# Use of Thermal Point Cloud for Thermal Comfort Measurement and Human Pose Estimation in Robotic Monitoring

Kaichiro Nishi Mitsuhiro Demura Jun Miura Shuji Oishi
Department of Computer Science and Engineering
Toyohashi University of Technology

## Abstract

*This paper describes applications of thermal point cloud to lifestyle support robots. 3D information is useful for recognizing human and objects based on their shapes, while thermal information is useful for assessing the residential and the human states as well as for detecting human. Combining these two kinds of information will be beneficial to the robots which live with and support people at home or in care houses. This paper shows two applications of thermal point cloud. One is thermal comfort measurement based on predictive mean vote (PMV) which uses, as one of the factors, the amount of clothing estimated by thermal information. The other is human pose estimation only by depth images, which has an advantages in terms of privacy and insensitivity to illumination changes. We developed methods for these applications and show experimental results.*

## 1. Introduction

Service robots are expected to operate in a near future in our daily life as robotics technologies is becoming matured and ready for deployment. As we are facing the *aged society*, one promising application is *monitoring*, in which a robot lives with and takes care of the elderly who lives alone.

Monitoring people has been an important application in robotics and computer vision. One approach is so-called *smart house* [14, 30, 16] which uses many embedded cameras and sensors, usually put on ceilings and walls, to monitor the state and the activities of residents inside. Since such an approach requires pre-installed sensors and cannot easily applied to normal houses. Another approach is to use wearable devices such as a thermometer and a cardiometer for health monitoring [21]. Using such devices makes it possible to take direct and reliable real-time data, but may impose a physical and/or mental burden on people.

One of the goals of the monitoring task is to examine if the physical states of a residence, such as temperature, illuminance, and air cleanness, are in a comfortable condition for a resident [13, 22]. Another goal is to examine if the resident is in a good health. Based on these examinations, responses such as advice to the resident, operation of appliances (e.g., air conditioner), or alert to the local hospital will be made. Using a mobile robot is a promising way to achieving such goals. By installing various sensors on an autonomous mobile robot, it can observe wherever and whatever it wants to collect necessary information by automatically detecting targets and navigating itself to them. While some previous works including ours [7] deal with examining physical state of a residence, this paper deals with estimating the human state.

Image data are useful for human detection and human state estimation. However, it may sometimes suffer from privacy issues as well as sensitivity to illumination conditions. Using depth images instead will address these issues, but it may make it difficult to reliably extract human regions, due to a scarcity of informative features. We therefore additionally use thermal information for locating humans, which is known to be effective especially in indoor scenes (e.g., [25]). We combine these two kinds of information into *thermal point cloud* data and apply them to human state estimation. This paper deals with two example problems, (i) thermal comfort measurement and (ii) pose estimation with various postures and occluding objects. These problems are effectively solved by using thermal point clouds. This is the contribution of the paper.

The rest of this paper is organized as follows. Section 2 describes related work. Section 3 describes human detection using a thermal point cloud. A calibration method between a thermal and a depth camera is also described. Section 4 describes an application of thermal images to estimating the human's comfort. Section 5 describes a method of estimating pose only from depth images using a deep neural network under large occlusions. An approach to generating training depth images for large occlusion cases is also described. Section 6 concludes the paper and discusses future work.

## 2. Related Work

### 2.1. Generating thermal point cloud

3D point cloud has recently been widely used for many mapping and recognition tasks thanks to the development of inexpensive RGB-D cameras and point cloud processing libraries. To add thermal information to a point cloud (i.e., to generate a thermal point cloud), a calibration between a thermal and a depth camera is necessary.

Rangel et al. [23] used a board with circular holes for a thermal-depth calibration. These holes can easily be detected by a depth camera, but could be difficult for a thermal cameras when a sufficient temperature difference does not exist between the board and the background. Rzeszotarski and Więcek [24] put aluminium sheets on white regions in a checker board for a thermal-RGB camera calibration. Aluminium sheets provide high reflection of infrared rays and are good for making markers for infrared cameras, but this approach cannot be applied to depth cameras. Vidas et al. [29] use non-colinear straight lines in the scene for calibration. They assume there are such lines visible from both a thermal and an RGB-D camera. Oreifej et al. [19] describes a method of calibrating three modalities, an optical camera, a thermal camera, and a 3D LIDAR, based on the optical-thermal and the optical-LIDAR calibration.

