 2D metrics are given in pixels (pp); 3D metrics are given in metrical units

Metrics	Application	Definition	Formulae
Time (T)	2D/3D (s)	Total time to perform a task	T
Path length (PL)	2D(pp)	Total path covered by the instrument in image space	$\int_{t=0}^{T} \frac{d|r(t)|}{dt} dt$
	3D(m)	Total path covered by the instrument in the setting	
Depth (D)	3D(m)	Total path length traveled in the instrument's axis direction	$\int_{t=0}^{T} \sqrt{\left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt$
Average speed (S)	2D(pp/s)	Rate of change of the instrument's position in the image	$\frac{1}{T}\int_{t=0}^{T} \frac{d|r(t)|}{dt}$
	3D (mm/s)	Rate of change of the instrument's position in the setting	
Average Acceleration (A)	3D (mm/s²)	Rate of change of the instrument's velocity within the setting	$\frac{1}{T}\int_{t=0}^{T} \frac{d^2|r(t)|}{dt^2}$
Motion Smoothness (MS)	3D (m/s³)	Abrupt changes in acceleration resulting in jerky movements of the instrument	$\sqrt{\frac{T^5}{2 \cdot PL^2}\int_{t=0}^{T} \left(\frac{d^3|r(t)|}{dt^3}\right)^2}$
Economy of area (EOA)	2D(-)	Relationship between the maximum image area occupied by the instrument and the total path length	$\frac{\sqrt{[(\text{Max}_t(x) - \text{Min}_t(x)) \cdot (\text{Max}_t(y) - \text{Min}_t(y))]}}{PL}$
	3D(-)	Relationship between the maximum surface area (task plane) occupied by the instrument and the total path length	
Economy of volume (EOV)	3D(-)	Relationship between the maximum volume occupied by the instrument in the setting and the total path length	$\frac{\sqrt[3]{[(\text{Max}_t(x) - \text{Min}_t(x)) \cdot (\text{Max}_t(y) - \text{Min}_t(y)) \cdot (\text{Max}_t(z) - \text{Min}_t(z))]}}{PL}$
Search time (ST)	2D (%)	Percentage of time spent in the "search zone" (chickpeas pick-up zone)	$\frac{1}{T}\sum_{t=1}^{T} I_A(x(t), y(t)) / A = \{(x,y) \in R^2 \mid x < 1/3 * \text{Max}_t(x(t)) \wedge y > 1/2 * \text{Max}_t(y(t))\}$
Idle time (IT)	3D (%)	Percentage of time where the instrument is considered to be still	$\frac{|\Im|}{T} : \Im = \left\{ t \in (0, \ldots T) \mid \sqrt{\left(\frac{dx(t)}{dt}\right)^2 + \left(\frac{dy(t)}{dt}\right)^2 + \left(\frac{dz(t)}{dt}\right)^2} \leq 5 \right\}$

Construct validation results are shown in Fig. 5 and Table 2. For 2D metrics, average speed and economy of area present significant differences along the three groups. Furthermore, Mann-Whitney analysis reveals that differences occur mostly between novices and both other groups (R, E). Search time did not reach the significance levels required.

For 3D metrics, construct validation was obtained for path length, depth, average speed, average acceleration, economy of area, and economy of volume. Paired comparisons once more showed that the greatest differences occur between novices and the other two groups. Comparison between 2D and 3D metrics show a strong correlation between them ($\rho > 0.85$) but also reflect an increment in the statistical significance between groups, which is more prominent in the case of path length and economy of area (Table 3).

For concurrent validation, Fig. 6 and Table 2 reflect the correlation between metrics. It can be seen that a strong correlation exists for all considered metrics, except for motion smoothness. Overall, results show that performance of EVA reproduces the measurements obtained with the TrEndo tracking system.

Discussion

We have demonstrated the feasibility of employing solely video based tracking of laparoscopic instruments for the assessment of psychomotor skills in a training setup. The EVA tracking system was validated using a peg transfer task and ten motion-related metrics. The results showed a strong correlation with the measurements obtained with the TrEndo tracking system.

An important result of this study is the high level of similarity between measurements obtained by EVA and the TrEndo tracking system. Correlation values between metrics reflect the validness and potential of video-based tracking. Only motion smoothness metric presented significant differences between systems. A possible reason for this may reside in the post-processing stage of EVA, which effectively applies low pass filtering of the signal in ways that may dampen the influence of jerkiness in movements. It is important to mention that for both systems motion smoothness did not achieve construct validation and that this particular metric has been already been proved not valid for several basic tasks [42, 43].

