Supervised Scene Illumination Control in Stereo Arthroscopes for Robot Assisted Minimally Invasive Surgery

Shahnewaz Ali, Yaqub Jonmohamadi, Yu Takeda, Jonathan Roberts, Ross Crawford, and Ajay K. Pandey

Abstract—Minimally invasive surgery (MIS) offers many advantages to patients but it also imposes limitations on surgeons ability, as no tactile or haptic feedback is available. From medical robotics perspective, visualizations issues specific to MIS such as limited field of view and the lack of automatic exposure control of the surgical area make it challenging when it comes to tracking tissue, tools and camera pose as well as in perceiving depth. Lighting plays an important role in 3D reconstruction and variations due to internal illumination conditions are known to degrade vital visual information. In this work, we describe a supervised adaptive light control system to solve some of the image visualization problems of MIS. Our proposed method is able to classify underexposed and over-exposed frames and adjust lighting condition automatically to enrich image quality. Our method uses support vector machines to classify different illumination conditions. Visual feedback is provided by gradient information to assess image quality and justify classifier decision. The output of this system has been tested against two cadaver knee experiment data with an overall accuracy of 97.75% for under-exposed and 89.11% for over-exposed classes. Hardware implementation of this classifier is expected to result in adaptive lighting for robot assisted surgery as well as in providing support to surgeons by freeing them from manual adjustments to lighting controls.

Index Terms— Intelligent light intensity control, support vector machine, knee arthroscopy, MIS, 3D reconstruction, robotic-assisted surgery.

1. INTRODUCTION

Contrary to open surgery, minimally invasive surgery (MIS) scenes are observed through a miniaturized endoscopic camera inserted into the body through a small incision and only a small region of the surgical scene becomes visible. Moreover, this procedure provides no haptic and tactile sensing hence, vision is the only source of information available to surgeons. In other words, surgeons have to operate on 3D surface based on 2D images projected on screen. This combined with a limited field of view adds further complexity to MIS. In recent years, many studies have aimed at increasing the accuracy for MIS [1]–[5]. It is usually well established that vision system plays an important role and robust vision is essential to avoid unintentional injury during MIS. Some recent studies suggest that simply extending the visualization of surgical area is able to significantly increase accuracy and reducing operating time. However, 3D reconstruction of the surgical scene is desired to further improve surgical outcomes [1], [6], [7]. Furthermore, 3D reconstruction is expected to have a great impact on the robotic-assisted surgery, such as tissue - tool tracking, and pose estimation.

Stereoscopic vision, simultaneous localization and mapping (SLAM) and photometric stereo are widely used to reconstruct 3D surfaces along with the alternative methods such as scanning technology [40]–[43], [46]. All these methodologies often encounter difficulties primarily due to the complex optical nature of anatomic constructs and also due to the lack over the lighting control [1], [46], [47]. It is also true for
that compromises overall performance the whole system [1].

Localization and pose estimation also encounter same problems vision related approach such as segmentation and registration, dark image. Apart from the 3D reconstruction process, other low light intensity that is often marked as an underexposed or lower value. Similar problem also arises when a scene contains can be a possible solution when light intensity is changed to a the whole 3D-reconstruction process. Partial reconstruction information, no features, and in situations like this stereo correspon-
dence search converges to an ambiguous result that lead to fail the whole 3D-reconstruction process. Partial reconstruction can be a possible solution when light intensity is changed to a lower value. Similar problem also arises when a scene contains low light intensity that is often marked as an underexposed or dark image. Apart from the 3D reconstruction process, other vision related approach such as segmentation and registration, localization and pose estimation also encounter same problems that compromises overall performance the whole system [1].

In conclusion, the light source or lighting condition has a great influence on the MIS visualization procedure. Surgeons or robotic platforms require a stable vision of the internal surgery scene. In knee arthroscopy, the presence of the diffuser (water), its anatomical structure, and camera motion are the most common factors that restrict uniform distribution of light intensity. Moreover, the control system must respond as fast as possible due to the camera motion. For a robust surgical vision this fundamental problem has to be solved.

