A Study of Expert/Novice Perception in Arthroscopic Shoulder Surgery

Myat Su Yin
Faculty of ICT
Mahidol University
Nakhon Pathom, Thailand
+66 2 441 0909
myatsu.yin@mahidol.ac.th

Peter Haddawy
Faculty of ICT
Mahidol University
Nakhon Pathom, Thailand
+66 2 441 0909
peter.had@mahidol.ac.th

Benedikt Hosp
Dept. of Computer Science
University of Tuebingen
Tuebingen, Germany
+497 071 297 8996
benedikt.hosp@uni-tuebingen.de

Paphon Sa-ngasoongsong
Fac. of Medicine, Ramathibodi Hospital
Mahidol University
Bangkok, Thailand
+66 2 201 1589
paphonortho@gmail.com

Thanwarat Tanprathumwong
Faculty of ICT
Mahidol University
Bangkok, Thailand
+66 2 441 0909
321view@gmail.com

Madereen Sayo
Faculty of ICT
Mahidol University
Nakhon Pathom, Thailand
+66 2 441 0909
madereen.ss@gmail.com

Supawit Yangyuenpradorn
Faculty of ICT
Mahidol University
Nakhon Pathom, Thailand
+66 2 441 0909
bcc32909@gmail.com

Akara Supratak
Faculty of ICT
Mahidol University
Nakhon Pathom, Thailand
+66 2 441 0909
akara.sup@mahidol.ac.th

ABSTRACT

Arthroscopic shoulder surgery is an advanced orthopedic surgical procedure, which is particularly challenging due to the complex anatomy of the shoulder, and tight spaces for navigation, which also limits the view from the arthroscope. In carrying out arthroscopy, the ability to quickly and effectively navigate through the joint to reach a desired location is essential. Novices often experience confusion in trying to triangulate the information from arthroscopy output with the background knowledge of anatomy while orienting and navigating the instruments. In this paper, we report on the results of the first cadaveric eye-tracking study of arthroscopic surgery in which we investigate differences in perception between experts and novices. Novices' perception is analyzed with cognitive load analysis throughout the procedure and specifically, during the portions of the procedure in which subjects are observed to be confused. In investigating such portions, the gaze data analysis is supplemented with head rotations and acceleration information from gyroscope and accelerometer sensors from the eye tracker. We also use the gathered eye tracking metrics to construct a model to classify subjects into expert/novice. We find statistically significant relations between head movement as well as pupil diameter and periods of confusion. We identify a subset of the metrics that we use to build a simple classifier that is able to distinguish between novices and experts with accuracy of 84%.

CCS Concepts

• Applied Computing → Life and Medical Science → Health Informatics • Human-centered computing → Human computer interaction (HCI) → HCI design and evaluation methods → Usability testing

Keywords

Eye-tracking, arthroscopic surgery, cadaver, classification, perception, attentional focus

1. INTRODUCTION

Arthroscopic shoulder surgery is an advanced orthopedic surgical procedure, which is particularly challenging due to the complex anatomy of the shoulder, and tight spaces for navigation, which also limits the view from the arthroscope. It is used to treat a number of disorders such as repair of torn tendons and rectifying chronic dislocation, as well as for diagnosis. In all of these procedures, the ability to quickly and effectively navigate through the joint to reach the desired location is essential. An important aspect of navigation is the ability to quickly recognize anatomical landmarks and to focus attention on the appropriate region of the

*ICMHI 202*0, August 14–16, 2020, Kamakura City, Japan

ACM ISBN 978-1-4503-7776-8/20/08…$15.00
https://doi.org/10.1145/3418094.3418135

arthroscope image. For assessment and training it is important to have an objective assessment of such perceptual and attentional aspects and to detect portions of the procedure where students may become confused.

In this paper, we report on the results of the first cadaver-based study to analyze and compare expert and novice eye movement patterns in performing arthroscopic surgery. We study the diagnostic arthroscopic shoulder surgery task since it involves navigating to various parts of the shoulder and inspecting them and thus allows us to focus purely on navigation skills. The existing studies on comparing eye movement patterns between experienced surgeons and novices have predominantly used VR training simulators [1–6], still images of the surgery [7,8] or physical box trainers [9]. We use so-called soft cadavers, which are specially prepared so as to retain the natural tissue properties. This means that our study is able to capture important aspects of the surgery such as tactile feedback and surgical setup not captured by simulations. Our work is also the first to study arthroscopic shoulder surgery. Previous eye-tracking studies of surgery have concentrated predominantly on laparoscopic surgery which usually involves anatomy of the abdomen. In contrast, the diagnosis of the shoulder requires the surgeon to navigate the arthroscope through bones and muscles inside the rounded shoulder joint.