Fig. 4 Tip motion representation examples. *Top*: results obtained by a novice (2D and 3D). *Bottom*: results obtained by an expert (2D and 3D). 2D motion is represented in pixels, 3D motion in cm
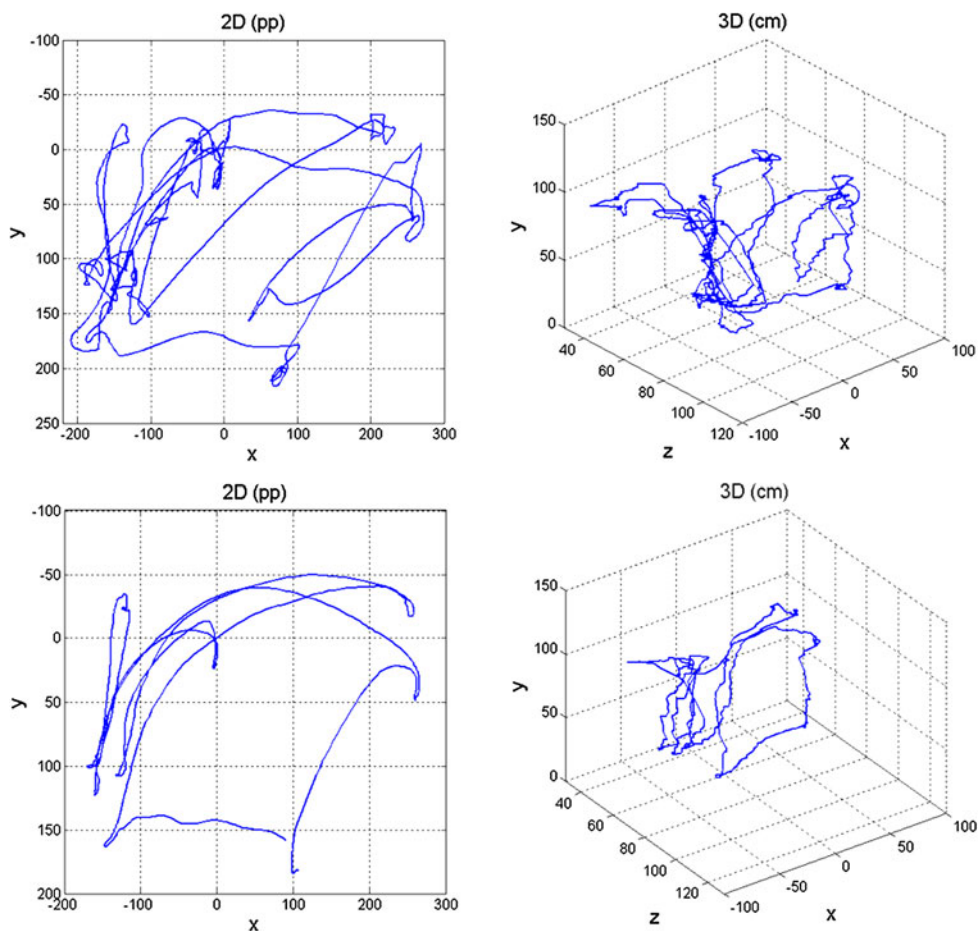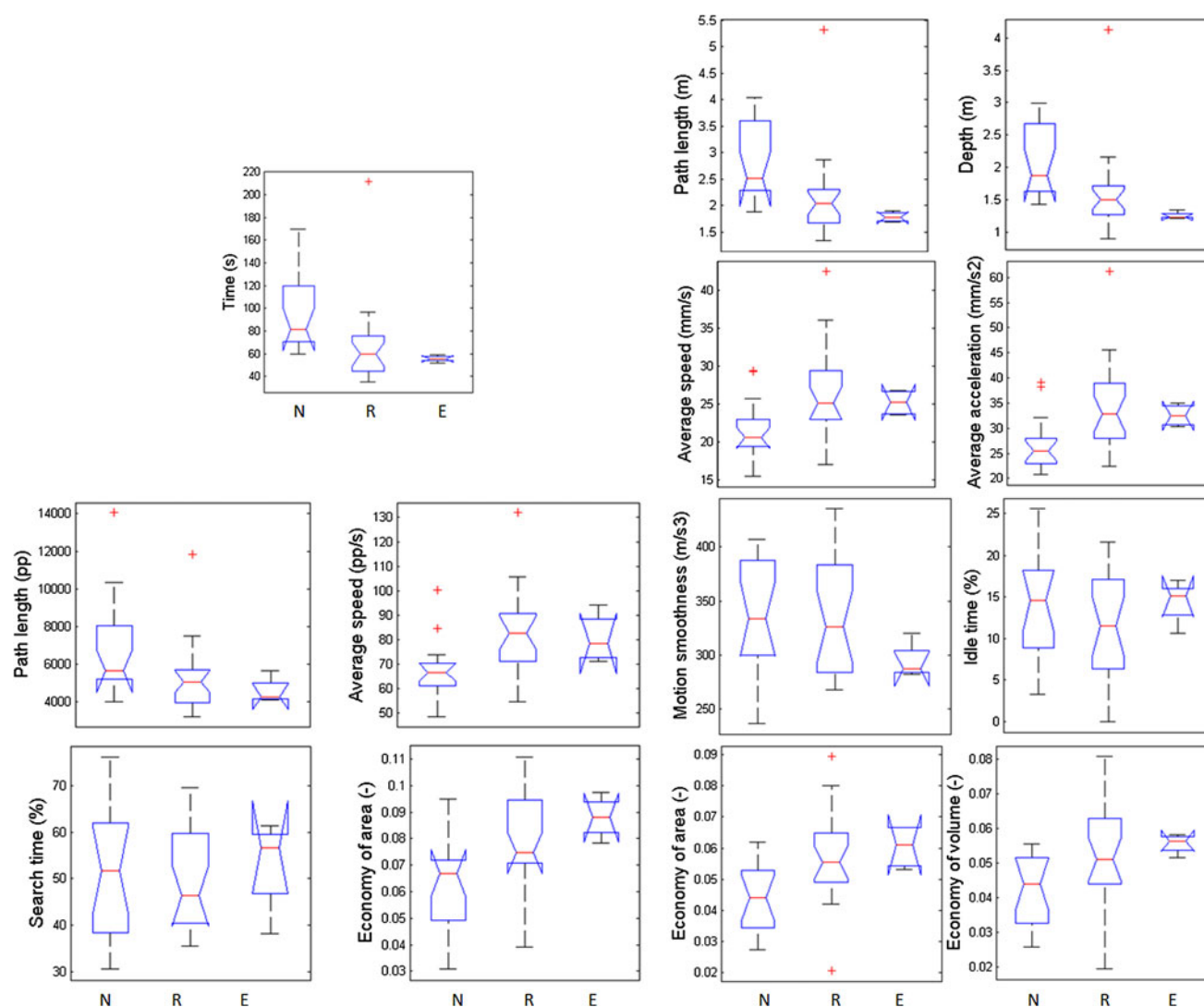
Fig. 5 Construct validation: representation of scores. *Left top*: Time. *Left bottom*: 2D metrics. *Right*: 3D metrics. Results are expressed as notched box diagrams, in which every box distinguishes lower quartile value, median and upper quartile value. Significance is shown where the notched sections of the boxes do not overlap each other

