Estimation of Human Orientation using Coaxial RGB-Depth Images

Fumito Shinmura[1], Daisuke Deguchi[2], Ichiro Ide[3], Hiroshi Murase[3] and Hironobu Fujiyoshi[4]
[1]*Institute of Innovation for Future Society, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan*
[2]*Information & Communications, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan*
[3]*Graduate School of Information Science, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan*
[4]*Department of Robotics Science and Technology, Chubu University, 1200 Matsumoto-cho, Kasugai, Aichi, Japan*

Keywords: Human Orientation Estimation, Single-chip RGB-ToF Camera, RGB-D.

Abstract: Estimation of human orientation contributes to improving the accuracy of human behavior recognition. However, estimation of human orientation is a challenging task because of the variable appearance of the human body. The wide variety of poses, sizes and clothes combined with a complicated background degrades the estimation accuracy. Therefore, we propose a method for estimating human orientation using coaxial RGB-Depth images. This paper proposes Depth Weighted Histogram of Oriented Gradients (DWHOG) feature calculated from RGB and depth images. By using a depth image, the outline of a human body and the texture of a background can be easily distinguished. In the proposed method, a region having a large depth gradient is given a large weight. Therefore, features at the outline of the human body are enhanced, allowing robust estimation even with complex backgrounds. In order to combine RGB and depth images, we utilize a newly available single-chip RGB-ToF camera, which can capture both RGB and depth images taken along the same optical axis. We experimentally confirmed that the proposed method can estimate human orientation robustly to complex backgrounds, compared to a method using conventional HOG features.

1 INTRODUCTION

Vision based human behavior recognition is widely used in security surveillance (Hu et al., 2004), consumer demand research (Hu et al., 2009), and gesture recognition (Chen and Koskela, 2014). Since the estimation of the human orientation is important for improving the accuracy of human behavior recognition, we are focusing on the accurate estimation of human orientation. A common approach for obtaining human orientation refers to the walking trajectory estimated by an object tracking technique. However, in the case of a moving camera, it is difficult to obtain a correct human walking trajectory. Therefore, this paper proposes a method to estimate human orientation using a single-shot image. In this paper, the human orientation is defined as shown in Fig. 1.

The human orientation estimation problem has been approached by many research groups, and various methods have been proposed. Gandhi and Trivedi used Histogram of Oriented Gradients (HOG) features and a support vector machine (SVM) for the estimation (Gandhi and Trivedi, 2008). Weinrich et al. used HOG features and a decision tree with SVMs for the estimation (Weinrich et al., 2012). Methods for human pose estimation have been also proposed, and the human orientation can be understood from the human pose. Straka et al. used skeleton graph extraction and skeleton model fitting (Straka et al., 2011), and Shotton et al. used body part classification and offset joint regression by randomized forests for estimating human pose (Shotton et al., 2013).

In the above methods, an RGB image is often used for the estimation of human orientation. Most of these methods calculate image features, such as intensity gradients and human body textures. These image features are effective to distinguish shapes and appearances of each pose. However, the image features are easily affected by a complicated background, since it has variable texture patterns, including a texture similar to a human body. In such case, the distinction between the outline of a human body and the texture of a background is difficult. Accordingly, the accuracy of orientation estimation is often degraded by the background scene.

On the other hand, some techniques using depth information have also been reported. In these methods, a depth image is utilized to represent 3D rela-

Shinmura F., Deguchi D., Ide I., Murase H. and Fujiyoshi H..
Estimation of Human Orientation using Coaxial RGB-Depth Images.
DOI: 10.5220/0005305301130120
In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* (VISAPP-2015), pages 113-120
ISBN: 978-989-758-090-1

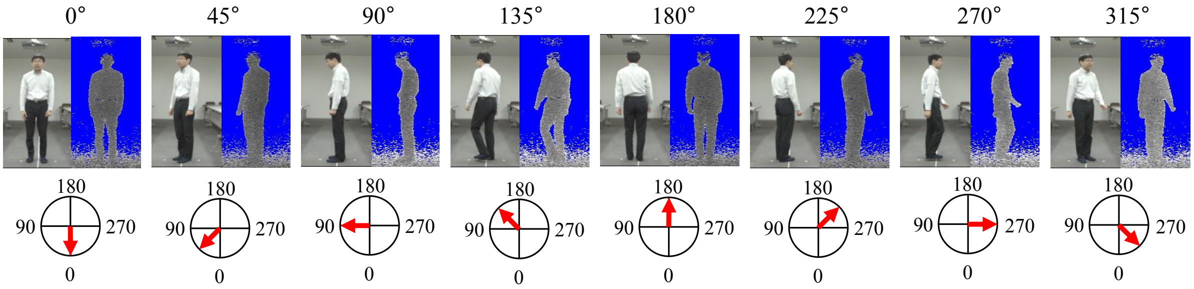


Figure 1: Example of images of human orientation.