Experts can usually smoothly maneuver the arthroscope instruments with the automaticity developed through experience. In contrast, novices often experience confusion in trying to locate the anatomical landmarks from the magnified view of the operating site on the arthroscope output. Previous studies in the area of human-computer interfaces and intelligent tutoring have found pupil size and head movement to be associated with periods of confusion [10,11]. We sought to determine whether these metrics can also be used to detect confusion during shoulder arthroscopy and found positive relationships between both and novice states of confusion. Ours is the first study to attempt to use objective metrics to detect confusion during surgery.

An effective assessment instrument should be able to distinguish between performance of subjects with varying levels of experience and expertise. We thus analyze the differences in gaze metrics between experts and two groups of novices of varying experience. We identify a small subset of the metrics with good discriminatory power and use them to build a simple classifier that is able to distinguish between novices and experts with high accuracy. This leads us to conclude that there are significant differences in perceptual parameters between novices and experts in arthroscopic surgery that could be used for objective assessment as well as tutoring

2. RELATED WORK

Arthroscopic skills are difficult to acquire because they require use of multiple tools, using both hands while viewing the surgical site on a two-dimensional display, with constant vigilance to the operating environment [12]. Arthroscopic surgery is taught as a core component in a majority of orthopedic residency programs. Cadavers are often the first choice of surgeons for practice because they provide a real anatomical experience [13]. Other methods that have been tested with varying success in orthopedic teaching include interactive computer simulation [14], physical simulation environments [15] and virtual reality simulators [16,17]. Approaches in assessing arthroscopic surgical skills include Global Rating Scales [18], motion analysis [19], virtual reality simulators [16, 17], and simple bench model arthroscopic simulators [20].

Eye tracking studies comparing experts and novices have been carried out in a number of surgical domains. Tien et al. [21] compared the gaze behaviors of experts and junior surgeons during key stages of a live open inguinal hernia repair. They found that experts have a higher fixation frequency and concluded that it could be due to lower mental demand resulting from automaticity developed through practice. Similar findings are reported by Erridge et al. [22] during live laparoscopic gastric bypass surgery. Novices were found to pay less attention to the operative site but more to the sterile field. A number of studies of eye movement patterns of experts and novices [3,23,24] found that experts tend to fixate on the target more often than the instruments. Meanwhile, Law et al. [23]reported that novices either alternate their gaze between the target and instruments, focus on objects in between the target and the instruments, or follow the instrument on its way to the target. A study by Hermens et al. [24] also found differences in eye movement statistics between experts and novices. The experts in their study reportedly had lower saccadic rates and higher peak velocity, independent of where these eye movements were aimed. Similarly, in a study of global eye movement parameters of expert and non-expert participants, Kocak et al. [9] found that experts had significantly lower saccade rates and higher peak velocity than non-experts.

Beyond analysis of eye movement metrics, a number of studies have used the metrics to build models to classify subjects into expert and novice. Eye metrics and tool motion data have been considered as features in assessing the skill of a surgeon while performing functional endoscopic sinus surgery [25]. Hidden Markov models were built for seven different surgeries in two levels of expertise using the eye-gaze locations and the surgical tools motions. The findings revealed that eye-gaze data contains the skill-related structures, and combining it with the surgical tool motion data improves the classifier performance. Richstone et al. [26] used eye movement metrics to develop models to classify surgeons into experts and nonexperts. In a simulated surgery they achieved 91.9% and 92.9% accuracy with the linear discriminant analysis and neural network analysis, respectively and 81.0% and 90.7% accuracy in a live operating room setting. Eivazi et al. [27] used a random forest classifier to classify micro-surgeons in the cutting and suturing tasks and achieved a 70% recognition rate for the detection of expert and novice groups. Rose and Pedowitz [28] investigate the assessment of basic arthroscopy skills using virtual reality modules developed through task deconstruction. Participants with the most arthroscopic experience performed better and were more consistent than novices on all 3 virtual reality modules. Greater arthroscopic experience correlates with more symmetry of ambidextrous performance.