EVA showed good construct validity both for 2D and 3D metrics. Time, path length, and depth are common parameters, which have been validated in many clinical studies [11, 44–48]. Average speed and acceleration are less widespread metrics, but also showed good discriminative scores [11, 44, 49]. Additionally, three new metrics were introduced in this study: search time, economy of area, and economy of volume. Search time was derived from direct scrutiny of performance and was based on the observation that inexperienced participants had more trouble when trying to grasp the chickpeas from the pick-up zone. This 2D-exclusive metric did not prove to be significant in the end. However, it illustrates a new kind of metrics that might be defined from image analysis. Economy of area and economy of volume were introduced to quantify the efficiency in use of the working space, both of

the camera's field of view (2D) and the real surgical setting (3D). Results for both metrics demonstrated their construct validation for the task proposed. They point at the relevance of acknowledging the surgeon's spatial dominion of the workspace, which in the operating room may become more complicated due to camera movements and orientation. Further studies are needed to investigate whether the three new metrics proposed here are valid for different tasks.

A new aspect investigated in this work included the analysis of the relevance of depth information in skills assessment by comparing correspondent metrics in 2D and 3D spaces. Correlation analysis between metrics registered both in image and real space (path length, average speed, and economy of area) yield high values of similarity. However, we must consider that box trainers often are a

Table 2 Validation results. Construct validation is marked with x where significant differences across the three groups occur (Kruskal–Wallis, K–W). For pairs of groups, construct validation is explicitly indicated by said groups initial letter (Mann-Whitney, M-W). Concurrent validation is presented as Pearson's correlation (ρ)

Metrics	Construct		Concurrent
	Construct (K–W)	Construct (M-W)	Correlation
Time	X	N-R, N-E	1*
2D metrics			
Path length		N-R	-
Av. speed	X	N-R, N-E	-
Search time			-
Economy of area	X	N-R, N-E	-
3D metrics			
Path length	X	N-R, N-E	0.97*
Depth	X	N-R, N-E	0.82*
Av. speed	X	N-R	0.94*
Av. acceleration	X	N-R, N-E	0.9*
Idle time			0.71*
Motion smoothness			0.03
Economy of area	X	N-R, N-E	0.86*
Economy of volume	X	N-R, N-E	0.85*

*Reflects correlation at $p < 0.05$

Table 3 Comparison between 2D and 3D metrics. For each metric, variation of p value obtained in construct validation is given (ΔP). Positive values (bold letter font) reflect higher statistical differences for 3D metrics. Correlation between 2D and 3D counterparts is given for each metric by means of Pearson's coefficient (ρ)

	ΔP				**Correlation**
	All	N-R	N-E	R-E	
Path length	**0.05**	**0.03**	0,146	**0.506**	0,92
Average speed	−0.001	0	−0.035	**0.162**	0,864
Economy of area	**0.006**	**0.006**	**0.007**	**0.119**	0,887

best-case scenario for training, where the camera is fixed, and thus depth-related skills may not be so determining as in the operating room. Depth information could be more or less relevant as an assessment metric depending on the viewpoint of the camera. Another fact to consider is that the 2D image space may vary depending on the endoscope's position, which may or may not have influence on performance. Nevertheless, we have shown that statistical significance improved from 2D to 3D for path length and economy of area, while remaining more or less constant for mean speed. This, along with other validation studies regarding depth information [34, 35, 47, 48], suggests the

importance of this information and the need for further exploration of their relevance.

Motion analysis has been proven an important asset in the assessment of MIS psychomotor skills, especially during the first stages of training [12, 16, 34]. In this study, we have shown that passive tracking of the laparoscopic instruments can be a reliable source of information for psychomotor skills assessment. However, due to its simplicity, video-based tracking technology can be easily applied to other areas of surgical career and lifelong learning. As such, EVA could be proved useful not only for the assessment of residents but also in the selection of candidates suited for surgical training. In aviation, for example, selection of pilot trainees is based on a series of tests that cover all abilities believed to be relevant for a "good pilot" [50]. Those include: intelligence tests, mental agility, ability to prioritize multiple inputs, and spatial orientation. Similar selection of candidates for surgical training should be done based on candidates' aptitudes. Van de Loo et al. [51] and Gilligan et al. [52] found that future "good surgeons" are determined by a compendium of factors, such as intelligence, operative skill, work attitude, stability, and cooperation. In this context, the EVA tracking system might be used to assess operative (psychomotor) skills. Additionally, EVA could be used as a part of recertification and lifelong learning programs for continuous education of surgeons; due to the potential portability of the system, EVA could be introduced in more complex scenarios, such as the operating room.

In the current state, however, the EVA tracking system presents further challenges that need to be addressed before its full deployment. The first challenge deals with computational cost. Currently, the system's Matlab implementation limits the processing time to approximately 3-4 seconds per frame. Because analysis was performed offline, real time was not a constraint in this case. However, to provide immediate feedback to the trainee, real-time computations are desired [53]. Several possible alternatives are being considered at present, involving software (use of computer vision libraries, such as OpenCV, application of predictive tracking solutions based on Kalman filtering [54]), and hardware (use of graphic processing units – GPUs [55]) solutions.