### 2.2. Thermal comfort measurement

Thermal comfort is human satisfaction with the thermal environment and influenced by various factors such as physical, physiological, and psychological processes [4]. One of the indices of assessing thermal comfort is PMV (predicted mean vote), determined in ISO 7730. PMV is also used for estimating another measure of thermal comfort, namely, PPD (predicted percent of dissatisfied). PMV mainly depends on four environmental factors (air temperature, mean radiant temperature, air velocity, and relative humidity) and two personal factors (clothing insulation and activity level) [13].

For measuring and estimating environmental factors, using a robot as a mobile base is an interesting research area. Previous researches deal with various applications such as odor map making [11], gas leak position localization [5], temperature and illuminance distribution mapping [7].

The activity level of a person is related to his/her metabolic rate and thus to the thermal comfort. The relationships between the metabolic rate and various activities such as seating, standing, and cooking have been analyzed [4]. Therefore activity recognition techniques (e.g., [27, 8]) could be adopted for thermal comfort measurement in robotic monitoring.

The clothing insulation, which is the other personal factor, is usually measured using a thermal mannequin [4]. The clothing insulation for various materials and designs have also been compiled in a database, which can be used for an on-line thermal comfort measurement. Matsumoto et al. [12] presented a method of measuring thermal comfort from the estimation of the material and the weight of clothing.

### 2.3. Human pose estimation

Pose estimation is an important function for a monitoring robot to know the state of a person. For example, unusual postures such as lying and crouching could be a sign of emergency.

Human pose estimation has been one of the important problems in computer vision. A large degrees of freedom of human structure and frequent occlusions sometimes make the pose estimation be a difficult task. For a robust and reliable estimation, various methods have been proposed [15, 10]. Thanks to recent advances in deep learning techniques, many image-based methods have been proposed, for example, for joint position estimation [28, 6] and part segmentation [18].

As stated above, depth image-based pose estimation has an advantage in terms of privacy and insensitivity to illumination conditions. Shotton et al. [26] developed a human pose estimation method using depth-based features with a random forest classifier. Although the method shows a nice performance, its applicability to unusual poses and heavily-occluded situations is limited.

We proposed a method of generating human depth images with pixel-wise body part labels by a combination of computer graphics and motion capture techniques [17]. We have shown that a deep neural network trained by using the generated images can recognize well a variety of human poses in real scenes.

## 3. Human Detection using Thermal Point Cloud

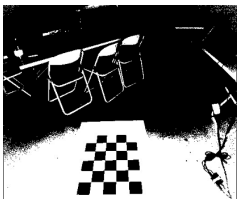
### 3.1. Thermal-depth camera calibration

One method to calibrate two cameras is to use a marker board (e.g., checker board). The relative pose between the cameras is calculated by combining two extrinsic parameters obtained by a readily-available calibration routine. In the case of thermal-depth camera calibration, making a board which is visible from both cameras is an issue.

We use PI-160 (Optris, $160 \times 120$ pixels) and Kinect v2 (Microsoft) as a thermal and a depth camera, respectively. Since it is difficult to make a board which is visible from both cameras, we additionally use an RGB camera of Kinect v2 and calculate the relative pose between the thermal and the depth cameras from those of the thermal-RGB and the RGB-depth camera pair; we use different calibration boards for each pair.

A usual checker board is used for the RGB-depth calibration, because the depth camera of Kinect v2 can also

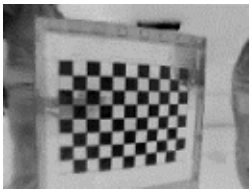
(a) RGB image.



(b) NIR image.

Figure 1: Calibration board captured by the RGB and the NIR camera.

(a) RGB image.



(b) Thermal image.

Figure 2: Calibration board captured by the RGB and the thermal camera.

produce NIR (near infrared) images, in which the board is visible. Fig. 1 shows images of the board in both camera images.

For the thermal-RGB calibration, we follow the work by Rzeszotarski and Więcek [24]. In addition to putting aluminium sheets on white regions, we heat up the board to make the temperature difference clearer for the thermal camera. Fig. 2 shows images of the board, which are visible by both the RGB and the thermal camera. Note that this aluminium-pasted board is hard to be properly observed in NIR images, and therefore we cannot directly calibrate the thermal-depth camera pair using this board.

