
Towards Intelligent Surgical Microscopes: Surgeons' Gaze and Instrument Tracking

Shahram Eivazi

University of Tübingen
Perception Engineering, Sand
14, 72076 Tübingen, Germany
shahram.eivazi@mnf.uni-
tuebingen.de

Enkelejda Kasneci

University of Tübingen
Perception Engineering, Sand
14, 72076 Tübingen, Germany
enkelejda.kasneci@uni-
tuebingen.de

Wolfgang Fuhl

University of Tübingen
Perception Engineering, Sand
14, 72076 Tübingen, Germany
wolfgang.fuhl@uni-tuebingen.de

Abstract

After many decades of research, the presence of intelligent user interfaces is unquestionable in any modern operating room (OR). For the first time, we aim to bring proactive intelligent systems into microsurgery OR. The first step towards an intelligent surgical microscope is to design an activity-aware microscope. In this paper, we present a novel system that we have built to record both eyes and instruments movements of surgeons while operating with a surgical microscope. We present a case study in micro-neurosurgery to show how the system monitors the surgeon's activities. We achieved about 1 mm accuracy for gaze and instrument tracking. Now real-time ecologically valid data can be used to design, for example, a self-adjustable microscope.

Author Keywords

eye tracking; intelligent surgical microscope; hands-free.

ACM Classification Keywords

H.5.2 [User Interfaces]: Prototyping

Introduction

To date, advances in microsurgery have enabled the development of more complex medical procedures. Although patients benefit from the use of new surgical microscope by reducing recovery time, the risk of hemorrhaging, and expo-

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
IUI'17 Companion, March 13-16, 2017, Limassol, Cyprus
ACM 978-1-4503-4893-5/17/03.
<http://dx.doi.org/10.1145/3030024.3038269>



Figure 1: A neurosurgeon removes hands from the operative field to adjust the surgical microscope.

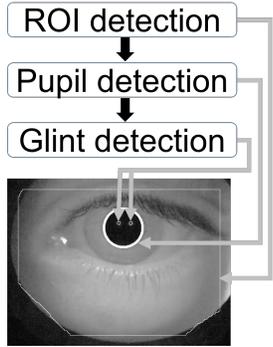


Figure 2: On the top the workflow of the main detection steps are shown. Down is the result for each detection step.

$$ROI(x, y, W) = \begin{cases} 1, & GM < LM(x, y, W) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Where GM is the mean value of the downsampled image, LM is the local mean in the window W and x, y is the position of the inspected pixel.

sure to infections, for surgeons the procedure of viewing an anatomical structure through a high magnification increases the complexity of operation. For example, Eivazi et al. [1, 2] showed that manual interaction with a neurosurgical microscope (positioning, zooming, and focusing) resulted to disruptive events and hazard situations (Figure 1).

We believe, an intelligent user interface is a promising answer to the increasing needs of real time computer-assisted microsurgical systems. As such, we take the first step towards intelligent surgical microscope by detecting surgeons' activities during a micro-neurosurgery procedure. In this paper, we evaluate a state-of-the-art eye and instrument tracker for surgical microscope.

System design

We used the recent Eivazi et al. [3] surgical microscope eye tracker. We extended their system to include two infrared light (LED) glints (cornea reflection to the light) for head movement compensation.

To determine the gaze point of a subject looking through the microscope, we first extract a Region of Interest (ROI) in the image around the subject's eye (Figure 2). Our algorithm downscales the image and then calculates the mean gray value. The decision whether an image pixel belongs to the ROI is taken based on the mean value of the surrounding pixels as described by Equation.

Next, to automatically detect the pupil, we employed the EISE algorithm [6] with a validity threshold of 50. The EISE algorithm operates on Canny-edge filtered eye images and is one the best performing algorithms among the state-of-the-art approaches in both head-mounted [6] and remote eye-tracking eye images [5]. In order to detect glint points we used the second step of the EISE algorithm related to blob detection.

For calibration, our system calculates the vector between the pupil center and the center of both glints. A least squares polynomial fit is calculated based on the vectors as input variables and the calibration point positions as polynomial result. We used a three dimensional polynomial for the left and right eye position, resulting thus in four polynomials. To map the correct points in the calibration phase to the correct points of the calibration grid, we used k-means clustering and standard deviation based outliers removal in each cluster.