The second challenge is related with robust management of occlusions and insertion/withdrawals of the instruments. These aspects do not significantly affect on controlled scenario, such as box trainers, but are of vital importance when considering its use during real interventions. In this sense, an interesting approach, although one that needs to take into account sterilization issues in the operating room, is the use of different artificial color markers as a complementary system to aid in the instrument identification
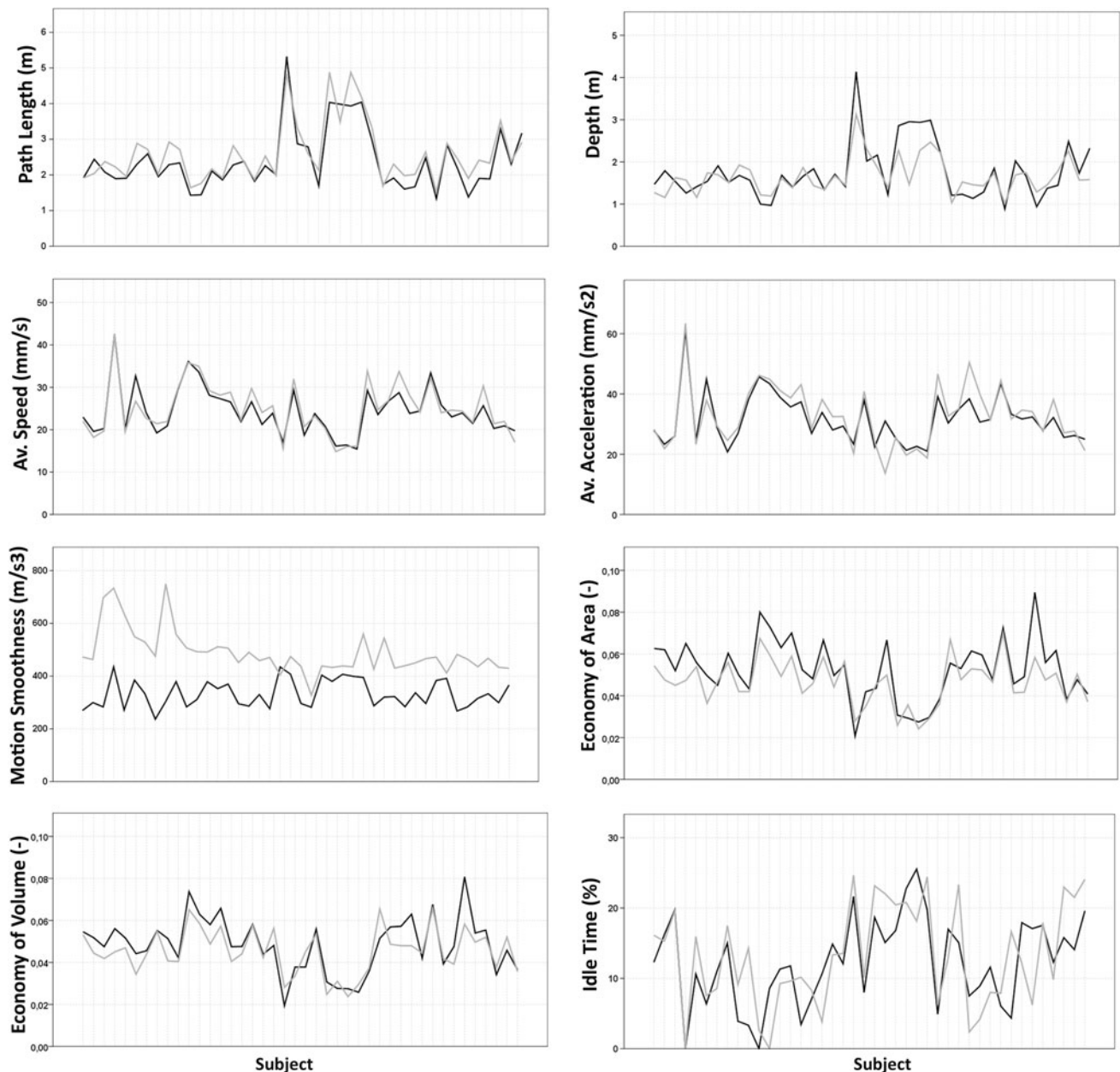
Fig. 6 Concurrent validation scores per subject, in order of participation in the study (x-axis). *Black*: scores obtained by EVA. *Gray*: scores obtained by TrEndo

process [23]. In this study occlusions were not a potential source of error as only one instrument was used, and the laparoscopic grasper employed was always visible on the screen. Moreover, the post-processing filtering was able to deal with outlier points caused by incorrect detections.

Further exploration of metrics will be addressed in future studies. New motion metrics, suited for video tracking, need to be developed and validated from a clinical point of view. For example, metrics related with the instrument's orientation might be explored [56, 57]. Thorough analysis of known metrics, such as time, path length, and speed, must be pursued with different tasks,

because, as shown in [7], depending on the skill assessed (e.g., grasping, cutting), a metric may be more or less relevant. To ensure a complete evaluation, quality metrics, such as end-product analysis or errors, also should be considered [8]. Because the purpose of this study was to validate the tracking capabilities of EVA for motion analysis, qualitative information was not included. However, the importance of such information has already been proven [58, 59]. With all this in mind, validation studies will need to be performed, involving different tasks and settings, as well as larger sample groups to confirm or decline the findings of this work.