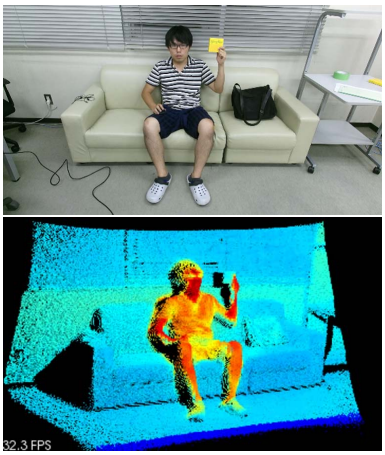
By combining the two calibration results, we have a transformation (i.e., relative pose) between the thermal and the depth camera. Using the transformation, we can attach thermal data to the point cloud. Fig. 3 shows the camera settings on a mobile robot and an example thermal point cloud. The temperature is shown in a pseudo color mapping, from red (warm) to blue (cool).

### 3.2. Human detection using thermal point cloud

We currently use a combination of a thresholding for thermal data and the Euclidean distance clustering for detecting a human using a thermal point cloud. We set a temperature threshold to $23^{\circ}\mathrm{C}$ for human region detection, and that for clustering is set to $80mm$. The size filtering is then adopted to extract only points of humans.

Fig. 4 shows a detection result. In the thermal image, since there are many warm regions and some of them are connected to the front human region, human detection using only thermal image may fail. Using the thermal point cloud,


(a) Robot and cameras.




(b) Test scene and thermal point cloud.

Figure 3: Camera settings and an example thermal point cloud.

(a) Scene. (b) Thermal image. (c) Human region.

Figure 4: Human detection using thermal point cloud.

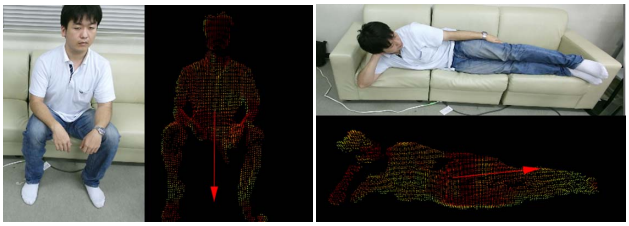


Figure 5: A simple state estimation, sitting or lying.

the human region is correctly extracted. Fig. 5 show other examples, with a principal axis estimation of the extracted region for a simple human state estimation (sitting or lying).

## 4. Thermal Comfort Measurement

### 4.1. Predicted mean vote

We first briefly explain how to calculate PMV based on [4]. Table 1 shows the correspondence between PMV and *the thermal sensation scale*.

PMV basically depends on the following values: thermal insulation (or thermal resistance) of clothing (clo-value), activity level (met-value), air temperature, air velocity, radiant temperature, and humidity. Among these, the air tem-

Table 1: PMV and thermal sensation scale.

| PMV | Thermal Sensation Scale |
|---|---|
| +3 | hot |
| +2 | warm |
| +1 | slightly warm |
| 0 | neutral |
| -1 | slightly cool |
| -2 | cool |
| -3 | cold |

perature and the humidity are measured by on-robot sensors. The radiant temperature is assumed to be the same as the air temperature. The air velocity is set to be a small value for indoor environments. The activity level is determined based on the human posture, such as sitting and standing, and actions taken. The clo-value can be approximately measured using thermal images as explained below.

### 4.2. Estimating clo-value

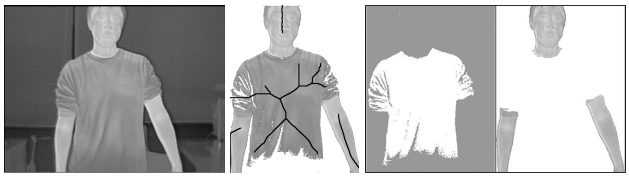
The clo-value of a clothing is usually measured on a thermal mannequin, but that cannot be used for on-line measurements in a daily situation. We instead adopt an estimation method based on thermal measurements. The clo-value $I_{clo}$ is given by [3]:

$$I_{clo} = (1/0.155 \cdot h)(t_s - t_{cl})/(t_{cl} - t_o), \qquad (1)$$

where $h$ is the heat transfer coefficient of the human body (set to $0.052$), $t_o$ is the operative temperature (approximated here by the air temperature), $t_s$ is the skin surface temperature, and $t_{cl}$ is the clothing surface temperature. The last two temperature values are measured using thermal images as follows.