tionships of body parts. Shotton et al. proposed the depth comparison feature (Shotton et al., 2013) which is defined as the difference of depth between two pixels. Since the depth of the human body and the background differ significantly, a human body region can easily be extracted. However, since body textures cannot be obtained from the depth image, it is difficult to distinguish between the front view and the back view of the human body.

Since a difference of depth occurs on the outline, but not on the pattern, the outline of human body and the texture of background can easily be distinguished by using depth images. Thus methods using depth images are expected to overcome the drawback of RGB image features. Therefore, combining RGB and depth images should also be effective for estimating human orientation.

In recent years, RGB-D cameras such as Microsoft Kinect have become commercially available. Liu et al. proposed a method for estimating human orientation using an RGB-D camera (Liu et al., 2013). Their method uses a viewpoint feature histogram for static cues and scene-flow information for motion cues, and combines these cues using a dynamic Bayesian network system. Human orientation is estimated based on these combined cues. However, this method cannot be applied to a moving camera and cannot estimate orientation when human tracking fails. Therefore, it is necessary to develop a method for estimating human orientation from a single-shot image. Additionally, although this kind of RGB-D camera can obtain both RGB and depth images simultaneously, the RGB and depth images are captured along different optical axes. Therefore, it is important to consider the difference of optical axes when calculating features for orientation estimation.

To overcome the above problems, this paper proposes a method for estimating human orientation from a single-shot image. In addition, the proposed method introduces the usage of a newly available single-chip RGB-ToF camera, which can capture both RGB and depth images along the same optical axis. By using this camera, this paper proposes an RGB-D image feature that considers the coaxial characteristics of the camera. Contributions of this paper are as follows:

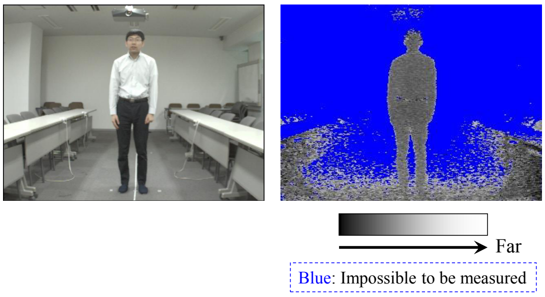
(1) First research on human orientation estimation using a single-chip RGB-ToF camera.

(2) Proposal of Depth Weighted HOG (DWHOG) feature, which is a variant of the HOG feature that suppresses the influence of the background texture.

In the following, the single-chip RGB-ToF camera is explained in section 2. The proposed method is presented in section 3. The results of the experiments are discussed in section 4. Finally, we conclude this paper in section 5.

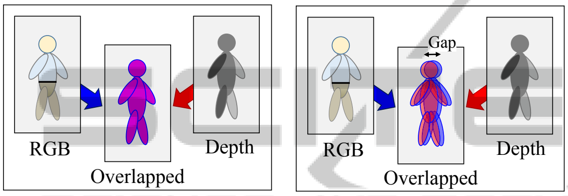
2 SINGLE-CHIP RGB-ToF CAMERA

In this paper, we make use of a single-chip RGB-ToF (RGB-D) camera for human orientation recognition. Currently, RGB-D cameras are widely used in many applications, such as consumer games with Kinect. However, most of them consist of a separate RGB camera and a depth camera, so they do not share the same optical axis. Therefore, it is difficult to integrate information from both cameras to improve the accuracy of human orientation recognition. To overcome this problem, we introduce the usage of a newly available single-chip coaxial RGB-ToF camera (Panasonic MN34901TL). Figure 2 shows the images acquired by this camera. This camera can coaxially acquire an RGB image and an infrared image, and it has an ability to measure the target depth by the Time-of-Flight (ToF) principle using infrared light. This camera allows us to use spatially aligned RGB and depth data simultaneously as shown in Fig. 3, and by combining the two, we expect to improve the accuracy of human orientation estimation.



(a) RGB image. (b) Depth image.