Conclusions

The EVA tracking system has been introduced as a valid means for assessing MIS psychomotor skills of surgeons. The potential use of this system in surgical training programs can offer a valid alternative to sensor-based systems, providing a transparent and portable alternative to capture and analyze motion data for skills assessment. The study has presented its use on a box trainer; however, porting the system into the operating room for training purposes in real surgeries should be feasible in future versions. Additionally, a whole new range of applications for EVA are being sought, from cognitive skills training to image-guided surgery. To establish the clinical interest and viability extent of the EVA tracking system for all its intended uses, new studies on robustness and cost-effectiveness will be performed.

References

1. Cuschieri A (2005) Laparoscopic surgery: current status, issues and future developments. Surgeon 3:125–130
2. Aggarwal R, Moorthy K, Darzi A (2004) Laparoscopic skills training and assessment. Br J Surg 91:1549–1558
3. Harden RM, Stevenson M, Downie WW, Wilson GM (1975) Assessment of clinical competence using objective structured examination. Br Med J 1:447–451
4. van Sickle KR, Ritter EM, McClusky DA 3rd, Lederman A, Baghai M, Gallagher AG, Smith CD (2007) Attempted establishment of proficiency levels for laparoscopic performance on a national scale using simulation: the results from the 2004 SAGES minimally invasive surgical trainer-virtual reality (MIST-VR) learning center study. Surg Endosc 21:5–10
5. Usón J, Sánchez-Margallo FM, Pascual S, Climent S (2010) Formación en Cirugía Laparoscópica Paso a Paso, 4th edn. Minimally Invasive Surgery Centre Jesús Usón, Cáceres
6. Sánchez-González P, Oropesa I, Romero V, Fernández A, Albacete A, Asenjo E, Noguera J, Sánchez-Margallo FM, Burgos D, Gómez EJ (2010) TELMA: technology enhanced learning environment for minimally Invasive surgery. Procedia Comp Sci 3:316–321
7. Oropesa I, Sánchez-González P, Lamata P, Chmarra MK, Pagador JB, Sánchez-Margallo JA, Sánchez-Margallo FM, Gómez EJ (2011) Methods and tools for objective assessment of psycho-motor skills in laparoscopic surgery. J Surg Res 171:e81–e95. doi:10.1016/j.jss.2011.06.034
8. Fried GM, Feldman LS (2008) Objective assessment of technical performance. World J Surg 32:156–160
9. Chmarra MK, Bakker NH, Grimbergen CA, Dankelman J (2006) TrEndo, a device for tracking minimally invasive surgical instruments in training setups. Sens Actuat A-Physical 126:328–334
10. Rosen J, Brown JD, Barreca M, Chang L, Hannaford B, Sinanan M (2002) The Blue DRAGON–a system for monitoring the kinematics and the dynamics of endoscopic tools in minimally invasive surgery for objective laparoscopic skill assessment. Stud Health Technol Inform 85:412–418
11. Sokollik C, Gross J, Buess G (2004) New model for skills assessment and training progress in minimally invasive surgery. Surg Endosc 18:495–500
12. Datta V, Mackay S, Mandalia M, Darzi A (2001) The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. J Am Coll Surg 193:479–485
13. Peters TM (2006) Image-guidance for surgical procedures. Phys Med Biol 51:R505–R540. doi:10.1088/0031-9155/51/14/R01
14. Sánchez-González P, Cano AM, Oropesa I, Sánchez-Margallo FM, del Pozo F, Lamata P, Gómez EJ (2011) Laparoscopic video analysis for training and image guided surgery. Minim Invasive Ther Allied Technol. doi:10.3109/13645706.2010.541921
15. Sánchez-Margallo JA, Sánchez-Margallo FM, Pagador JB, Gómez-Aguilera EJ, Sánchez-González P, Usón J, Moreno J (2010) Video-based assistance system for training in minimally invasive surgery. Minim Invasive Ther Allied Technol 20:197–205
16. van Sickle KR, McClusky DA 3rd, Gallagher AG, Smith CD (2005) Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. Surg Endosc 19:1227–1231
17. Pellen MG, Horgan LF, Barton JR, Attwood SE (2009) Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 23:130–139
18. Ritter EM, Kindelan TW, Michael C, Pimentel EA, Bowyer MW (2007) Concurrent validity of augmented reality metrics applied to the fundamentals of laparoscopic surgery (FLS). Surg Endosc 21:1441–1445
19. Krupa A, Gangloff J, Doignon C, de Mathelin MF, Morel G, Leroy J, Soler L, Marescaux J (2003) Autonomous 3-D positioning of surgical instruments in robotized laparoscopic surgery using visual "servoing.". IEEE T Robotic Autom 9:842–853
20. Allen BF, Kasper F, Nataneli G, Dutson E, Faloutsos P (2011) Visual tracking of laparoscopic instruments in standard training environments. Stud Health Technol Inform 163:11–17
21. Speidel S, Delles M, Gutt C, Dillmann R (2006) Tracking of instruments in minimally invasive surgery for surgical skill analysis. Med Imag Augment Real 4091:148–155
22. Tonet O, Ramesh TU, Megali G, Dario P (2006) Tracking endoscopic instruments without localizer: image analysis-based approach. Stud Health Technol Inform 119:544–549
23. Bouarfa L, Akman O, Schneider A, Jonker PP, Dankelman J (2011) In vivo real-time tracking of surgical instruments in endoscopic video. Minim Invasive Ther Allied Technol DOI: 10.3109/13645706.2011.580764
24. Voros S, Long J, Cinquin P (2006) Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. Med Image Comput Comput Assist Interv 4190:535–542
25. Climent J, Marés P (2004) Automatic instrument localization in laparoscopic surgery. Electron Lett Comput Vis Image Anal 4:21–31