II. PROBLEM FORMULATION

Considering anatomical structure, in some MIS procedures such as in knee arthroscopy, tissues further away from the arthrooscope often receive less light that causes formation of dark pixels (under exposed) whilst near objects encounter pixel saturation (over exposed). This situation is depicted in Fig.1. Although the overexposed and the underexposed events frequently occur in all MIS procedures, in current work this problem is analyzed in the context of the knee arthroscopy.

Owing to the presence of curved bone structure and narrow spaces, different areas of a surgical scene could receive different levels of illumination; which result in over and under exposures. Example of this scenario is shown in Fig.2. Moreover, some tissues reflect more light than others. A representative example is the Anterior Cruciate Ligament (ACL) that often causes pixel saturation. In addition, the distal parts of the anatomy, e.g. bone and soft tissue joints, create a more complex situation as depicted in Fig. 1-(a) where some parts remain over exposed and under exposed in the same frame. While the total scene intensity may remain the same, increasing the light intensity saturates parts that are closer to camera and decreasing light increases underexposed problem for the distant parts. Automatic methods to balance both over exposed and under exposed fractions with minimum frame loss are therefore highly desired.

Light intensity is controlled manually and surgeons have to adjust light intensity multiple times that interrupts with the surgery process flow. Clearly, manual approach to lighting control is not feasible for robotic-assisted surgery. Robotic imaging systems would require robust lighting control for depth estimation.

In this work, we present a supervised light control system that imitates like an endoscopic camera illumination system and provides real time feedback to improve stereo system accuracy and robustness by recommendations to adjust scene illumination condition to enhance scene features and image contrast even if the scene is considered at normal exposure. We have trained and validate our system with two distinct cadaver experiments.

III. RELATED WORK

At present, no work has been found so far that resolves the illumination problem automatically in arthroscopic or endoscopic imaging systems. A big volume of articles has been found to control camera exposure time by controlling the amount of light that falls on an image sensor. It has some similarities related to MIS visualization problems stated above. However, it is worth noting that most of these approaches are based on static images. In this context, camera response versus scene light intensity is evaluated through the equation described below.

\[ \sum = K \times L \times G \times S \]  

where, \( \sum \) is the sum value of imaging signal, \( K \) is a constant, \( L \) is the brightness of the object, \( G \) is the gain of the automatic gain control (AGC) circuit, \( S \) represents shutter speed of the camera. The automatic exposure bracketing based method is widely used to adjust exposure time [10], [12], [13]. Pour- reza et.al presented an automatic exposure selection method based on exposure bracketing [10]. It requires \( n \) number of image frames with different exposure values. In their method, they used two building blocks namely, scene analysis and exposure time. The clustering method was used to divide the image frame into three parts and exposure time was calculated through the camera response function. Radiometric based exposure control has been proposed by Kim et.al for outdoor scene analysis [11]. The radiometric camera response function is individual to the camera sensor [14], [15] that often generates camera specific solutions. However, this method also requires a set of still images. A statistical measure such as mean, mean sample value, etc. are used to intercept scene intensity distribution for normal exposure images. Impact of lighting on image formulation is significant, some literatures address light noise conditions with the recommendation to
the use of filters, artificial intelligence algorithms and new optoelectronics devices [44], [45].

Alternatively, neural networks have been studied to estimate chromaticity of image intensity [19]. Geoffroy et al. proposed a method that uses convolutional neural network (CNN) to estimate illumination intensity for High Dynamic Range images. In their work CNN was used to predict illumination for low dynamic range images. Usually a CNN network requires a large number of data sets for training. Additionally, deep learning requires a dedicated number of processing units to achieve real time performance [20].

Some other methods assess image quality at different exposure values [21]. Images having different exposure values can be achieved by transforming an image using gamma correction or tone mapping curve [21], [22]. In MIS context these approaches are not always valid. Camera motion, tissue movement, non-uniform illumination, surface curvature and distance between camera and surface limits the utility of almost all static image-based approaches. In endoscopic imaging process we need to minimize the frame loss, whereas exposure bracketing methods always have a number of redundant frames and camera navigation process has to remain in a fix position for multiple times for image stabilization. Moreover, most of the literature does not consider underwater or turbid environment encountered in MIS. In this context, scattering is a one of the most considerable problem where background light saturates image sensor making it difficult to visualize the object of interest [16].