Figure 2: Example of images acquired from the RGB-ToF camera. The blue regions in the depth image are impossible to be measured since they are outside the sensor range.



(a) Coaxial. (b) Non-coaxial.

Figure 3: Coaxial and non-coaxial RGB-D images.

3 ESTIMATION OF HUMAN ORIENTATION BY SINGLE-CHIP RGB-ToF CAMERA

This section presents the method that we propose for human orientation estimation using the coaxial RGB-D camera. Figure 4 shows the process flow of the proposed method.

3.1 Basic Idea

The proposed method consists of the following four steps:

1. Noise reduction of the depth image by referring to the RGB image.
2. Computation of features from coaxial RGB-Depth images.
3. Construction of a classifier for human orientation estimation.
4. Estimation of human orientation.

Given RGB and depth images from the RGB-ToF camera, we first reduce noise from the depth images. Since depth images are easily affected by noise from the environment and the sensor itself, it is very important to reduce them. As we noticed that RGB images can be observed with slightly less noise in comparison with depth images, the proposed method tries to reduce the noise of the depth image by using information from the coaxially obtained RGB image, with a cross-bilateral filter (Pestschnigg et al., 2004; Yang et al., 2013). Here, the coefficients of the cross-bilateral filter are calculated from the RGB image.

Next, the RGB-D features are computed. The appearance features such as the shape and the texture of a human body can be obtained from the RGB image. Since the appearance of the human body changes depending on orientation, we decided to employ the HOG feature to represent the appearance. However, the RGB image features are easily affected by the background texture. Therefore, the influence of the background is reduced by emphasizing the human body outline provided by the depth image. The HOG features are weighted according to the magnitude of depth gradient. Weighting the region with a large depth gradient, which corresponds to the outline of a human body, with a larger weight provides robustness to a textured background, while preserving texture areas within the human body outline. These are expected to allow robust estimation to complex backgrounds.

Finally, human orientation is classified into a discrete number of direction in orientation estimation. In this paper, the human body orientation estimation problem is approached by classifying human orientation into eight directions ($0°, 45°, \ldots, 315°$) as shown in Fig. 1. Therefore, a multi-class SVM is used for estimating the human orientation. For implementation, we use LIBSVM (Chang and Lin, 2011), which is a library for support vector machines.

3.2 Training Phase

In the training phase, we construct the estimator for human orientation, as shown in the left side of Fig. 4. The depth image is first smoothed, and the RGB-D features are computed from each RGB and depth image pairs. These computed features are combined, and the estimator is constructed by training these features.

3.2.1 Noise Reduction of the Depth Image Referring to the RGB Image

Assuming the image coordinate of the pixel of an interest to be $\boldsymbol{x}$, the noise reduction using coaxial RGB-D characteristic can be formulated as a cross-bilateral filter (Pestschnigg et al., 2004) as