While no work has investigated detection of confusion during surgery, detection of cognitive affective states such as confusion and boredom has been studied in the field of Intelligent Tutoring Systems. Pachman and colleagues [29] used eye tracking for early detection of confusion in a digital learning environment. In their study, the participants were asked to solve problems while their eye trajectories were recorded and this data was triangulated with self-ratings of confusion and cued retrospective verbal reports. Delucia and colleagues [30] sought to determine whether eye movements reflect confusion while users completed tasks with two simulated devices. They measured confusion using a subjective Likert measure in which subjects were asked to rate their agreement with the statement “I was confused” and were not able to find consistent common correlation patterns between the variables for both devices, but they found that higher confusion ratings were positively correlated with the total fixation time on the whole screen, mean fixation duration and task completion time. Lallé and colleagues [31] included pupil diameter and head distance to the target as the predictors of the user’s confusion. They studied various combinations of gaze, pupil diameter, head distance and mouse events as predictors. The authors concluded that features of pupil size are strong predictors of confusion, which is consistent with the fact that pupil size is correlated with cognitive load, which plausibly correlates with confusion.

3. PARTICIPANTS, MATERIALS AND METHODS

After obtaining approval from the Mahidol University Institutional Review Board, a total of thirteen participants (4 Females) were recruited. They consisted of four fellows (two to ten years of experience) from the Department of Orthopaedics, Faculty of Medicine Ramathibodi Hospital, Mahidol University, and nine residents from the Orthopaedic Surgery Residency Program there. Five of the residents were in the third year and four in the fourth year. The residents were at an early stage of orthopedic training and were without prior arthroscopy experience. All the participants had normal or corrected-to-normal vision.

Eye gaze data was recorded using the Tobii Pro eye tracker (Tobii Glasses 2.0, Tobii Sweden), which was calibrated by looking at a marker placed near the arthroscopic output screen. The cadaver (Male, 52 years old) was set up in the beach-chair position. An expert surgeon prepared the arthroscope setup (ConMed Linvatec) and inserted the primary portals into the shoulder prior to the procedure. The arthroscope camera output was displayed on a 52-inch screen which was placed four feet away from the participant. ARUCO markers were also placed around the screen in order to identify the screen in a later stage. Each participant was first acquainted with the cadaver setup, the diagnostic shoulder arthroscopy steps, and the evaluation study protocol. Each participant was asked to navigate and diagnose twelve anatomical landmarks within the shoulder in sequence (Table I). The portion of the shoulder anatomy from viewing with the scope in the posterior portal and four visible landmarks 2, 7, 10 and 11 are shown in Figure 1. Among them, some are easy to navigate to and diagnose while some are more difficult. The landmarks which are categorized by the expert as hard to diagnose are highlighted and explanations are provided in Table 1.

For each landmark, the expert provided explicit verbal instructions with the name of the landmark (e.g. “Start Biceps tendon”) to navigate to and upon arrival at the landmark, the expert called out its name (e.g. “reached Biceps tendon”). The start and end times for each landmark navigation task were recorded as part of the data stream. Throughout the procedure, a think-aloud protocol was used and the participants were asked to describe their immediate objective, actions and any points at which they became confused (when they could not find the landmark or they did not recognize the part of the anatomy they were in).

In addition to the self-reported confusion, a member of the investigation team also monitored the participants and recorded portions of the performance as confusion in situations when a participant paused or made non-goal directed movements for a period of time which was followed by the attending surgeon’s assisting intervention. The study spanned two days, with the left shoulder of the cadaver used on the first day for six participants, and the right shoulder used on the second day for eight participants.

Table 1. Twelve anatomical landmarks to diagnose (The landmarks which are categorized by the expert as hard to diagnose are highlighted.)

1.	Rotator interval
2.	Biceps tendon & Biceps probe test: easy to find long head biceps (LHB) but difficult for use probe to handle LHB (need another hand to control the probe)
3.	Biceps anchor
4.	Labral superior to anterior
5.	IGHL
6.	Subscapularis tendon and insertion
7.	Anterosuperior cuff insertion (Supraspinatus)
8.	Posterosuperior cuff insertion (Infraspinatus): difficult to move from supraspinatus to infraspinatus (need to control the camera backward along the tendon)
9.	Bare area
10.	Inferior recess: difficult move from the posterior chamber downward direction to the inferior chamber
11.	Posterior labral: difficult to slide the camera from inferior chamber to posterior than to superior chamber (the camera could easily back out from the trocar due to the limited space)
12.	Back to Rotator interval

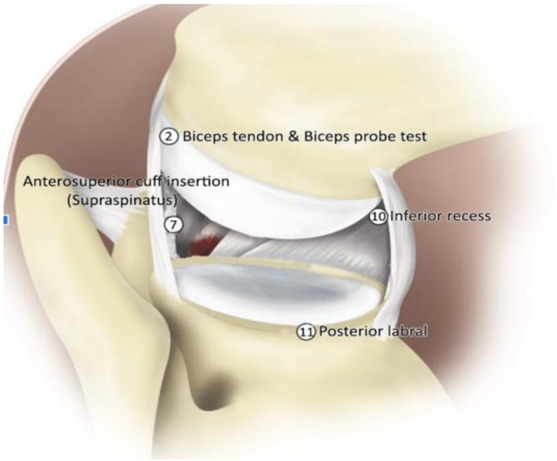


Figure 1. Portion of the shoulder anatomy with Landmarks 2, 7, 10, and 11

4. DATA PREPARATION

From a preliminary study, we found that while surgeons diagnose a landmark, they tend to look at the center of the scope image and