Assuming that an extracted human region in the thermal image (see Sec. 3.2) is composed of the skin and the clothing regions and that the former is warmer than the latter, we apply a binarization to that region for discriminating skins and clothings. We first get an initial binarization using Otsu's method [20] and then apply a thinning operation to the segmented regions for obtaining *markers* to be used for the subsequent processing of a marker-based watershed algorithm. The output of the watershed algorithm is the final segmentation. Fig. 6 shows the process of dividing a human region into skin and clothing regions. The average temperature is calculated in the respective regions and used for calculating the clo value in eq. (1).

We show experimental results for the clo-value calculation. Fig. 7 shows thermal images and Table 2 shows the calculation results. Subject A layered two T shirts, a shorter one and a longer one; according to [3], their clo-values are predicted as $0.09$ and $0.12$, respectively, and $0.21$ in total. Subject B had the same combination. Subject C wore only the shorter one. We calculate the clo-values from the skin surface, the clothing surface, and the air temperatures, and then compared them with the predicted ones. The difference between the two clo-values are between $0.1$ to $0.5$. The differences are not very small but at least qualitatively acceptable considering the simplicity of the method.



(a) Thermal image. (b) Markers. (c) Segmentation result.

Figure 6: Detecting skin and clothing regions.

(a) Subject A. (b) Subject B. (c) Subject C.

Figure 7: Three subjects in different clothings.

Table 2: Clo-value calculation results.

| Subject | $T_s$ ($^\circ C$) | $T_c$ ($^\circ C$) | $T_a$ ($^\circ C$) | $Clo_m$ | $Clo_t$ |
|---|---|---|---|---|---|
| A | 35.83 | 31.65 | 24.50 | 0.20 | 0.21 |
| B | 30.77 | 28.49 | 25.50 | 0.26 | 0.21 |
| C | 29.76 | 28.42 | 24.75 | 0.12 | 0.09 |

$T_s$: Skin surface temperature, $T_c$: Clothing surface temperature,
$T_a$: Air temperature,
$Clo_m$: Measured clo-value, $Clo_t$: Clo-value estimated from the material.

### 4.3. Measuring Thermal Comfort

We developed an experimental system which estimates the thermal comfort automatically. We use a mobile robot equipped with the thermal-depth camera system and various sensors such as thermocouples and a humidity sensor controlled by Arduino's. Fig. 8 shows an experimental scene in which the robot finds and moves to each person to estimate the clo-values, with thermal image processing results. We measured PMV values for each person at three different heights, $170cm$, $110cm$, and $10cm$ from the floor, to see the comfort at various body positions.

Table 3 summarizes the measurement results. ISO 7730

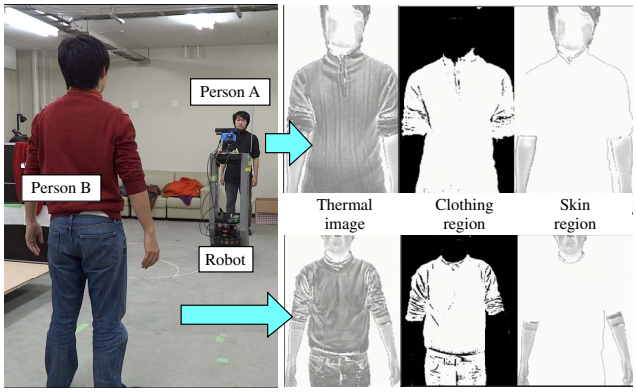



Figure 8: Thermal comfort measurement experiment.

Table 3: PMV measurement results.

| Person | Clo-value | PMV (170cm) | PMV (110cm) | PMV (10cm) |
|---|---|---|---|---|
| A | 0.21 | -2.52 | -2.86 | -2.64 |
| B | 0.22 | -2.43 | -2.75 | -2.54 |

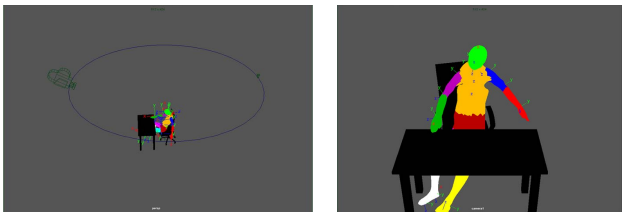
describes that a comfortable environment has the PMV value in $\pm 0.5$. The thermal environment of the experimental site is shown to be rather cool for both persons (see Table 1). To raise the lowest PMV value (-2.86) to be within the comfortable range ($\pm 0.5$), for example, the air temperature needs to be increased by nine degrees or the clo-value should be increased by 0.8 points. A monitoring robot could take an action such as turning on a heater or recommending the respective person to layer another clothing (e.g., sweater).