26. McKenna SJ, Charif HN, Frank T (2005) Towards video understanding for laparoscopic surgery: instrument tracking. Image and Vision Computing New Zealand Conference, Dunedin, New Zealand

27. Doignon C, Nageotte F, Maurin B, Krupa A (2008) Pose estimation and feature tracking for robot assisted surgery with medical imaging. In: Kragic D, Kyrik V (eds) Unifying perspectives in computational and robot vision, vol 8. Springer, New York, pp 79–101

28. Cano AM, Lamata P, Gayá F, del Pozo F, Gómez EJ (2006) New methods for video-based tracking of laparoscopic tools. In: Harders M, Székely G (eds) ISBMS 2006, LNCS, vol 4072. Springer, Heidelberg, pp 142–149

29. Cano AM, Sánchez-González P, Sánchez-Margallo FM, Oropesa I, del Pozo F, Gómez EJ (2008) Video-endoscopic image analysis for 3D reconstruction of the surgical scene. In: Sloten JV, Verdonck P, Nyssen M, Haueisen J (eds) 4th European Conference of the International Federation for Medical and Biological Engineering, IFMBE Proceedings vol 22. pp 923–926

30. Wolf R, Duchateau J, Cinquin P, Voros S (2011) 3D tracking of laparoscopic instruments using statistical and geometric modeling. In: Fichtinger G, Martel A, Peters T (eds) Medical image computing and computer-assisted intervention – MICCAI 2011, LNCS, vol 6891. pp 203–210

31. Satava RM, Cuschieri A, Hamdorf J (2003) Metrics for objective assessment. Surg Endosc 17:220–226

32. Chmarra MK, Jansen FW, Grimbergen CA, Dankelman J (2008) Retracting and seeking movements during laparoscopic goal-oriented movements. Is the shortest path length optimal? Surg Endosc 22:943–949

33. Chmarra MK, Dankelman J, van den Dobbelsteen JJ, Jansen FW (2008) Force feedback and basic laparoscopic skills. Surg Endosc 22:2140–2148

34. Chmarra MK, Klein S, de Winter JCF, Jansen FW, Dankelman J (2010) Objective classification of residents based on their psychomotor laparoscopic skills. Surg Endosc 24:1031–1039

35. Hiemstra E, Chmarra MK, Dankelman J, Jansen FW (2011) Intracorporeal suturing: economy of instrument movements using a box trainer model. J Minim Invasive Gynecol 18:494–499

36. Cano AM, Vara I, Sánchez-González P, Gómez EJ (2008) Laparoscopic image analysis for automatic tracking of surgical tools. In: Proceedings of computer assisted radiology and surgery (CARS 2008), vol 3. p S279