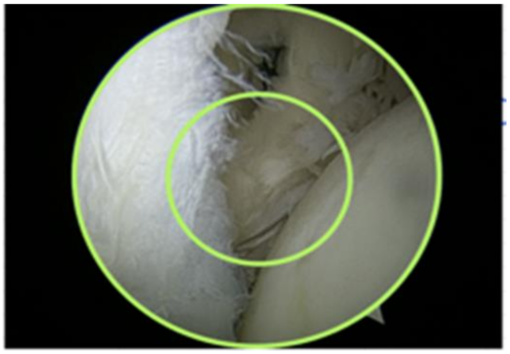


Figure 2. Inner and outer circles on the scope output

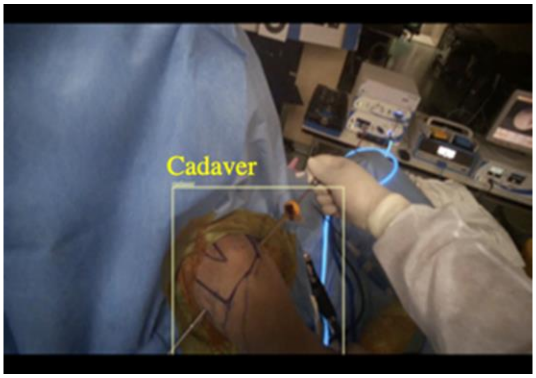


Figure 3. Detected cadaver shoulder on the video

tend to look at the area near the circumference of the scope image in the direction of the next landmark to visit before moving the scope. We, therefore, define four areas of interest (AOIs): the center area of the scope image (the inner circle) (Figure 2), the outer area of the scope image (outer circle) (Figure 2), the arthroscope output screen (outside of the scope image), and the shoulder area on the cadaver (Figure 3).

Eye-tracking metrics considered in this study are the rate and duration of fixations/saccades, the time to first fixation and the duration of the first fixation were calculated with the Tobii-I-VT Attention Filter using default parameters. Fixation is the visual gaze on a single location and saccades are the rapid movements of the eyes that abruptly change the point of fixation.

The field view videos of the eye tracker were processed to demark the AOI's. The arthroscope output screen was detected using ARUCO markers and the scope view on the screen was detected using a simple circle detection method (cv2.circle()). The cadaver area in the video frames was detected by using the YOLOV3 CNN object detection model [32] trained using transfer learning. The cadaver shoulder in the video frames was labeled using the video labeler app from Matlab (R2019b). We used the video frames from three participants for the left shoulder and from two participants for the right shoulder area.

5. ANALYSIS AND DISCUSSION

The two most commonly studied features of eye movement are fixations and saccades. Fixations are visual gazes on a single location whereas saccades are rapid eye movements between fixations. Among the large number of possible eye tracking metrics, those commonly used in medical studies are fixation rate (number of fixations per second), saccade rate (number of saccades per second), fixation duration (length of each fixation), saccade duration, average time to first fixation, and duration of first fixation [9,21,26,33,34]. We thus chose these metrics for the current study. Along with the eye metrics, we used the completion time as an objective measure of skill. We categorized participants into three groups: four experts as E, four third-year residents RY3, and five fourth-year residents as RY4.

5.1 Gaze Data Analysis

As shown in Table 2, the average fixation rate of experts is higher than novices, but the expert's average fixation duration is the lowest among all the groups. The average saccade rate and duration (ms) of experts is higher than the RY4 group. The expert's average time to the first fixation is the lowest among the three groups, the average fixation duration is less than that of RY4.

