https://doi.org/10.2352/ISSN.2470-1173.2020.6.IRIACV-016

Locating Mechanical Switches Using RGB-D Sensor Mounted on a Disaster Response Robot

Takuya Kanda* † *, Kazuya Miyakawa* ‡ *, Jeonghwang Hayashi* ‡ *Jun Ohya* ‡ *, Hiroyuki Ogata* § *, Kenji Hashimoto*, Xiao Sun* ‡ *,
***Takashi Matsuzawa* ‡ *, Hiroshi Naito* ‡ *, Atsuo Takanishi* ‡**

† ***Department of** Integrative Bioscience and Biomedical Engineering**, Waseda University, Tokyo, Japan***
‡ ***Department of Modern Mechanical Engineering, Waseda University, Tokyo, Japan***
§ ***Department of System Design Engineering, Seikei Univeristy, Tokyo, Japan***
**** Department of Mechanical Engineering Informatics, Meiji University, Kanagawa, Japan***

Abstract

To achieve one of the tasks required for disaster response robots, this paper proposes a method for locating 3D structured switches' points to be pressed by the robot in disaster sites using RGBD images acquired by Kinect sensor attached to our disaster response robot. Our method consists of the following five steps: 1)Obtain RGB and depth images using an RGB-D sensor. 2) Detect the bounding box of switch area from the RGB image using YOLOv3. 3)Generate 3D point cloud data of the target switch by combining the bounding box and the depth image.4)Detect the center position of the switch button from the RGB image in the bounding box using Convolutional Neural Network (CNN). 5)Estimate the center of the button's face in real space from the detection result in step 4) and the 3D point cloud data generated in step3) In the experiment, the proposed method is applied to two types of 3D structured switch boxes to evaluate the effectiveness. The results show that our proposed method can locate the switch button accurately enough for the robot operation.

1. Introduction

In recent years, large-scale disasters such as accidents at the Fukushima Daiichi Nuclear Power Plant have occurred. In such disaster sites, it is necessary to investigate the damage status and to perform emergent restoration works. However, such extreme sites are very dangerous for humans to work. Therefore, demand for disaster response robots that can work on behalf of humans in the extreme sites is increasing. To meet to this demand, we are developing a disaster response robot called "WAseda REsCuer-No.1" (WAREC-1) [1], which is a quadruped robot that can work in such extreme environments.

Among many restoration work capacities required for disaster response robots, switch operation is one of the most important functions. The switch operation is also adopted as one of the tasks in the Darpa Robotics Challenge (DRC) [2], which is a competition for disaster response robots held in the United States. The switch operation is also listed as one of the work capabilities required for disaster response robots at Council on Competitiveness-Nippon [3,4] in Japan.

At present, WAREC-1 can be remotely operated so as to turn on/off a switch like many other robots participated in DRC. Such a remote operation needs tele-communications between the operator and the robot. However, in extreme environments, tele-communication is not always fully available, because tele-communication disruptions could happen frequently. In such a case, it is difficult to remote-control the restoration work. In DRC, there is also a task that simulates environments in which tele-communication is not available in a good state by intermittently interrupting tele-communication between the operator and the robot. In such a situation, it is necessary for a robot to have a certain level of autonomy for carrying out restoration works. Therefore, this research focuses on estimating the switch position necessary for operating switches autonomously using the RGB-D sensor mounted on WAREC-1.

Related studies of autonomous switch operations can be found in the field of service robots. In these studies, the robot presses general elevator buttons autonomously to navigate people across floors. These studies' methods for detecting switches are to detect the button area as a bounding box using Convolutional Neural Network (CNN) [5], or to combine Faster RCNN and OCR [6]. Elevator switches, which are addressed by these studies, are usually flat. To press the flat switch, the robot just needs to move its hand to the center of the detected bounding box. Therefore, in the bounding box these studies do not need to locate the position that should be pressed with the robot hand. In contrast, switches to be pressed by a disaster response robot are not only flat switches such as elevator switches, but also mechanical switches having 3D (3-dimentional) structures. In case of such a mechanical switch, depending on the view point, the point to be pressed by the robot is not necessarily the center of the bounding box. Thus, in addition to detecting the bounding box of a switch area, a method for locating the mechanical switch's point to be pressed needs to be achieved

This paper proposes a method that detects a 3D-structured switch and calculate the position to be pressed using the information obtained by the RGB-D sensor attached to WAREC-1. This paper focuses on pushbutton switches and tactile switches among switches.