**IV. METHODOLOGY**

We defined the illumination control problem as a classification problem here. Therefore, the aim of this work is to identify whether the classification and regression methods can infer requirements specific to the level of illumination desired in visualizing the surgical scene in terms of light distribution. The whole scene light conditions are grouped into three classes, as follows; I) overexposed, II) underexposed, and III) normal. Usually the first two scene categories are overlapped.

This method provides a solution that establishes robust illumination condition for the stereo and the monocular endoscope. We define ‘area of interest’ as an overlapped region between the stereo image pair captured by the stereo endoscope developed in our lab. Area of interest plays a significant role on stereo matching process which is the leading mechanism of stereo vision. It is more reasonable that the area of interest preserves sufficient image context.

Image pixel values are grouped into three groups and different threshold values were defined to serve this purpose. Usually threshold values are set by experiments [37]–[39]. In this work, eight-bit gray scale threshold values are defined empirically through the observations of intensity responses of internal anatomy using vision processing algorithm such as rank transformation and stereo matching followed by the human visibility limits. Three thresholds are defined as follows:

\[
I(x,y) = \begin{cases} 
\text{Overexposed, if} & 255 \geq I(x,y) > 226 \\
\text{Underexposed, if} & 101 \geq I(x,y) \geq 0 \\
\text{NormallyExposed, if} & 226 \geq I(x,y) > 101 
\end{cases}
\]  

Here, \(I(x,y)\) represents pixel intensity of the image coordinate \(x, y\). The maximum intensity of a saturated pixel in eight-bit gray scale level is 255. However, when a pixel is about to be saturated it has the same effect on the human visual system or on the vision processing methods as it is depicted by Fig. 3. The underexposed pixels also have the same effect. Research has found that in middle gray tone pixels receive optimal exposure value which is 128 in eight-bit gray scale level [39]. In order to identify optimal value of these two extremes, arthroscopy video sequences are segmented in to two slices where applicable. Slices are marked as an overexposed and underexposed. A range of gray level value then applied on the whole video sequences in order to find the optimal gray level intensity values for these two extremes. When a certain gray level intensity achieved a maximum segmentation
accuracy, then that gray level value is selected as a threshold value. In this work, intersection over union (IoU) is used to evaluate segmentation accuracy at different threshold values. Underexposed, overexposed and normal indices describe the overall image statistical information. Underexposed index expresses relative underexposed pixel amount over the whole image. Similarly, overexposed and normal exposed indices express relative amount of over and normally exposed values over the whole image. These three indices define the whole dataset in 2D space.

Support Vector Machine (SVM) is used with radial basis function (gaussian kernel). SVM was originally proposed by V. N. Vapnik [8]. SVM is widely used to address classification as well as regression problem as it is described in literature [49]. It is a binary classifier but in conjunction with other SVM it can perform multi classification problem. It classifies data with possible highest inter-class distance. Unlike multi-layers perception such as deep learning networks, SVM can be trained with limited dataset. It consumes fewer hardware resources compared to deep network, yet it provides efficient classification result in our problem domain. Instead of implementing multi-layer perception and conditional probability computation, SVM provides support vectors that are used to compute class distance from the hyperplane in run time. Moreover, Light control system has to operate in real time with minimal frame lost. Trained SVM has these strong points.

The most complex problem is to identify overexposed and underexposed scenes. To classify these total three lighting condition (underexposed, normal, overexposed), SVM with radial basis function (RBF) is applied that is stated in equation (3).

\[ K(x, y) = e^{-\gamma |x-y|^2} \]  

(3)

The advantage of the radial basis function is that for given sufficient data, it is able to define an efficient hyperplane discriminant among the nonlinear data set. It does not have saturation problem. The nature of lighting distributions is non-linear as depicted in Figure 6, we chose nonlinear kernel function for the proposed classifier. Comparative study shows radial basis function outperforms with nonlinear data among others such as polynomial kernel [50], [51]. Random subset from shuffled sample dataset is compared against different non-linear kernel before RBF kernel is selected for this approach. Mean accuracies of cross validation dataset are 0.9559448, 0.82038217, and 0.98152866 for the SVM kernel Sigmoid, Polynomial and Radial basis function respectively. Optimization takes places to establish classification and marginal error tolerance depicted in Figure 6.