# 5. Pose Estimation using Depth Images with Occlusions

## 5.1. Generating training data with occlusions for a depth image-based pose estimation

We use our depth image-based human pose estimation method [17]. The method trains a deep neural network using a set of depth images with pixel-wise body part labels. In such a learning-based method, the amount and the quality of training data is a key to a high estimation performance. In a usual residential environment, people often use furniture or other apparatus and their bodies may be largely occluded. When we see a person working at a desk from his/her front side, for example, the lower part of the body is almost occluded. Such largely-occluded data should also be included in the training data.

Since our method [17] uses a computer graphics tool (i.e., Maya [2]) for generating data, we can create various largely-occluded situations by putting arbitrary objects



(a) Determining the camera pose. (b) View from the virtual camera.

Figure 9: Image generation using Maya.

around a human. As typical cases, we consider desks and chairs as objects occluding a human body. Fig. 9 shows how the labeled human body image is generated. We set the camera pose to the one on our robot system (30 degrees of downward looking angle and 110cm height, see Fig. 3(a)). We moved the camera on a circle around the vertical axis at the human model as shown in the left figure, and generate images from various viewing directions. The right figure shows the view from a camera pose, indicating the case where the upper legs are completely occluded.

We change the following three factors on the object placement to generate a variety of occlusions: the type of chairs, the existence of a desk, and the distance between a human and a desk. Fig. 10 shows the variations of scenes obtained by changing these factors. Object regions in the images are treated as background in the final images. Fig. 11 shows example pairs of depth image and body part labels.

## 5.2. Experimental results

### 5.2.1 Classifier and dataset

We used the same FCN (fully connected network) as the one used in [17]. The input to the network is a depth image with $212 \times 212$ pixels and the output is twelve labels (eleven for body parts and one for background). We generated 764,832 images and trained the network using three TITAN X GPU boards. The learning process converged about in seven days. The generated dataset and others can be downloaded at [1].

### 5.2.2 Evaluation using synthetic data

We generated 29,608 test images using in the same way as the training image generation. Fig. 12 shows example estimation results. Table 4 summarizes the results in a confusion matrix. These results show that the body parts are well recognized even under occlusion.

### 5.2.3 Results for real data

We then tested the trained classifier for the depth image regions extracted using the thermal point cloud (see Sec. 3.2).

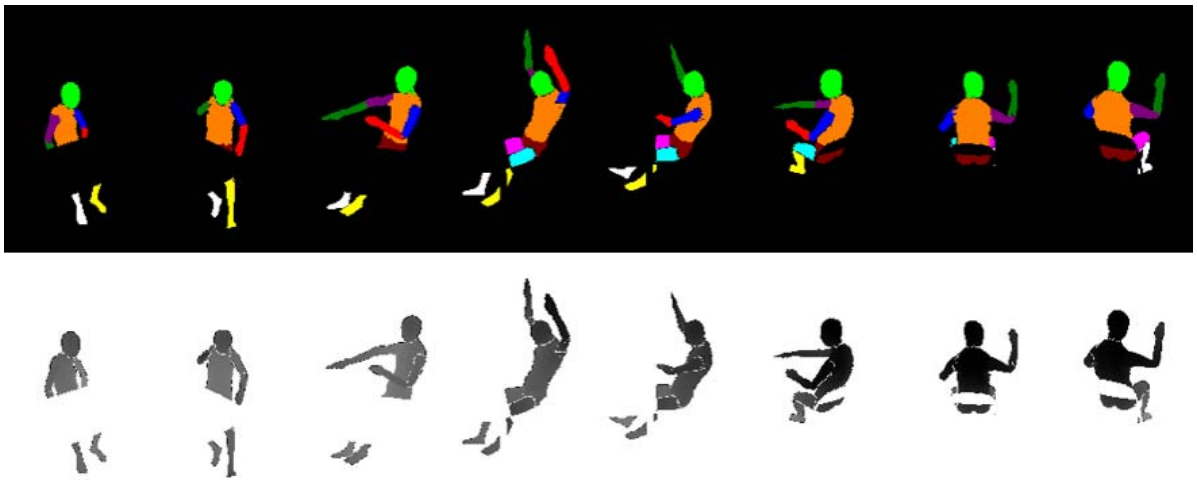


Figure 11: Example pairs of body part label images (first row) and depth images (second row).