2. Proposed Method

RGB and depth images acquired by an RGB-D sensor are used to calculate the switch position, which is necessary for an autonomous switch operation. An overview of our proposed method is shown in Fig. 1.

The proposed method consists of the following five steps:

1. An RGB image and a depth image that could contain the target switch are acquired using KINECT v2, which is an RGB-D sensor mounted on the body of WAREC-1.
2. The bounding box of the switch in the RGB image is detected using YOLOv3 [7], which is one of the object detection methods based on deep learning.
3. The center of the button's face, which is considered as the point to be pressed, is estimated within the bounding box using a CNN.
4. From the information on the bounding box and the depth image obtained from KINECT v2, depth information in the bounding box is extracted.
5. The shape of the switch is reconstructed as a 3D point cloud from the depth information, and after the noise of the 3D point cloud is removed, the center of the button's face is located using CNN.

2.1. Obtaining RGB-D information using KINECT v2

To calculate the switch position, RGB image and depth image are acquired by KINECT v2. KINECT v2 can acquire both RGB and depth images simultaneously with a resolution of 512 × 424 pixels. The effective measuring range of depth information of KINECT v2 is 0.5 [m]-8.0 [m]. KINECT v2's view angle is 70.6 x 60 degrees.

The RGB image is used to detect the switch. The depth image is used to calculate the switch position in 3D.

2.2. Detecting switch using YOLOv3

Since information such as the shape of the target switch in the disaster site is in general unknown in advance, versatility of the process for detecting different switches is required. Therefore, YOLOv3 is used to detect the switch area from the RGB image, as described in Section 2.1. The area, which could contain a switch is outputted as a bounding box. The RGB images of different switches to train YOLOv3 were collected beforehand from the Internet. Among the switches, we collected the pushbutton-type and tactile-type switch images that are the targets of the method proposed by this paper and used them as training data. The encoder used for feature extraction is Darknet53.

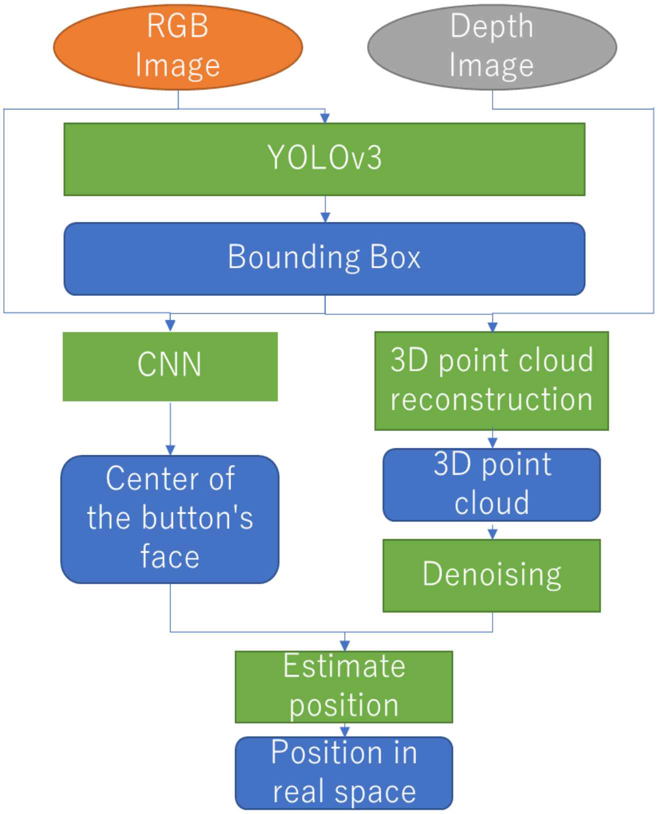
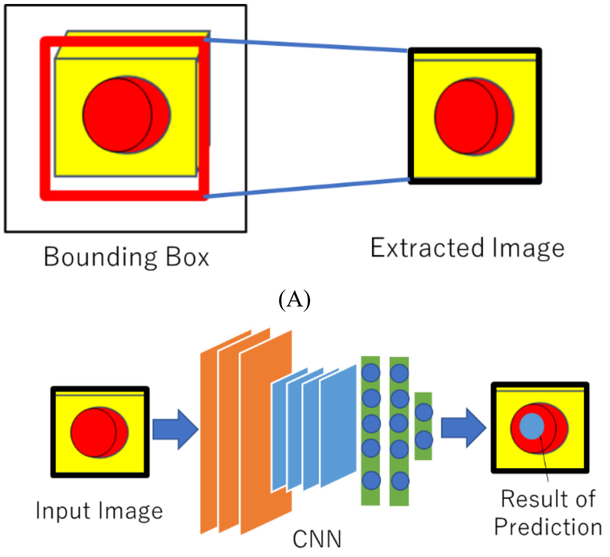