Feedback system can justify classifier outcomes in real environment. Under sufficient lighting conditions image contains more context features, such as edges. Based on this, feedback system evaluates present image context and past image context. That can be achieved by evaluating image features such as edges, corner. In underwater environment, it has been observed that two consequent image frames show a subtle gradient change. Therefore, an empirical threshold is required that calibrates the system. Figure 4 presents our approach to define individual threshold values across all image conditions. For each frame different level of gray scale intensity values are used to achieve binary segmentation. The outcomes of each threshold value are then compared to the ground truth mask and then IoU (Jaccard index) metrices are calculated. In the next step, overexposed index, underexposed index and normally exposed index is calculated that are used by the SVM classifiers, as detailed below in Figure 5. 1st SVM classifier determines whether scene need more illumination or less. If scene requires more illumination second SVM classifier determines whether it actually needs more illumination or maintain the current illumination condition. Figure 6. shows hyperplane and optimized SVM (RBF) margin for over and under exposed conditions and output to identify light intensity changes between two classes.
Fig. 4. Set (I) represents overexposed threshold selection process. Left most images (a) represent the original images captured by the stereo endoscope, the next column (b) represents the segmented images and the column (c) represents corresponding mask for each stereo image pair. Over exposed pixels are marked with green color on the segmented images and those pixels are manually selected for each video frames. (d) represents average IoU curve and the optimal value is obtained at intensity level \( \sim 226 \) value is selected. Likewise, Set (II) represents underexposed threshold selection process. Left most images (a) represent the original images captured by the stereo endoscope, the next column (b) represents the segmented images and the column (c) represents corresponding mask for each stereo image pair. Underexposed pixels are marked with blue color on the segmented images and those pixels are manually nominated for each video frames. (d) represents average IoU curve and the optimal value is obtained when intensity level \( \sim 101 \) is selected.

Fig. 5. System overview of light control system. Stereo pair is cropped into areas of interest. In the next step, overexposed index, underexposed index and normally exposed index are calculated that are used by the SVM classifiers.

In this work, any scene context changes are referred as a motion. Pixel wise comparison is futile due to illumination variation, displacement and rotations. Additionally, frame similarity index such as structural similarity index (SSIM) may not provide a good result in this situation. Moreover, most of these approaches are based on gray level image statistics such as mean and variance. To compare two pixels under different illuminations, in this article census transformation in grayscale is used.

\section{V. Training}

For training purpose, two cadaveric knee experiment video data are used. Each frame is manually labeled. Scene is marked as an overexposed image if it contains lack of image features due to pixel saturation otherwise it is label as an underexposed image. If the scene is classified as an overexposed image, it is then further marked as a normally exposed image if it contains enough features and contains minimum image saturation otherwise overexposed image due to dark pixel amount. Additionally, if image does not contain saturated pixels and increasing light intensity enhances image quality then it is labeled as an underexposed image even though it may not contain enough dark pixels. This training data set is used to train the second SVM.

Further SVM regularization is recommended to increase classifier accuracy. Initially approx. 2950 frames are used to create training dataset. Optimization process is performed over RBF gamma and C parameter to fit optimal classification boundary curvature and the penalty value for misclassification. The optimal gamma value is 0.01 and optimal C value is 1 for the classifier up (light intensity increase) and no change (light intensity no change) classes. Similarly, for the classifier of the intensity change up (increase) and down (decrease), the method shows optimal performance on gamma equal to 0.001 and C equal to 100. The training accuracy are approx. 97.75% and 89.11%. Higher gamma value is discarded.
Fig. 6. Hyperplane and optimized SVM (RBF) margin with subset of our sample data to identify light intensity change. (a) left image represents classifier margin of two events; intensity increase and intensity decrease. (b) middle represents classifier margin of two events; intensity increase and intensity no change classes. (c) right image expresses the example optimization process of SVM parameter, Gamma and C. Different values of gamma and c are used to select optimal values of them in order to make the class separable while avoiding marginal error.

to avoid data overfitting problem. During the optimization process minimization of classification error is more considered rather than the marginal error.