Table 4: Confusion matrix for synthetic data.

| Actual \ Estimated | HD | TRS | LUA | RUA | LFA | RFA | Hip | LUL | RUL | LLL | RLL | BG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HD | **0.97** | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| TRS | 0.00 | **0.95** | 0.00 | 0.01 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| LUA | 0.00 | 0.06 | **0.87** | 0.00 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 |
| RUA | 0.01 | 0.10 | 0.00 | **0.82** | 0.00 | 0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 |
| LFA | 0.00 | 0.01 | 0.02 | 0.00 | **0.88** | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.06 |
| RFA | 0.01 | 0.00 | 0.00 | 0.03 | 0.00 | **0.92** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 |
| Hip | 0.00 | 0.04 | 0.00 | 0.00 | 0.02 | 0.00 | **0.89** | 0.00 | 0.01 | 0.00 | 0.00 | 0.04 |
| LUL | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | **0.88** | 0.03 | 0.01 | 0.01 | 0.04 |
| RUL | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.03 | 0.00 | **0.92** | 0.00 | 0.01 | 0.03 |
| LLL | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | **0.88** | 0.05 | 0.06 |
| RLL | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.01 | **0.92** | 0.05 |
| BG | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | **1.00** |

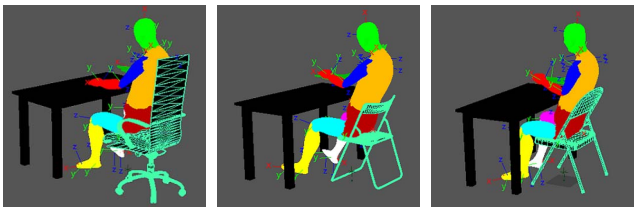
HD: Head, TRS: Torso, LUA: left upper arm, RUA: right upper arm, LFA: left forearm, RFA: right forearm
LUL: left upper leg, RUL: right upper leg, LLL: left lower leg, RLL: right lower leg, BG: background

Fig. 13 shows results for several real scenes. The columns shows the following from left to right: scene images, thermal point clouds, extracted human regions in the depth image, the estimation results. Note that RGB images in the first column are not used in pose estimation. The estimation results are reasonably good even in sometimes heavy occlusions.

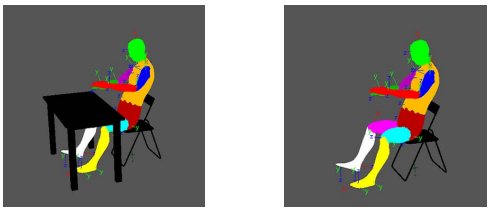
## 6. Conclusions and Discussion

This paper describes two applications of thermal point cloud in robotic monitoring task: thermal comfort measurement and human pose estimation. Thermal point cloud is generated by combining data from a calibrated pair of thermal and depth cameras and effective in extracting human regions. Thermal comfort is evaluated using PMV, which is measured based on a combination of environmental and personal factors. One of the personal factors, clothing insulation, is automatically estimated by analyzing an extracted human thermal image. Pose estimation is carried out using a depth image-based deep neural network. Considering frequent largely-occluded situations, we developed a method of generating training data with various occlusion cases. The classifier trained by using the data is shown to be effective in the experiments using a synthetic and a real data set.

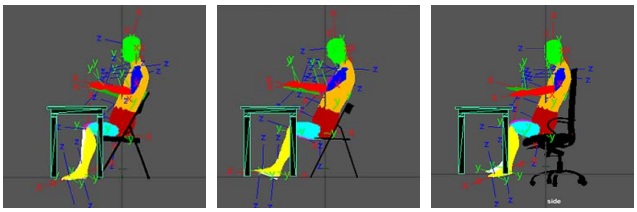
Currently, the factors in measuring thermal comfort are approximated or assumed; for example, we use a simple temperature difference-based method for estimating clo-



(a) Changing chair types.



(b) Switching the existence of a desk.



(c) Changing distances to the desk: 50cm, 60cm, and 70cm from left to right.

Figure 10: Object placement variations.