Fig. 1 Overview of the Proposed Method

2.3. Detecting the center of button's face using CNN

CNN is used to estimate the coordinates of the center of the button's face in the RGB image of the bounding box detected as a switch. As the target switch in this paper is not a flat switch, the center of the bounding box is not always a suitable position to press. Therefore, it is necessary to locate the center of the button's face in the bounding box. Since there are many variations in the shape of switches, we utilize CNN to detect the center of the button's face from the bounding box. Figure 2 shows an overview of detecting the center position of the button's face. Fig. 2 (A) shows the process of extracting the data to be inputted to the CNN. The input of the CNN is the image of the bounding box obtained by YOLOv3. Fig. 2 (B) shows the process of estimating the center position of the switch button's face using CNN. Concerning the training data, the part annotated as the switch area was cut out from the switch image data used for training the YOLOv3. Figure 3 shows the structure of the CNN.



(A)

(B)

Fig. .2 Flow of Detecting Center of Switch

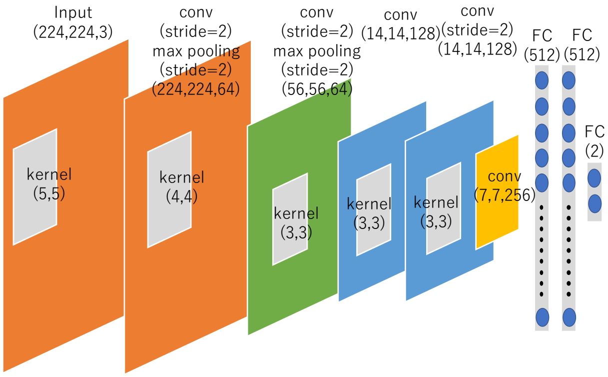
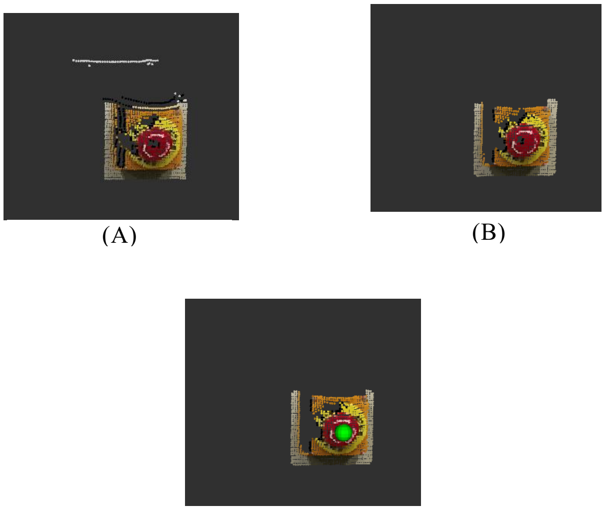


Fig. 3 Structure of CNN

2.4. Reconstructing 3D point cloud information of the switch area by overlaying RGB and depth images

The switch area in the depth image is estimated by overlaying the bounding box in the RGB image on the depth image. 3D point cloud of the switch is reconstructed based on the depth information of the area.