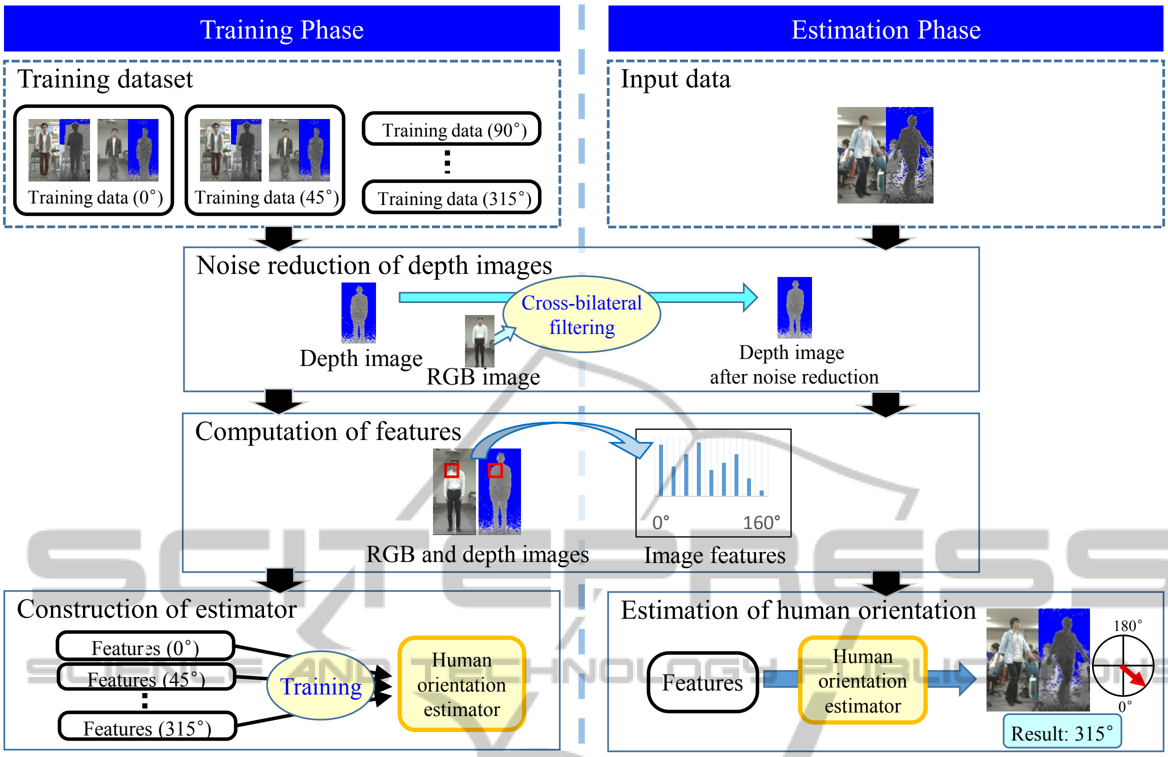


Figure 4: Process flow of the proposed method.

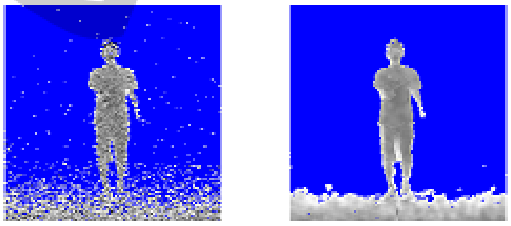
$$F(\boldsymbol{x}) = \frac{\sum_{\boldsymbol{x}' \in N(\boldsymbol{x})} w_d(\boldsymbol{x},\boldsymbol{x}') w_v(g(\boldsymbol{x}),g(\boldsymbol{x}')) f(\boldsymbol{x})}{\sum_{\boldsymbol{x}' \in N(\boldsymbol{x})} w_d(\boldsymbol{x},\boldsymbol{x}') w_v(g(\boldsymbol{x}),g(\boldsymbol{x}'))}, \quad (1)$$

$$w_d(\boldsymbol{x},\boldsymbol{x}') = \exp(-\frac{||\boldsymbol{x}-\boldsymbol{x}'||^2}{2\sigma_1^2}), \quad (2)$$

$$w_v(g(\boldsymbol{x}),g(\boldsymbol{x}')) = \exp(-\frac{(g(\boldsymbol{x})-g(\boldsymbol{x}'))^2}{2\sigma_2^2}), \quad (3)$$

where $N(\boldsymbol{x})$ are neighborhood pixels of $\boldsymbol{x}$, $\boldsymbol{x}'$ is the coordinates of a pixel in $N(\boldsymbol{x})$, σ_1 and σ_2 are smoothing parameters, respectively. Functions $f(\cdot)$ and $g(\cdot)$ represent the pixel values of the depth image and that of the RGB image, respectively. Function w_d is the weight assigned according to the spatial distance, and w_v is the weight assigned according to the difference of pixel values, as shown in Eqs. (2) and (3).

An example by applying this cross-bilateral filter to a depth image is shown in Fig. 5. Cross-bilateral filtering reduced noise while preserving the outline of the object, and filled in the pixel gaps. Figure 5 also shows how bumps on the outline of the object and the salt and pepper noise are removed after applying the cross-bilateral filter.



(a) Before application. (b) After application.

Figure 5: Example of depth images before and after applying the cross-bilateral filter.