Instrument tracking

Due to the extensive set of different tools in microsurgery, we developed a shape independent instrument tracking algorithm. The main detection features for the algorithm are the dark color of the instruments and their size. Therefore, the algorithm starts with a color transformation of the input image (Figure 3(a)). Equation 2 shows the formula used to transform the red, green, blue color image in a gray scale image.

$$GV(x, y) = (|\Delta I^-(x, y)|)^2 * (\emptyset I^-(x, y))^2 \quad (2)$$

$|\Delta I^-(x, y)|$ is the summed difference between all three channels as positive value. The I^- denotes that we used the inverted intensity image, meaning that high values are set to the corresponding low value and vice versa. With \emptyset we denote the average value between all three color channels ($\frac{Red+Green+Blue}{3}$). The power two of the inverted gray value and the difference behaves like a high pass filter. The result of this equation 2 can be seen in Figure 3(b).

The next step of the algorithm is to decide which pixel belongs to an instrument. This is done by inspecting an area surrounding the actual position, denoted with W in equation 3. The threshold (0.1) used in Equation 3 shows 10%

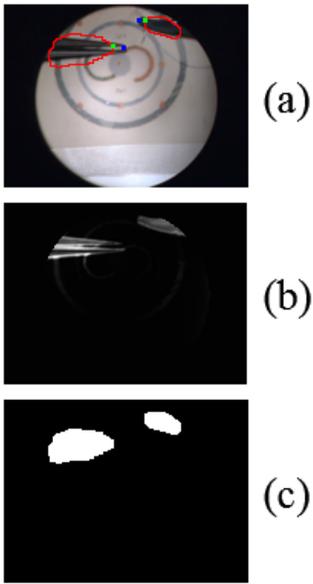


Figure 3: (a) the input image with registered results. The red outline is the automatic detected instrument. The green dot is the estimated instrument pointing position and the blue dot is the manual labeled instrument pointing position. The result of the color transform is shown in (b) and (c) is the binary image generated out of the color transformed.

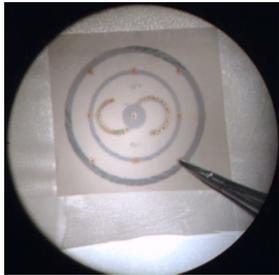


Figure 4: The actual experiment design. Participants were asked to cut and suture precisely along two inner green lines in the left and right side of central circle.

of the magnitude resulting from the color transformation. As such the result of the color transformation (equation 2) has to be normalized to the range 0 – 1. The result of this segmentation is shown in Figure 3(c).

$$BIN(x, y, W) = \begin{cases} 1, & \frac{\sum_{i=-W}^W \sum_{j=-W}^W GV(x+i, y+j)}{(2*W+1)^2} > 0.1 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Finally, all connected white pixels are grouped into a segment, if the amount of pixels is higher than 200. The outline of both found segments is shown in Figure 3(a) as red line. The estimation of the instrument pointing is done by calculating the Principal Component Analysis of all segment points. The first resulting vector is used for calculating the intersection with the outline and this point is then used as instrument pointing estimation. This is drawn in Figure 3(a) as blue dot. The green dot in Figure 3(a) is manual labeled instrument pointing position for evaluation purpose.

Evaluation

The experiment was conducted in the neurosurgery department of Helsinki University Central Hospital. Ten neurosurgeons were asked to cut precisely along a curved line drawn on the top of a latex glove sheet (Figure 4). Three Ps3Eye cameras with a sampling rate of 60Hz were used to record left and right eye (eye cameras), and the microscope field of view (scene camera).

For the purpose of this paper, we used a nine-point calibration and a nine-point evaluation procedure to report our gaze estimation accuracy (Figure 4). The evaluation procedure last for less than one minute. We had to remove two participants' data due to lack of glint in eye images. The results can be seen in Figure 5. An error of 0.02 corresponds

to 16 pixels in the scene image which is about 1mm. Participant 1 had the worst results as the eyelashes hide most of the glint points and therefore reduce the amount of data for fitting the prediction model. This was also same for Participant 8. To solve this we aim to change the direction of eye cameras in future work.

We evaluated the accuracy of the instrument tracking algorithm by manually annotating 1000 images randomly and compare that with our algorithm estimation. The manual labeling is a time consuming task and thus for the purpose of this paper we only use one video from a suturing task. In future work we will annotate all participants videos.