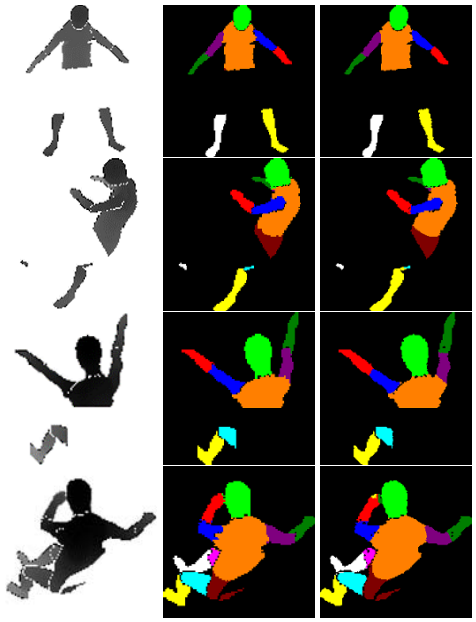


Figure 12: Estimation results for synthetic data. Left: input depth images, Center: correct part labels, Right: estimation results.

values. Although the obtained clo-values are reasonable, we need to consider a more complicated model (e.g., [9]) or more various combinations of clothings and environments. Applying pose estimation, activity recognition, and activity classification techniques to determine the activity level is also an interesting research direction.

The current pose estimation method seems to work well, but needs to be tested for a more variety of situations. Connecting the pose estimation results to assistive/supportive operations, such as vital sign recognition using a specific body parts (e.g., head part), is also necessary for real application situations.

## Acknowledgment

This work is in part supported by JSPS KAKENHI Grant Numbers 25280093/17H01799 and the Hibi Science Foundation.

## References

[1] AISL-HDIBPL dataset, `http://www.aisl.cs.tut.ac.jp/database_hdibpl.html`.

[2] Maya, `http://www.autodesk.com/products/maya/overview/`.

[3] *Mechanisms of Thermal Comfort*. The Society of Heating, Air-Conditioning and Sanitary Engineers of Japan, 2006.

[4] *ASHRAE Handbook of Fundamentals*. Americal Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2009.

[5] V. H. Bennetts, A. Lilienthal, A. Khaliq, V. P. Sesé, and M. Trincavelli. Towards real-world gas distribution mapping and leak localization using a mobile robot with 3d and remote gas sensing capabilities. In *Proceedings of 2013 IEEE Int. Conf. on Robotics and Automation*, pages 2327–2332, 2013.

[6] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In *Proceedings of 2017 IEEE Conf. on Computer Vision and Pattern Recognition*, 2017.

[7] S. Kani and J. Miura. Mobile monitoring of physical states of indoor environments for personal support. In *Proceedings of 2015 IEEE/SICE Int. Symp. on System Integration*, pages 393–398, 2015.

[8] H. Koppula, R. Gupta, and A. Saxena. Learning human activities and object affordances from rgb-d videos. *Int. J. of Robotics Research*, 32(8):951–970, 2013.

[9] J.-H. Lee, Y.-K. Kim, K.-S. Kim, and S. Kim. Estimating clothing thermal insulation using an infrared camera. *Sensors*, 16(3), 2016.

[10] Z. Liu, J. Zhu, J. Bu, and C. Chen. A survey of human pose estimation: the body parts parsing based methods. *J. of Visual Communication and Image Representation*, 32:10–19, 2015.

[11] L. Marques, A. Martins, and A. de Almeida. Environmental monitoring with mobile robots. In *Proceedings of the*

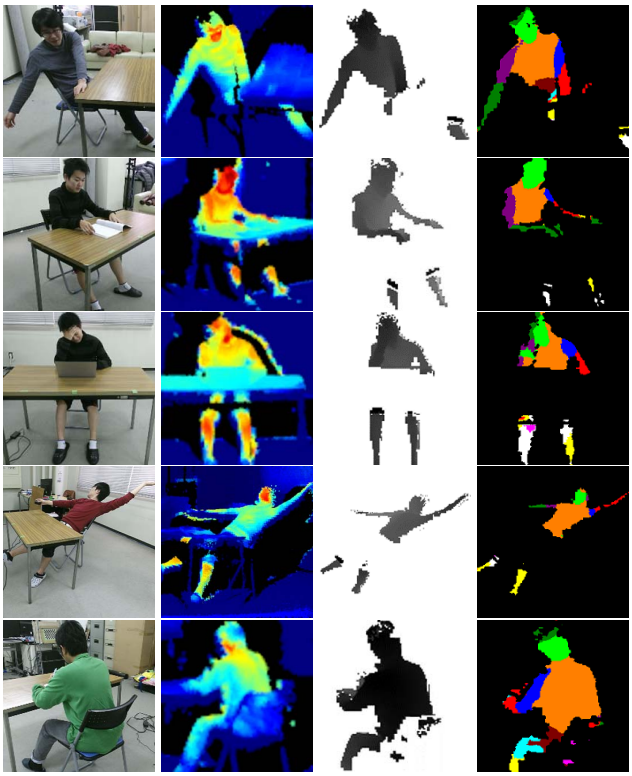


Figure 13: Pose estimation results in real scenes. Columns from left to right: scene images (not used for pose estimation), thermal point clouds, extracted human regions in the depth image, and the estimation results.