A. Data Collection

During our cadaveric experiments, stereo video sequences of knee Arthroscopy are collected. Stereo endoscope has manual light control unit. During arthroscopy surgeon changes light intensity several times in different part of the knee cavities. Those video sequences are used to validate model. Moreover, one set of dedicated video sequences has been recorded that contains a linear light intensity changes from low to high in some position during the camera trajectory.

VI. DESIGN OF THE ENDOSCOPIC STEREO CAMERA

Figure 7 presents the clinical context and a representative location of arthroscope within the knee cavity during MIS. Details of this stereo camera prototype developed in our lab is shown in Figure 8. A commercially available traditional arthroscope is also shown in this figure. Our stereo camera prototype is comprised of following elements;

- a pair of muC103A cameras together with their C8262 UVC interface modules,
- two white LED (T0402W) for illumination,
- 3D printed camera head for mounting cameras and the LED,
- 3D printed box for containing the wires and the circuits, and
- insertion tube.

muC103A is a CMOS technology camera sensor from OmniVision which can stream video at 30 fps with 400×400-pixel resolution and the field of view in 120 deg. The tip diameter of 1.52 mm makes this camera ideal for the endoscopic applications. The two cameras were mounted on the 3D printed head manually, hence some degree of misalignment between the two cameras was unavoidable. The T0402W LED consumes less than 20 mA at 3.3 V and is 1 mm wide. The baseline for the cameras was set 1.52 mm (distance between the optical center of two cameras is 1.52 mm).

VII. RESULT

Total 3,350 video frames are selected from our cadaveric knee experiments. Distorted and contaminated frames are discarded. During this training phase we select 1/3 of frames from each video file that contains different levels of difficulties such as shadow, bone joint and cavity, close contact between the camera and tissue etc. We use those frames for cross validation purposes. Table I shows the all outcomes during our training phase. During the test phase, fully trained classifier is tested against different arthroscopy video data along with cross validation data set. In some video frames different stereo camera is used. All the SVM graphs are created using scikit [29].

This methodology is compared against the camera exposure control methods in order find a benchmark. Camera exposure control system limits the amount of light that falls on digital camera sensor in order to achieve good quality of images. In this evaluation, increase of exposure time simulated as an increase of light intensity, decrease of exposure time as a decrease of light intensity, and no effect of exposure time interpreted as a no change of light intensity. Compared methods are mainly based on histogram analysis and scene quality analysis. However, compared to natural images MIS images has a number of limitations such as camera locomotion, distance between the camera surface, and no uniform light distribution etc.

Algorithm 1 is based on Gradient Information [21], [34]. Gradient information is computed as follows;

\[
\bar{m} = \begin{cases} 
\frac{1}{n} \log (\lambda (m_i - \delta) + 1) & \text{for } m_i \geq \delta \\
0 & \text{for } m_i < \delta \\
s.t. & N = \log (\lambda (1 - \delta) + 1)
\end{cases}
\]
Fig. 7. Stereo endoscope along with its illumination sources are used to scan the whole knee anatomy. The surface light distributions and the ratio of it to the different classes at all possible intensity levels are extracted to construct a 2D data set. This data set is used to train our classifier.

Fig. 8. The stereo endoscope prototype (a) together with the close-up view at the tip (b) and the 3D design (c) and the muC103A camera (d). The potentiometer shown by the red arrow is used to adjust the LED intensity. The endoscope circuits and extra wiring are contained inside the black box and connected to the computer using two USB cables.

In order to calculated gradient information of images having different exposure values, gamma function is used as proposed by the article. The exposure value is selected based on the following criteria;

$$\arg\max_{\gamma} M \left( I_{in}^\gamma \right)$$  \hspace{1cm} (5)

where $\gamma$ is the gradient information and $M$ mimic exposure time value. The aim is of this method is to maximize gradient information.

Algorithm 2 is based on third moment of image intensity histogram [35]. Third moment is calculated as follows proposed by the authors as follows;

$$\mu = \sum_{i=0}^{4} (i+1) \cdot x_i / \sum_{i=0}^{4} x_i$$  \hspace{1cm} (8)

where, $x_i$ is the sum of the values in region $i$ and $i$ denotes the region of the histogram. According their article, the image is correctly exposed when $\mu \approx 2.5$ [36].