Table 2. Eye gaze metrics

	Expert	RY3	RY4
Avg. fixation rate	3.01	1.62	1.93
Avg. saccade rate	0.71	0.39	0.82
Avg. fixation duration (ms)	411.24	490.37	466.33
Avg. saccade duration (ms)	29.90	35.77	28.24
Avg. time to first fixation (ms)	50.00	155.00	450.00
Avg first fixation duration (ms)	1,039.50	499.80	1,269.25

Overall, experts have higher fixation rates compared to the novices and the majority of their fixations fell on the scope image. To investigate the fixation patterns of the expert and novice in the inner and outer circles AOIs of the scope, we considered 80% of the process of navigating from one landmark to another into finding the general area of the landmark and another 20% as zeroing in on the landmark. We found that during the 80% portion experts and novices both tended to fixate more on the outer circle in a ratio of roughly 2:1. During the 20% portion the experts fixated on the inner circle with a ratio of 2:1 while the novices continued to fixate on the outer circle with roughly the same ratio as before. This shows that the experts adjust their focus of attention to suit the portion of the navigation task, while the novices keep their focus primarily in only one area. This could be explained by the fact that an expert would be expected to know that they are getting close to a landmark whereas a novice might not.

5.2 Confusion

With a handful of reference anatomical regions within the joint, novices often miss the target landmark to diagnose during the procedure. Failure to recognize landmarks may result in disorientation and confusion as a student seeks to navigate through the shoulder joint. Since previous studies in user interfaces and intelligent tutoring had identified significant relationships between user confusion and metrics of pupil diameter and head movement, we sought to determine whether such relationships exist in this surgical domain as well.

As head movement metrics, we used the gyroscope and accelerometer data available from the Tobii eye tracker. Six novice participants (3 RY3, 3 RY4) reported a total of 14 confusion points while navigating and diagnosing at landmarks 1, 3, 6, 7, 8, and 12. The number of confusion points per landmark ranged from one to five with the highest frequency of three times reported at the landmarks 1, 6 and 8.

The follow-up interviews with the experts revealed that novices might get confused in landmark 1 due to a lack of recall of the background knowledge. At landmark 1, instead of looking for the void triangular space of the rotator interval between the subscapularis and glenoid and supraspinatus, the novices tended to look at the nearby structure. While in landmark 6, the novices need to locate the insertion of supraspinatus on the humerus. In the experts' opinion, the novices mostly focus on the tendon part, while all experts specifically focus on the tendon insertion point. This may be related to the level of knowledge of the pathological area on this tendon. The infraspinatus at landmark 8 is a tendon posterior to supraspinatus tendon. These tendons are blended together and have the same texture. Therefore, the location of infraspinatus can be identified only by understanding the exact location of infraspinatus (posterior half of these blended tendons).

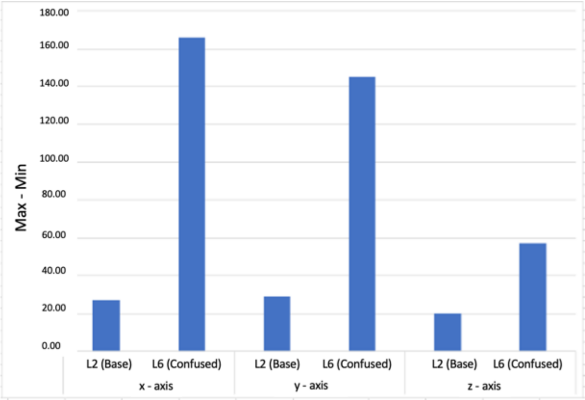
In terms of the time taken to complete the task, the experts completed the task with the least amount of time to diagnose at each landmark and had the least variation in task times. We observe that some landmarks require more time to navigate to and diagnose, particularly landmark 2 and 6 which are categorized as hard to diagnose. On average, the six novices who became confused took 1.5 times and 2 times longer than other novices in hard and easy landmarks, respectively.

The Percentage Change in Pupil Diameter (PCPD) is an objective measure of cognitive skills. Kruger et al. [25] studied PCPD as a measure of cognitive load and compared it with different cognitive load metrics including EEG, heart rate and blink rate when students were watching a recorded academic lecture, with and without subtitles. They found that higher cognitive loads were associated with higher PCPD values. We expect that the subject's cognitive load will increase while navigating the arthroscope in the landmarks where confusion was recorded. To determine that, we need a period of low cognitive load as a baseline. We used the period from the end of the previous landmark until the beginning of the current (confused) landmark as the baseline period since during that period the subject just is not actively navigating through the joint. The PCPD value was computed by subtracting the average diameter from the (confusion) landmark from the baseline diameter and divided it by the baseline diameter. From the six participants who became confused, the PCPD ranged from a minimum of 0.91% (left eye) and 0.97% (right eye) to a maximum of 1.22% (left eye) and 1.12% (right eye). On average, during the periods of confusion the pupil diameter changed by 1.02% in the left eye and 1.03% in the right eye relative to the baseline. The minimum values came from two novices at five different landmarks; all others had positive change in PCPD.