The coordinates of the switch region in the real space can be obtained by converting the depth image information into the 3D point cloud information.



(A)

(B)

(C)

Fig. 4 Preprocessing for Points Cloud

2.5. Detecting the position to be pressed using 3D point cloud processing

The 3D point cloud data obtained in Section 2.4 is processed using Point Cloud Library [8]. Areas that do not correspond to the switch are removed by applying a pass through filter in advance. Then, the influence of measurement errors of the depth sensor is reduced by removing outlier points using a statistical outlier filter. Thereafter, the coordinates of the center of the button's face in real space obtained in Section 2.3 are calculated.

Figure 4 shows an example of the results of each process. Fig. 4 (A) shows the point cloud information after 3D point cloud reconstructed from the depth image, (B) shows the point cloud after removing outlier points, and (C) shows the result of calculating the center coordinates of the switch operation part.

3. Experiment

3.1. Experimental environment

We carried out an experiment to evaluate the proposed method. The switch box and KINECT v2 were geometrically arranged as illustrated in Fig. 5. The distance *d* between the switch and Kinect and the orientation θ are to be estimated. Two types of switch boxes are used in the experiment. One is the switch box with a mushroom-shaped pushbutton switch (*i. e.* an emergency switch) [9], and the other is the switch box [10] with two rectangular ON and OFF buttons. As a whole, three buttons on the two switch boxes were used for the experiments.

Figure 6 shows a photo and the size of each switch box. The distance *d* between the center of the button's face and KINECT v2 was set to 60 [cm] and 80 [cm], and the angle θ was set to 0 [deg], -45 [deg], 45 [deg]. The position was measured 100 times for each switch to evaluate the measurement results.

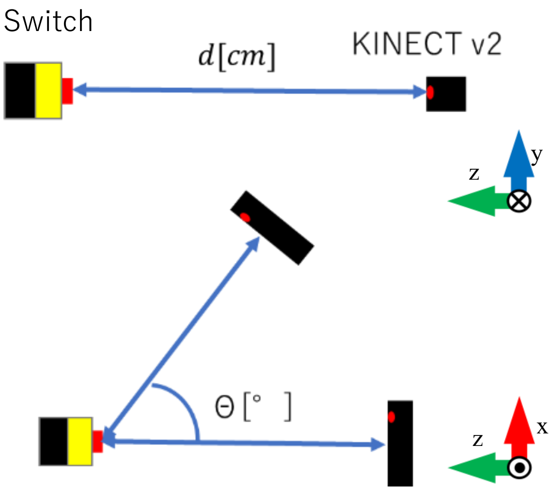


Fig. 5 Experiment Environment

(A)Emergency Switch (B) Power Open / Close Switches

Fig. 6 Target Switches

3.2. *Experimental result*

Tables 1 to 6 show the average and maximum errors of each coordinate of the detection result of the center of the button's face obtained in each experiment.

Table 1 Emergency Switch 60[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	2.6	3.1
	Y	1.1	2.1
	Z	0.5	1.2
0[deg]	X	1.0	1.5
	Y	0.6	1.1
	Z	1.0	1.2
45[deg]	X	0.5	1.1
	Y	0.2	0.5
	Z	0.5	1.2

Table 2 Emergency Switch 80[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	0.7	1.1
	Y	0.8	1.1
	Z	0.5	1.2
0[deg]	X	0.6	3.1
	Y	0.8	0.9
	Z	0.8	1.2
45[deg]	X	0.3	1.1
	Y	0.4	0.7
	Z	0.5	2.2