Each image has a resolution of 640x480 pixels. As evaluation metric we used the relative error which is the Euclidean distance normalized by the image diagonal (800 pixel). The algorithm detected 1499 instruments correct and missed 64 from a total of 1541 labeled instruments. In addition, it misclassified 8 segments as instruments. Undetected instruments are from images where only a small part of the instrument was present, and therefore the region of this segment was too small. The misclassified instruments are from splitting up an instrument into two. This happens if the instrument is opened and the connecting part of the tool is not present in the image. The results of the instrument tracking is shown in Figure 6. The average error of 0.02 correspond to 16 pixels in the scene image (i.e., 1mm).

Implications

The eye and instrument tracker presented in this work provides valuable opportunities to record detailed surgeons activities in real-time microsurgery. We can now start to paint a picture of various ways in which our system helps researchers to develop computer-assistant microsurgical systems without the need to distract surgeons from their

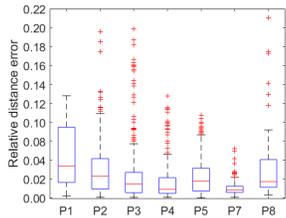


Figure 5: The relative error for the calibration tracking. The target image has a resolution of 640x480 and the diagonal for error normalization is therefor 800 pixels.

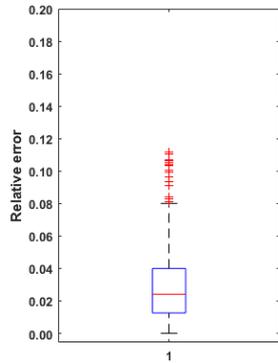


Figure 6: The relative error for the instrument tracking. The input image has a resolution of 640x480 and the diagonal for error normalization is therefor 800 pixels.

routine tasks.

The broad implication is to apply our system for the self-adjustable motorized surgical microscope [7]. Such a system offers solutions to numerous issues related to adjusting the microscope settings manually by reducing the number of interruptions. We envision to use gaze and location of instruments to direct the microscope movement commands and automatically adjust the focus and zoom level in the target of interest.

Moreover, given the complexity of microsurgery and longitudinal training, the reason for applying our system as a tool to evaluate the surgeon performance is obvious. Recently, we have seen growing interest in analysis of expert and novice differences in microsurgery [4]. However, to date none of these studies is conducted based on a surgical microscope. Thus, we propose our tracking measures to be used as a suitable objective metric for the automatic assessment of one's level of expertise in microsurgery.

Acknowledgements

Dr. Ahmad Hafez from Helsinki university central hospital has designed the novel surgical task presented in this study. We thank all surgeons who participated in our study.

REFERENCES

1. Hoorieh Afkari, Shahram Eivazi, Roman Bednarik, and Susanne Mäkelä. 2014. The potentials for hands-free interaction in micro-neurosurgery. In *Proc. of 8th Nordic Conference on Human-Computer Interaction*. ACM, 401–410.
2. Shahram Eivazi, Hoorieh Afkari, Roman Bednarik, Ville Leinonen, Markku Tukiainen, and Juha E Jääskeläinen. 2015a. Analysis of disruptive events and precarious situations caused by interaction with neurosurgical microscope. *Acta neurochirurgica* 157, 7 (2015), 1147–1154.
3. Shahram Eivazi, Roman Bednarik, Ville Leinonen, Mikael von und zu Fraunberg, and Juha E Jääskeläinen. 2015b. Embedding an Eye Tracker Into a Surgical Microscope: Requirements, Design, and Implementation. *IEEE Sensors Journal* 16, 7 (2015), 2070–2078.
4. Shahram Eivazi, Roman Bednarik, Markku Tukiainen, Mikael von und zu Fraunberg, Ville Leinonen, and Juha E Jääskeläinen. 2012. Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings. In *Proc. of the Symposium on Eye Tracking Research and Applications*. ACM, 377–380.
5. W. Fuhl, D. Geisler, T. Santini, and E. Kasneci. 2016a. Evaluation of State-of-the-Art Pupil Detection Algorithms on Remote Eye Images. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2016 ACM International Symposium on Wearable Computers*. ACM.
6. Wolfgang Fuhl, Tiago Santini, Thomas Kübler, and Enkelejda Kasneci. 2016b. ElSe: Ellipse Selection for Robust Pupil Detection in Real-World Environments. In *Proc. of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. ACM, New York, NY, USA, 123–130.
7. Sven R Kantelhardt, Markus Finke, Achim Schweikard, and Alf Giese. 2013. Evaluation of a completely robotized neurosurgical operating microscope. *Neurosurgery* 72 (2013), A19–A26.