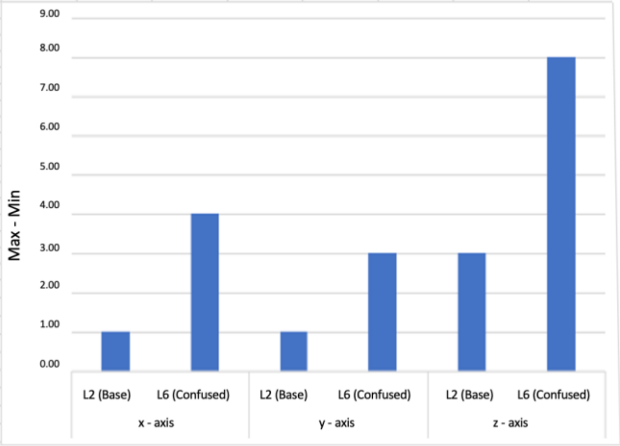
We investigated the head movement of the novice participants during the landmarks with confusion using the information from the gyroscope and accelerometer sensors of the eye tracker. Confusion was not reported in landmark 2 (L2: Biceps tendon & Biceps probe test) for any of the novices and hence it was considered as the baseline. We compared the head rotation and acceleration information between novices with and without reported confusion by computing the differences between the minimum and maximum values in x-, y- and z-axes. The differences are compared with the baselines using a paired t-test for each participant with confusion reported. The differences are significant in all three axes for head movements from the accelerometer as well as in y- and z-axes from the gyroscope sensors (p-value = 0.05). As shown in Table 3, the average differences between the two groups are substantial in the x-axis for head rotations and the z-axis for acceleration.

Table 3. Comparison in average differences in minimum and maximum values in three axes between novices with confusion reported, novices without confusion

Sensor		**x – axis**	**y – axis**	**z – axis**
Gyroscope	Novices with confusion	106.67	52.50	28.17
	Novices without confusion	22.93	30.12	13.08
Accelero-meter	Novices with confusion	2.00	2.50	4.67
	Novices without confusion	1.37	0.99	1.83



(a) The differences between minimum and maximum in



(b) The differences between minimum and maximum in three axes at the baseline landmark (L2) and the landmark with confusion (L6) from the accelerometer sensor

Figure 4. Comparison of head movements from gyroscope and accelerometer sensors between the baseline and the landmark with the confusion marker (L6) for a novice (RY4) in x, y, z axes

Figure 4 shows (a) the rotation from the gyroscope sensor and (b) shows the acceleration from the accelerometer along the x, y, z axes of a novice participant (RY4). As shown in the figure, this particular novice rotates the head along the x-axis and moves along the z-axis while navigating the arthroscope to the landmark 6 and performing the insertion (Subscapularis tendon and insertion).

6. CLASSIFICATION

In order to evaluate whether the eye-tracking metrics can be used to assess level of expertise in arthroscopic shoulder surgery, we sought to build models to classify participants as novice or expert. Due to the small size of the data set, we used leave-one (participant)-out to validate the classifiers. We applied Synthetic Minority Over-sampling Technique (SMOTE) repeatedly to the remaining twelve participants' gaze features. In each iteration, we randomly selected three novices and four experts, and generated one instance of novice with SMOTE and added it back to the novice data pool. The process was repeated until we reached a total of 100 novices. In the same manner, we generated expert data instances until we achieved 100, resulting in a balanced dataset with 200 instances. Features considered for the classification model included twelve gaze metrics extracted from the eye data including fixation and saccade rates for the whole procedure and three AOIs, average fixation and saccade rates, time to first fixation, and duration of first fixation. We selected the best five features using the information gain ratio (Table 4).

With the logistic regression model, we achieved a classification accuracy of 84%. The logistic regression model misclassified an expert and an R3 novice who have similar fixation rates (gaze points/sec) and an R3 novice with similar time to first fixation with an expert. The results show that in the domain of arthroscopic shoulder surgery, although the differences in eye-movement data are multidimensional, the two groups of participants can be classified with high accuracy by a simple model.