Table 3 Power Open Switch 60[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	0.3	0.7
	Y	0.7	2.5
	Z	1.0	1.6
0[deg]	X	0.0	5.9
	Y	0.4	8.8
	Z	0.2	0.8
45[deg]	X	0.7	0.9
	Y	0.6	1.8
	Z	1.4	1.8

Table 4 Power Open Switch 80[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	0.2	0.4
	Y	0.5	1.1
	Z	0.8	1.2
0[deg]	X	0.6	0.98
	Y	0.9	1.6
	Z	0.7	1.4
45[deg]	X	0.0	0.5
	Y	0.8	1.2
	Z	0.2	0.3

Table 5 Power Close Switch 60[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	0.2	0.6
	Y	0.3	0.6
	Z	0.2	0.9
0[deg]	X	0.3	0.8
	Y	0.8	1.7
	Z	0.2	0.9
45[deg]	X	0.4	0.7
	Y	0.7	1.0
	Z	0.2	0.3

Table 6 Power Close Switch 80[cm]

	Axis	Mean error [cm]	Maximum error [cm]
-45[deg]	X	0.2	0.4
	Y	0.0	1.3
	Z	1.1	1.4
0[deg]	X	0.6	01.3
	Y	0.6	1.2
	Z	0.7	1.3
45[deg]	X	0.5	0.8
	Y	0.8	1.5
	Z	0.3	1.0

4. Discussion

4.1. Detection of the switch operation center using YOLO v3 and CNN

Figure 6 shows the switch detection results in the experiment. Though there was no significant error from the switch operation part in the detection by YOLO v3, there were some cases where the entire switch operation area is not within the detected bounding box. The reason for this is that there may be cases where CNN cannot obtain enough feature information from the image to detect the center of the button’s face accurately. In addition, since the size of the switch is relatively small in the RGB image, it can be considered that when the center of the button’s face is estimated using the CNN, the pixel resolution becomes rough and sufficient features cannot be obtained.

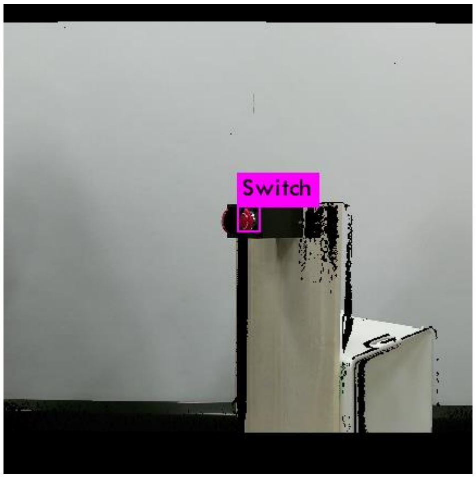


Fig. 6 Switch Detection Result using YOLO v3

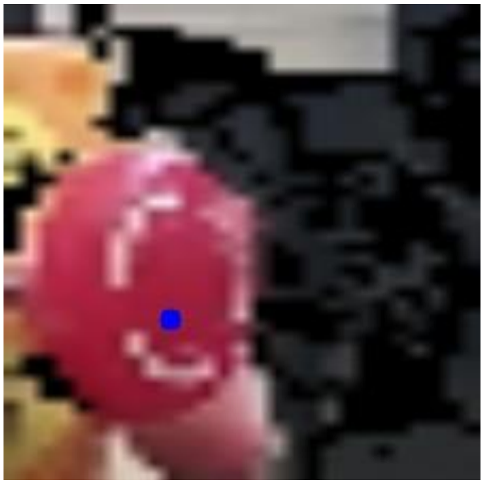


Fig. 7 Center of Switch Operating Part Detection Result using CNN

Furthermore, the center of the button’s face can be accurately detected even when the image of the mechanical switch is taken from an oblique direction, and the center of the operation unit is not at the center of the bounding box (Fig. 7).

4.2. 3D point cloud processing

In this paper, the center of the button’s face is estimated based on only one point information. Therefore, the measurement result is greatly affected by the noise of the depth sensor.

Since the outlier removal filter is not robust enough to remove the noise, it is necessary to improve a method that is more robust to measurement errors of depth information by using the surrounding point cloud information rather than calculating based on only one point cloud information.