Confusion matrices for lighting control in stereo frames and monocular frames is shown in Figure 9. Model occurrences are our manually class label and predicted occurrences are endoscopic illumination controller outcomes. For class intensity increase, according to the normalized confusion matrix it shows 97% accuracy. Most significant error in that case is 0.009% that is decrease light intensity. Similarly, decrease intensity event achieved 97.8% classification accuracy and no change in light intensity event achieved 55.6% classification accuracy. No change light intensity is hard in stereo situation. In our test stereo pair image encountered different illumination in most of situation caused by shadow, surface curvature of internal anatomy etc. Lighting control outcomes for over-exposed, under-exposed and normal exposed classes are presented in Figure 11. Upward arrow, downward arrow and dotted line annotations to stereo images indicate controller output for each case.

Figure 12. compares the outcome of methods evaluated. Our proposed method received average error of 7.37% and 4.56 % for stereo endoscope and monocular endoscope respectively. All the methods fail to identify no change event for monocular endoscope and except Algorithm 1. It receives 32.7 % error to identify no change error. Where the proposed method is able to detect no change event with 71.27% average accuracy (16.36% in error in stereo and 41.1% error in monocular). Considering overall accuracy to control illumination event the
Fig. 9. (a) Normalized Confusion Matrix for Stereo camera light control system. Model occurrences are our manually class labels and predicted occurrences are endoscopic illumination controller outcomes for each scene. All the diagonals are true positive. (b) Mono camera light control Confusion Matrix. All the diagonals are true positive. Increase light intensity classified by our system with approx. 95% accuracy during the test. Similarly, in approx. 97% situations decrease light intensity events are classified with true positive. Intensity no change events accuracy is about 82%.

Fig. 10. It compares the performance of our method over pure image post processing-based approach. Column (a) represents the accuracy of the depth map when image is degraded by overexposed pixels. The achieved PSNR and SSIM are 2.73 and 0.29 respectively where well exposed image achieved (b) achieved 5.37 and 0.58. Overexposed image then corrected using correction of brightness (c), decreased contrast (d), increased contrast (e) and through the correction of histogram equalizer (f). The achieved PSNR and SSIM are shows image corrected through the control of illumination provides better solution.

proposed method outperforms state-of-the-art exposure control methods.

VIII. CONCLUSION

We have explored the feasibility of support vector machine classifier for endoscopic camera light intensity control problem in real time. Our results offer good prospects in inferring challenges imposed in surgical scene visualization into a scene classification problem. Confusion matrices represents the overall performance of our implemented illumination controller. However, accuracy decreases for the ‘no change’ class category. In knee arthroscopy this event occurs rarely hence, very few training data was available for train this category. Additionally, in stereo vision system, two viewpoints receive different level of illumination distribution mostly due to the anatomical structure hence, this specific class category encounters more error. A set of training video sequences under slowly varying light intensity at every possible intensity value is recommended for this approach for further improvement.

Image patch-based similarity index under radiometric, rotation and illumination changes are expected to further enhance feedback in real time. It might be useful to combine patch similarity or dissimilarity index along with the degree of changes...
Fig. 11. Light control outcomes for binocular stereo camera on knee arthroscopy video sequences. Each image block contains one stereo image pair, basically left image from left camera and right image from right camera. Image annotations with upward arrow indicates scene requires more light intensity, similarly downward arrows indicate scene contain overexposed portion and need low light intensity and double dashed sign indicates scene meets sufficient illumination condition.

Fig. 12. Evaluation of the illumination control methods.

provided by content indexing methods. Maximum similarity regions with positive improvement of those reason can infer an intensity event as a true decision, as investigated in this work.

REFERENCES

[1] J. Song, J. Wang, L. Zhao, S. Huang, and G. Dissanayake, “MIS-SLAM: Real-time large scale dense deformable SLAM system in minimal invasive surgery based on heterogeneous computing,” 2018, arXiv:1803.02009. [Online]. Available: http://arxiv.org/abs/1803.02009

[2] X. Du et al., “Combined 2D and 3D tracking of surgical instruments for minimally invasive and robotic-assisted surgery,” Int. J. Comput. Assist. Radiol. Surgery, vol. 11, no. 6, pp. 1109–1119, Jun. 2016.