Table 4. Selected features with the information gain ratio

Feature	Gain Ratio	Min, Max, Mean
Time to first fixation (ms)	0.482	Expert = 25,75, 69.27 Novice = 75, 1075, 181.22
Fixation_rate	0.418	Expert = 2.01, 3.66, 3.37 Novice = 1.26, 2.67, 1.64
Fixation_rate_AOI_ Inner	0.381	Expert = 9.62, 15, 12.08 Novice = 5.98, 13.30, 10.03
Average_Fixation_ Duration_(ms)	0.358	Expert = 221.73, 705.98, 302.40 Novice = 354.21, 640.85, 489.45
Average_Saccade_ Duration_(ms)	0.306	Expert = 26.08, 35, 28.45 Novice = 25.47, 42.94, 34.24

7. CONCLUSION

The required skill set for arthroscopy is complex, due to an indirect view of the surgical site through the arthroscope, limited tactile feedback, and complex hand-eye-coordination. The operative time, probe path length, and number of movements are commonly utilized as surrogate markers for assessing skills. While previous studies have centered around the dexterous aspects of motor skills, we investigate cognitive aspects by studying the differences in perception between participants of differing experience [16]. During the arthroscopic surgery, surgeons rely primarily on visual information. Perception and attention are two separate but related processes. Initially attention occurs, and perception follows.

This study has shown that there are significant differences between expert and novice focus of attention during the arthroscopic navigation task both overall and during particular portions of navigation. We investigated a number of other questions such as the relationship between user confusion and metrics of pupil diameter and head movement, as well as whether the eye-tracking metrics can be used to classify the experts and novices. In contrast to the existing studies, the gaze measures in our study are collected with the cadaver specimens which provide the most realistic experience. We have demonstrated the potential of eye-tracking to provide reliable tools for automatic performance assessment in arthroscopic shoulder surgery. This leads us to the conclusion that gaze data carries important information about the skills of arthroscopic surgeons which could contribute to automated objective assessment. The future steps of this research include the development of an intelligent training system in the virtual reality environment that dynamically detects novice confusion and classifies surgeon's performance based on eye-movement data.

8. ACKNOWLEDGMENT

This work was partially supported through a fellowship from the Hanse-Wissenschaftskolleg Institute for Advanced Study (HWK), Delmenhorst, Germany to Su Yin for collaborative work with the University of Bremen, and through a study group grant from HWK to Haddawy. It was also partially supported through a grant from the Mahidol University Office of International Relations to the MIRU joint unit.

9. REFERENCES

[1] Geoffrey Tien, M. Stella Atkins, Bin Zheng, Colin Swindells. Measuring situation awareness of surgeons in laparoscopic training. ETRA '10: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications. 2010. pp. 149–152.

[2] Sodergren MH, Orihuela-Espina F, Clark J, Teare J, Yang G-Z, Darzi A. Evaluation of orientation strategies in laparoscopic cholecystectomy. Ann Surg. 2010;252: 1027–1036.

[3] Wilson M, McGrath J, Vine S, Brewer J, Defriend D, Masters R. Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. Surg Endosc. 2010;24: 2458–2464.

[4] Tien G, Zheng B, Atkins MS. Quantifying surgeons' vigilance during laparoscopic operations using eyegaze tracking. Stud Health Technol Inform. 2011;163: 658–662.

[5] Zheng B, Tien G, Atkins SM, Swindells C, Tanin H, Meneghetti A, et al. Surgeon's vigilance in the operating room. Am J Surg. 2011;201: 673–677.

[6] Atkins MS, Jiang X, Tien G, Zheng B. Saccadic delays on targets while watching videos. Proceedings of the Symposium on Eye Tracking Research and Applications. New York, NY, USA: Association for Computing Machinery; 2012. pp. 405–408.

[7] Sodergren MH, Orihuela-Espina F, Mountney P, Clark J, Teare J, Darzi A, et al. Orientation strategies in natural orifice translumenal endoscopic surgery. Ann Surg. 2011;254: 257–266.

[8] Eivazi S, Bednarik R, Tukiainen M, von und zu Fraunberg M, Leinonen V, Jääskeläinen JE. Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings. Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA ’12. New York, New York, USA: ACM Press; 2012. p. 377.

[9] Kocak E, Ober J, Berme N, Melvin WS. Eye motion parameters correlate with level of experience in video-assisted surgery: objective testing of three tasks. J Laparoendosc Adv Surg Tech A. 2005;15: 575–580.

[10] Ehlers J, Strauch C, Huckauf A. A view to a click: Pupil size changes as input command in eyes-only human-computer interaction. Int J Hum Comput Stud. 2018;119: 28–34.

[11] McCuaig J, Pearlstein M, Judd A. Detecting Learner Frustration: Towards Mainstream Use Cases. Intelligent Tutoring Systems. Springer Berlin Heidelberg; 2010. pp. 21–30.

[12] Tauro J, Pedowitz R. Arthroscopic Skills Training Modalities. In: Karahan M, Espregueira-Mendes J, Akan HK, editors. Motor Skills Training in Orthopedic Sports Medicine. Berlin, Heidelberg: Springer Berlin Heidelberg; 2017. pp. 53–64.