5. Conclusion

This paper has proposed a method for locating the switch position information necessary for autonomous switch operation using the RGB-D sensor attached to the disaster response robot and evaluated the proposed method. The proposed method can be applied not only to a planar switch but also to three-dimensionally structured switches such as a mechanical switch. The proposed method contributes to operating the three-dimensional switch, which is required for disaster response robots. In the experiment, we measured the position of two types of mechanical switches by setting the distance between the switch and RGB-D sensor to 60 [cm] and 80 [cm] and the orientation to 0, -45 and 45 [deg]. From the experimental results, it turned out that it is possible to estimate the distance within mean error of 2.6[cm] in the X direction, 1.1[cm] in the Y direction, and 1.4[cm] in the Z direction.

But, to operate a switch, the information about the direction of the button's face is also required. In this research, we proposed a method that assumes that both RGB image information and depth image information can be obtained. However, there is a situation where RGB information cannot be obtained because of insufficient lighting environments at the disaster site, or it is difficult to obtain depth information using infrared rays due to smoke or dust. Therefore, in the future, we would like to develop a method that can autonomously perform restoration work even in an environment where such RGB image information or depth image information cannot be obtained sufficiently.

Acknowledgement

The authors acknowledge the support of Mr. Li Qi, Mr. Kota Umebayashi, Ms. Reina Yoshizaki of Waseda University and Mr. Keishi Nishikawa of Mitsubishi Electric Corporation.

References

[1] Kenji Hashimoto, Shunsuke Kimura, Nobuaki Sakai, Shinya Hamamoto, Ayanori Koizumi, Xiao Sun, Takashi Matsuzawa, Tomotaka Teramachi, Yuki Yoshida, Asaki Imai, Kengo Kumagai, Takanobu Matsubara, Koki

Yamagushi, Gan Ma, Astuo Takanishi, "WAREC-1 - a Four-Limbed Robot Having High Locomotion Ability with Versatility in Locomotion Styles," 2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR2017), pp. 172-178, 2017.

[2] Adam Norton, Willard Ober, Lisa Baraniecki, Eric McCann, Jean Scholtz, David Shane, Anna Skinner, Robert Watson, and Holly Yanco, "Analysis of human–robot interaction at the DARPA Robotics Challenge Finals." The International Journal of Robotics Research, Vol. 36, Issue 5-7, pp. 483-513, 2017.

[3] Council on CompetitivenessNippon (COCN).: Establishment plan for a disaster response robot center. The 2013 Report of Council on Competitiveness Nippon (COCN), 2013.

[4] Asama, H., Tadokoro, S., Setoya, H.: In: COCN (Council on Competitiveness-Nippon) Project on Robot Technology Development and Management for Disaster Response. IEEE Region 10 Humanitarian Technology Conference 2013, 2013.

[5] Zijian. Dong, Delong. Zhu and Max Q. -H. Meng, "An autonomous elevator button recognition system based on convolutional neural networks," 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 2533-2539 , 2017.

[6] Zhu, Delong, Tingguang Li, Daniel Ho, Tong Zhou and Max Q. -H. Meng. “A Novel OCR-RCNN for Elevator Button Recognition.” 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3626-3631, 2018.

[7] Joseph Redmon, Ali Farhadi “YOLOv3: An Incremental Improvement.” arXiv:1804.02767, 2018.

[8] Point Cloud Library (last access 30th September 2019) http://pointclouds.org/documentation/tutorials/passthrough.php.

[9] UXCell corporation, Red Sign Mushroom Emergency Stop Push Button Switch Station 1 NO 1 NC 10A 660V, http://www.uxcell.com/self-locking-emergency-stop-red-mushroom-switch-push-button-station-p-156304.html, (last access 30th September 2019).

[10] PATLITE corporation, BS210B2, https://www.patlite.co.jp/product/detail1000019260.html, (last access 30th September 2019).