37. Brown DC (1971) Close-range camera calibration. Photogramm Eng 37:855–866

38. Bouguet JY (2010) Camera Calibration Toolbox for Matlab - Calibrating a stereo system, stereo image rectification and 3D stereo triangulation (2010). Pasadena, CA (USA): California Institute of Technology; 2010 Available at: http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example5.html. Accessed 27 Oct 2011

39. Gonzales RC, Woods RE (2002) Digital Image Processing, 2nd edn. Prentice-Hall, Englewood Cliffs, NJ

40. Jain AK, Dubes RC (1981) Algorithms for clustering data. Prentice-Hall, Upper Saddle River, NJ

41. Aggarwal R, Grantcharov T, Moorthy K, Milland T, Papasavas P, Dosis A, Bello F, Darzi A (2007) An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. Ann Surg 245:992–999

42. Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS. Surg Endosc 20:104–112

43. Chmarra MK, Kolkman W, Jansen FW, Grimbergen CA, Dankelman J (2007) The influence of experience and camera holding on laparoscopic instrument movements with the TrEndo tracking system. Surg Endosc 21:2069–2075

44. Yamaguchi S, Konishi K, Yasunaga T, Yoshida D, Kinjo N, Kobayashi K, Ieiri S, Okazaki K, Nakashima H, Tanoue K (2007) Construct validity for eye–hand coordination skill on a virtual reality laparoscopic surgical simulator. Surg Endosc 21: 2253–2257

45. Verdaasdonk EGG, Stassen LPS, Schijven MP, Dankelman J (2007) Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. Surg Endosc 21:1406–1412

46. Larsen CR, Grantcharov T, Aggarwal R, Tully A, Sørensen JL, Dalsgaard T, Ottesen B (2006) Objective assessment of gynecologic laparoscopic skills using the LapSimGyn virtual reality simulator. Surg Endosc 20:1460–1466

47. Stylopoulos N, Cotin S, Maithel S, Ottensmeyer M, Jackson P, Bardsley R, Neumann P, Rattner D, Dawson S (2004) Computer-enhanced laparoscopic training system (CELTS): bridging the gap. Surg Endosc 18:782–789

48. Egi H, Okajima M, Yoshimitsu M, Ikeda S, Miyata Y, Masugami H, Kawahara T, Kurita Y, Kaneko M, Asahara T (2008) Objective assessment of endoscopic surgical skills by analyzing direction-dependent dexterity using the hiroshima university endoscopic surgical assessment device (HUESAD). Surg Today 38:705–710

49. Megali G, Sinigaglia S, Tonet O, Cavallo F, Dario P (2006) Understanding expertise in surgical gesture by means of Hidden Markov Models. In: Proc IEEE RAS EMBS Int Conf Biomed Robot Biomechatron 2006:625–630

50. Bell JA (1998) Royal air force selection procedures. Ann R Coll Surg Engl 70:270–275

51. Van de Loo RPJM (1998) Selection of surgical trainees in The Netherlands. Ann R Coll Surg Engl 70:277–279

52. Gilligan JH, Treasure T, Watts C (1996) Incorporating psychometric measures in selecting and developing surgeons. J Manag Med 10:5–16

53. Lamata P, Gomez EJ, Bello F, Kneebone RL, Aggarwal R, Lamata F (2006) Conceptual framework for laparoscopic VR simulators. IEEE Comput Graph Appl 26:69–79

54. Kalman RE (1960) A new approach to linear filtering and prediction problems. Trans ASME, J Basic Eng 82:35–45

55. Sinha SN, Frahm JM, Pollefeys M, Gene Y (2011) Feature tracking and matching in video using programmable graphics hardware. Mach Vision Appl 22:207–217

56. Rosen J, Brown JD, Chang L, Sinanan MN, Hannaford B (2006) Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete Markov model. IEEE Trans Biomed Eng 53:399–413

57. Bann SD, Khan MS, Darzi AW (2003) Measurement of surgical dexterity using motion analysis of simple bench tasks. World J Surg 27:390–394

58. Sherman V, Feldman L, Stanbridge D, Kazmi R, Fried G (2005) Assessing the learning curve for the acquisition of laparoscopic skills on a virtual reality simulator. Surg Endosc 19:678–682

59. van Dongen KW, Tournoij E, van der Zee DC, Schijven MP, Broeders IAMJ (2007) Construct validity of the LapSim: can the LapSim virtual reality simulator distinguish between novices and experts? Surg Endosc 21:1413–1417