[13] Karahan M, Kerkhoffs GMMJ, Randelli P, Tuijthof GJM, editors. Effective Training of Arthroscopic Skills. Springer, Berlin, Heidelberg; 2015.

[14] Tofte JN, Westerlind BO, Martin KD, Guetschow BL, Uribe-Echevarria B, Rungprai C, et al. Knee, Shoulder, and Fundamentals of Arthroscopic Surgery Training: Validation of a Virtual Arthroscopy Simulator. Arthroscopy. 2017;33: 641–646.e3.

[15] Tuijthof GJM, van Sterkenburg MN, Sierevelt IN, van Oldenrijk J, Van Dijk CN, Kerkhoffs GMMJ. First validation of the PASSPORT training environment for arthroscopic skills. Knee Surg Sports Traumatol Arthrosc. 2010;18: 218–224.

[16] Pedowitz RA, Esch J, Snyder S. Evaluation of a virtual reality simulator for arthroscopy skills development. Arthroscopy. 2002;18: E29.

[17] Gomoll AH, Pappas G, Forsythe B, Warner JJP. Individual skill progression on a virtual reality simulator for shoulder arthroscopy: a 3-year follow-up study. Am J Sports Med. 2008;36: 1139–1142.

[18] Hoyle AC, Whelton C, Umaar R, Funk L. Validation of a Global Rating Scale for Shoulder Arthroscopy: A Pilot Study. Shoulder & Elbow. 2012;4: 16–21.

[19] Howells NR, Brinsden MD, Gill RS, Carr AJ, Rees JL. Motion analysis: a validated method for showing skill levels in arthroscopy. Arthroscopy. 2008;24: 335–342.

[20] Goyal S, Radi MA, Ramadan IK-A, Said HG. Arthroscopic skills assessment and use of box model for training in arthroscopic surgery using Sawbones - “FAST” workstation. SICOT J. 2016;2: 37.

[21] Tien T, Pucher PH, Sodergren MH, Sriskandarajah K, Yang G-Z, Darzi A. Differences in gaze behaviour of expert and junior surgeons performing open inguinal hernia repair. Surg Endosc. 2015;29: 405–413.

[22] Erridge S, Ashraf H, Purkayastha S, Darzi A, Sodergren MH. Comparison of gaze behaviour of trainee and experienced surgeons during laparoscopic gastric bypass. Br J Surg. 2018;105: 287–294.

[23] Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment | Proceedings of the 2004 symposium on Eye tracking research & applications. In Proceedings of the 2004 symposium on Eye tracking research & applications. pp. 41–48.

[24] Hermens F, Flin R, Ahmed I. Eye movements in surgery : A literature review. 2013;6. doi:10.16910/jemr.6.4.4

[25] Ahmidi N, Hager GD, Ishii L, Fichtinger G, Gallia GL, Ishii M. Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery. Med Image Comput Comput Assist Interv. 2010;13: 295–302.

[26] Richstone L, Schwartz MJ, Seideman C, Cadeddu J, Marshall S, Kavoussi LR. Eye metrics as an objective assessment of surgical skill. Ann Surg. 2010;252: 177–182.

[27] Eivazi S, Slupina M, Fuhl W, Afkari H, Hafez A, Kasneci E. Towards automatic skill evaluation in microsurgery. Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion. dl.acm.org; 2017. pp. 73–76.

[28] Rose K, Pedowitz R. Fundamental arthroscopic skill differentiation with virtual reality simulation. Arthroscopy. 2015;31: 299–305.

[29] Pachman M, Arguel A, Lockyer L, Kennedy G, Lodge J. Eye tracking and early detection of confusion in digital learning environments: Proof of concept. 1. 2016;32. doi:10.14742/ajet.3060

[30] DeLucia PR, Preddy D, Derby P, Tharanathan A, Putrevu S. Eye Movement Behavior During Confusion: Toward a Method. Proc Hum Fact Ergon Soc Annu Meet. 2014;58: 1300–1304.

[31] Lallé S, Conati C, Carenini G. Predicting Confusion in Information Visualization from Eye Tracking and Interaction Data. IJCAI. 2016. pp. 2529–2535.

[32] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement. arXiv [cs.CV]. 2018. Available: http://arxiv.org/abs/1804.02767

[33] Merali N, Veeramootoo D, Singh S. Eye-Tracking Technology in Surgical Training. Journal of Investigative Surgery. 2019. pp. 587–593. doi:10.1080/08941939.2017.1404663

[34] Atkins MS, Tien G, Khan RSA, Meneghetti A, Zheng B. What do surgeons see: capturing and synchronizing eye gaze for surgery applications. Surg Innov. 2013;20: 241–248.