Author Biography

Takuya Kanda received B.S. degree from the Department of Modern Mechanical Engineering, the School of Creative Science and Engineering, Waseda University, Japan, in 2018. Currently, he belongs to the Department of Integrative Bioscience and Biomedical Engineering, the Graduate School of Advanced Science and Engineering, Waseda University as an M.S. student. His research areas include deep learning, image processing, robot vision.

Kazuya Miyakawa received B.S. degree from the Department of Modern Mechanical Engineering, the School of Creative Science and Engineering, Waseda University, Japan, in 2018. Currently, he belongs to the at the Department of Modern Mechanical Engineering, Waseda University as an M.S. student. His research areas include deep learning, image processing, robot vision.

Jeonghwang Hayashi belongs to the Department of Modern Mechanical Engineering, the School of Creative Science and Engineering, Waseda University, Japan. His research areas include deep learning, image processing, robot vision.

Jun Ohya earned B.S, M.S. and Ph.D. degrees in Precision Machinery Engineering from the University of Tokyo, Japan, in 1977, 1979 and 1988, respectively. He joined NTT Research Laboratories in 1979. NTT sent him to the University of Maryland, USA, for one year from 1988. He transferred to ATR, Kyoto, Japan, in 1992. In 2000, he joined Waseda University as a professor. In 2005, he was a Guest Professor at the University of Karlsruhe Germany. Since 2014, he has been a professor at the Department of Modern Mechanical Engineering, Waseda University. His research areas include computer vision and machine learning. Dr. Ohya is a member of IEEE, IIEEJ, IEICE, IPSJ and VRSJ.

Hiroyuki Ogata received his BS, MS and PhD degrees in Precision Machinery Engineering from the University of Tokyo, Japan, in 1988, 1990 and 2000. He worked in NTT Research Laboratories from 1988 to 2002. Then, he moved to Seikei University as an assistant professor. Currently, he is a professor at the Department of System Design Engineering, Seikei University. His research interest includes robotics, machine learning, and psychometrics. Dr. Ogata is a member of JRSJ, JSPE, IEICE, and BSJ. He serves as a board member for JART.

Kenji Hashimoto received B.E. and M.E. degrees in 2004 and 2006, respectively, and the Ph.D. degree from the Department of Integrative Bioscience and Biomedical Engineering in 2009, all from Waseda University, Japan. During the Ph.D. course, he was funded by the Japan Society for the Promotion Science (JSPS) as a Research Fellow. He was a Postdoctoral Researcher at the Laboratoire de Physiologie de la Perception et de l 'Action in UMR 7152 College de France-CNRS, France from 2012 to 2013. He was an Assistant Professor from 2015 to 2017, and an Associate Professor from 2017 to 2018 at the Waseda Institute for Advanced Study, Waseda University. From April 2018, he joined the faculty of the Department of Mechanical Engineering Informatics, Meiji University as an Associate Professor. His research interests include legged robots and humanoid robots. He is a member of IEEE, RSJ, JSME and Japan Council of IFToMM.

Xiao Sun received B.E. and M.E. degree from the School of Automation in Beijing Institute of Technology, China in 2011 and 2014, respectively. He received Ph.D. degree from the Department of Integrative Bioscience and Biomedical Engineering of Waseda University, Japan in 2018. He was a Research Associate from 2017 to September 2019 and now he is an Assistance Professor in the Department of Modern Mechanical Engineering, Waseda University. His research interests include legged robots, ladder climbing robots and robot motion planning.

Takashi Matsuzawa received M.S. degree from the Department of Modern Mechanical Engineering, the Graduate School of Creative Science and Engineering, Waseda University, Japan in 2017. Currently, he belongs to the Department of Integrative Bioscience and Biomedical Engineering, the Graduate School of the Advanced Science and Engineering, Waseda University as an Ph.D. candidate and Research Fellowship for Young Scientists of Japan Society for the Promotion of Science (JSPS). His research areas include legged robots, motion planning and locomotion on rough terrain.