[3] R. M. Sousa, M. Wang, P. Santos, and F. Morgado-Dias, “NanEye—An endoscopy sensor with 3-D image synchronization,” IEEE Sensors J., vol. 17, no. 3, pp. 623–631, Feb. 2017.

[4] M. Hu, G. Penney, P. Edwards, M. Figl, and D. J. Hawkes, “3D reconstruction of internal organ surfaces for minimal invasive surgery,” in Medical Image Computing and Computer-Assisted Intervention–MICCAI (Lecture Notes in Computer Science), vol. 4791, N. Ayache, S. Ourselin, and A. Maeder, Eds. Berlin, Germany: Springer, 2007.

[5] T. Collins and A. Bartoli, “3D reconstruction in laparoscopy with close-range photometric stereo,” in Medical Image Computing and Computer-Assisted Intervention–MICCAI (Lecture Notes in Computer Science), vol. 7511, N. Ayache, H. Delingette, P. Golland, K. Mori, Eds. Berlin, Germany: Springer, 2012.

[6] S. M. D. Sørensen et al., “Three-dimensional versus two-dimensional vision in laparoscopy: A systematic review,” Surgical Endoscopy, vol. 30, pp. 147–153, Mar. 2016.

[7] V. Velayutham et al., “3D visualization reduces operating time when compared to high-definition 2D in laparoscopic liver resection: A case-matched study,” Surgical Endoscopy, vol. 30, pp. 147–153, Mar. 2016.

[8] V. N. Vapnik, “An overview of statistical learning theory,” IEEE Trans. Neural Netw., vol. 10, no. 5, pp. 988–999, Sep. 1999.

[9] T. Kuno, H. Sugiuara, and N. Matoba, “A new automatic exposure system for digital still cameras,” IEEE Trans. Consum. Electron., vol. 44, no. 1, pp. 192–199, Feb. 1998, doi: 10.1109/30.663747.

[10] R. Pourreza-Shahri and N. Kehtarnavaz, “Exposure bracketing via automatic exposure selection,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Quebec City, QC, Canada, Sep. 2015, pp. 320–323, doi: 10.1109/ICIP.2015.7350812.

[11] S. Joo Kim, J.-M. Frahm, and M. Pollefeys, “Radiometric calibration with illumination change for outdoor scene analysis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Anchorage, AK, USA, Jun. 2008, pp. 1–8, doi: 10.1109/CVPR.2008.4587648.
Shahnewaz Ali is a Ph.D. Researcher with the Queensland University of Technology, Australia, and undertaking his research on medical robotics. His current research interest addresses visualization challenges related with robotic vision within the body, more specifically within knee-joints that includes perception and awareness as well as technological developments required for robotic imaging system for MIS imaging systems.

Yaqub Jonmohamadi received the Ph.D. degree in neuroimaging, with focus on EEG analysis, from the University of Otago, New Zealand, in 2014. He started his first postdoctoral fellowship in multimodal neuroimaging with the University of Auckland, New Zealand, in 2015. He is currently undertaking his second postdoctoral fellowship in medical robotics and medical imaging with the Queensland University of Technology, Australia.

Yu Takeda received the Ph.D. degree from the Hyogo College of Medicine in 2018. He studied medicine from the Hyogo College of Medicine, Japan, from 2003 to 2009. He is currently a Visiting Researcher with the Queensland University of Technology, Australia, in the field of robot-assisted and autonomous surgery. He is also an Orthopedic Surgeon.

Jonathan Roberts is a Professor of Robotics with the Queensland University of Technology (QUT). He is also a Chief Investigator with the Australian Centre for Robotic Vision (ACRV). His main research interests include field robotics, medical robotics, performance robotics, and more recently design robotics.

Ross Crawford holds the position of the Chair in Orthopedic Research with the Queensland University of Technology. He is an internationally recognized Orthopedic Surgeon and an Expert in robotic-assisted knee and hip replacement surgery.

Ajay K. Pandey is a Senior Lecturer of Robotics and Autonomous Systems with the School of Electrical Engineering and Robotics. His research interests include the application of photonics and molecular electronics principles for designing intelligent vision systems and tactile sensors for medicine and robotics.