*2005 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 3624–3629, 2005.

[12] H. Matsumoto, Y. Iwai, and H. Ishiguro. Estimation of thermal comfort by measuring clo value without contact. In *Proceedings of 2011 IAPR Conf. on Machine Vision Applications*, pages 491–494, 2011.

[13] A. Melikov. Design and assessment of indoor environment. In *Proceedings of CLIMA 2000*, 1997.

[14] I. Mikic, K. Huang, and M. Trivedi. Activity monitoring and summarization for an intelligent meeting room. In *Proceedings of IEEE Workshop on Human Motion*, 2000.

[15] T. Moeslund, A. Hilton, and V. Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104:90–126, 2006.

[16] T. Mori, S. Tominaga, H. Noguchi, M. Shimoasaka, R. Fukui, and T. Sato. Behavior prediction from trajectories in a house by estimating transition model using stay points. In *Proceedings of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 3419–3425, 2011.

[17] K. Nishi and J. Miura. Generation of human depth images with body part labels for complex human pose recognition. *Pattern Recognition*, 2017.

[18] G. Oliveira, A. Valada, C. Bollen, W. Burgard, and T. Brox. Deep learning for human part discovery in images. In *Proceedings of 2016 IEEE Int. Conf. on Robotics and Automation*, 2016.

[19] O. Oreifej, J. Cramer, and A. Zakhor. Automatic generation of 3d thermal maps of building of interiors. *ASHRAE trans.*, 120(2), 2014.

[20] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-9(1):62–66, 1979.

[21] A. Pantelopoulos and N. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Trans. on Systems, Man, and Cybernetics Part C: Applications and Reviews*, 40(1):1–12, 2010.

[22] D. Raimondo, S. Corgnati, and B. Olesen. Evaluation methods for indoor environmental quality assessment according to EN15251. *REHVA European HVAC J.*, 04/2012:14–19, 2012.

[23] J. Rangel, S. Soldan, and A. Kroll. 3D thermal imaging: Fusion of thermography and depth cameras. In *Proceeings of 12th Int. Conf. on Quantitative InfraRed Thermography*, 2014.

[24] D. Rzeszotarski and B. Więcek. An integrated thermal and visual camera system for 3d reconstruction. In *Proceedings of 11th Int. Conf. on Quantitative InfraRed Thermography*, 2012.

[25] C. Schönauer, E. Vonach, G. Gerstweiler, and H. Kaufmann. 3D building reconstruction and thermal mapping in fire brigade operations. In *4th Augmented Human Int. Conf.*, pages 202–205, 2013.

[26] J. Shotton, T. Sharp, A. Kipman, A. Fitzgiboon, M. Finocchio, A. Blake, M. Cook, and R. Moore. Real-time human pose recognition in parts from single depth images. *Communications of the ACM*, 56(1):116–124, 2013.

[27] M. Tenorth, F. de la Torre, and M. Beetz. Learning probability distributions over partially-ordered human everyday activities. In *Proceedings of 2013 IEEE Int. Conf. on Robotics and Automation*, pages 4539–4544, 2013.

[28] Z. Toshev and C. Szegedy. Deeppose: Human pose estimation via deep neural networks. In *Proceedings of 2014 IEEE Conf. on Computer Vision and Pattern Recognition*, pages 1653–1660, 2014.

[29] S. Vidas, P. Moghadam, and M. Bosse. 3D thermal mapping of building interiors using an rgb-d and thermal camera. In *2013 IEEE Int. Conf. on Robotics and Automation*, pages 2303–2310, 2013.

[30] C. Wu, A. Khalili, and H. Aghajan. Multiview activity recognition in smart homes with spario-temporal features. In *Proceedings of the 4th ACM/IEEE Int. Conf. on Distributed Smart Cameras*, pages 142–149, 2010.