Hiroshi Naito received B.S. degree from the Department of Modern Mechanical Engineering, the School of Creative Science and Engineering, Waseda University, Japan, in 2018. Currently, he belongs to the Department of Modern Mechanical Engineering, the Graduate School of the Creative Science and Engineering, Waseda University as a master’s student. His research areas include humanoid robots, manipulation planning and robot hand.

Atsuo takanishi is a Professor of the Department of Modern Mechanical Engineering as well as the director of the Humanoid Robotics Institute, Waseda University. He received the B.S.E. degree in 1980, the M.S.E. degree in 1982 and the Ph.D. degree in 1988, all from Waseda University. His current researches are related to Humanoid Robotics and its applications in medicine and well-being, such as the biped walking/running humanoids, the emotion expression humanoids, the flute/saxophone player humanoids, the ultrasound medical inspection robots, the airway management training humanoids, etc. He recently initiated new projects on disaster response four limbed robots and mobile robots for animal monitoring. He was the former President of the Robotics Society of Japan (RSJ) from 2015 to 2016 and is the Chairman of the Japanese Council of the International Federation for the Promotion of Mechanism and Machine Science (Jc-IFToMM). He is a member of IEEE, RSJ, the Society of Mastication Systems, the Robot Revolution Initiatives, and a Vice President of the Fukuoka Prefectural Robotics and Advanced System Industry Development Council, etc. He is a fellow of RSJ, the Japanese Society of Mechanical Engineers (JSME) and a senior member of IEEE.

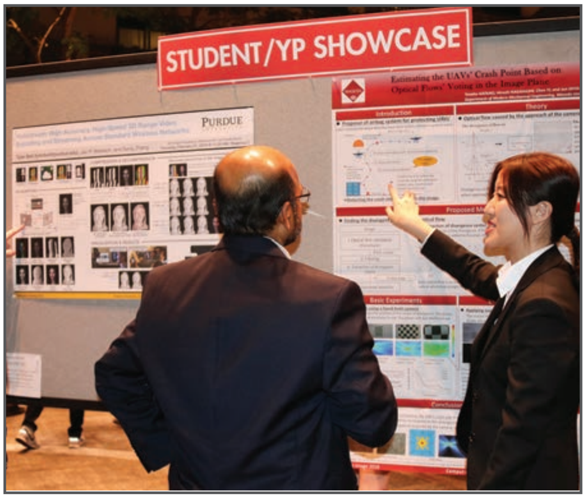
JOIN US AT THE NEXT EI!

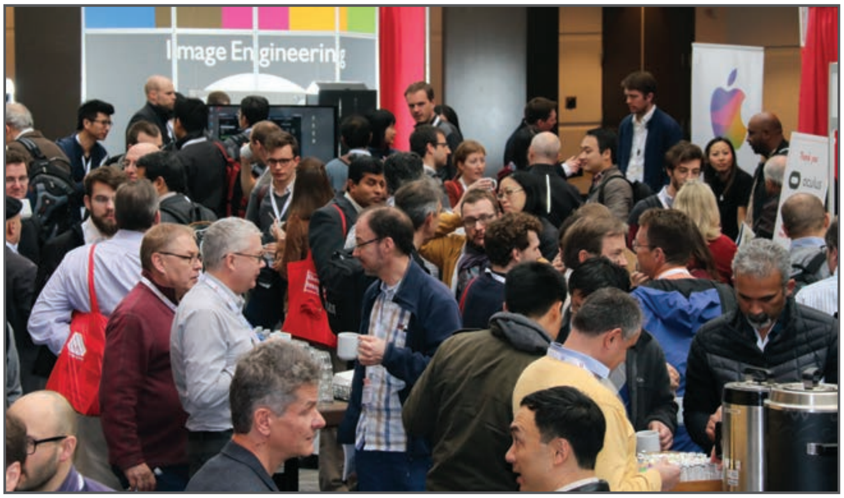
IS&T International Symposium on

Electronic Imaging

SCIENCE AND TECHNOLOGY

Imaging across applications . . . Where industry and academia meet!





• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •
• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •

www.electronicimaging.org

IS&T
imaging.org