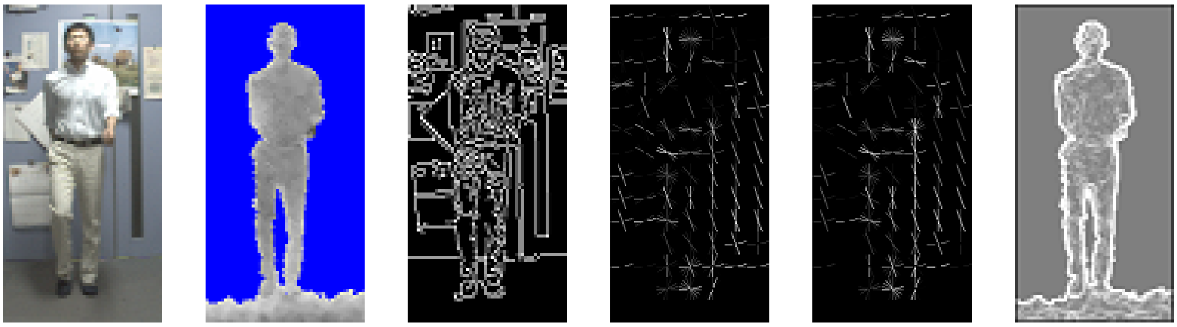
3.2.2 Computation of Depth Weighted HOG (DWHOG) Features

This paper proposes a HOG feature weighted by the magnitude of depth gradient, which we name the "Depth Weighted HOG (DWHOG)".

The original HOG feature is computed from RGB images as in (Dalal and Triggs, 2005). The gradient strengths of HOG features are generally computed as

$$m(\boldsymbol{x}) = \sqrt{\left(\frac{d}{dx}g(\boldsymbol{x})\right)^2 + \left(\frac{d}{dy}g(\boldsymbol{x})\right)^2}, \quad (4)$$

where $\boldsymbol{x}$ is the image coordinate of the pixel, $m(\boldsymbol{x})$ is the gradient strength at $\boldsymbol{x}$. $\frac{d}{dx}g(\boldsymbol{x})$ and $\frac{d}{dy}g(\boldsymbol{x})$ are horizontal and vertical differential values of intensity



(a) Input RGB image. (b) Input Depth image. (c) RGB edge image. (d) Conventional HOG feature. (e) DWHOG feature. (f) Depth weight of DWHOG feature.

Figure 6: Example of images visualizing image features and depth weight.

at $\boldsymbol{x}$, respectively.

In the proposed method, the gradient strengths are weighted as

$$m(\boldsymbol{x}) = W_d(\boldsymbol{x}) \sqrt{\left(\frac{d}{dx}g(\boldsymbol{x})\right)^2 + \left(\frac{d}{dy}g(\boldsymbol{x})\right)^2}. \quad (5)$$

Here, the weight $W_d(\boldsymbol{x})$ is computed using depth values, such as

$$W_d(\boldsymbol{x}) = \frac{1}{1+e^{-0.01n(\boldsymbol{x})}}, \quad (6)$$

$$n(\boldsymbol{x}) = \sqrt{\left(\frac{d}{dx}f(\boldsymbol{x})\right)^2 + \left(\frac{d}{dy}f(\boldsymbol{x})\right)^2}, \quad (7)$$

where $\frac{d}{dx}f(\boldsymbol{x})$ and $\frac{d}{dy}f(\boldsymbol{x})$ are horizontal and vertical differential values of depth at $\boldsymbol{x}$, respectively. The regions having a large depth gradient corresponds to the human body outline. Therefore, the weighting method of Eqs. (5) – (7) enhances the features on the outline of the human body.

The images visualizing features and depth weights are shown in Fig. 6. Figures 6(a) and (b) are the input RGB image and the input depth image, respectively. Figure 6(c) displays the RGB edge image obtained by Canny edge detector (Canny, 1986), where we can see many large intensity gradients besides the outline of a human body. Figure 6(d) shows the visualized conventional HOG feature, and Fig. 6(e) shows the visualized DWHOG feature. These images display the principal directions with large intensity gradient in each block region, and lines with brighter colors represent the directions with large intensity gradients. We can see that the HOG feature has larger intensity gradient on backgrounds than the proposed DWHOG feature. Therefore, although the conventional HOG feature is influenced by gradient directions in backgrounds, the DWHOG feature selects the gradient directions corresponding to the outline of the human body. Figure 6(f) shows the depth weight which is used when the DWHOG feature is computed. Regions with brighter colors represent those with larger weight. It can be confirmed that a large weight is calculated at the outline of the human body.

3.2.3 Construction of a Human Orientation Estimator

A multi-class SVM classifier for eight directions is constructed as the human orientation estimator. The implementation for SVM multi-class classification is the one-against-one method (Chang and Lin, 2011). For training, training data of each orientation are prepared, and the classifier learns DWHOG features computed from these data.

3.3 Human Orientation Estimation Phase

In the estimation phase, the human orientation is estimated from input images by using the constructed estimator, as shown in the right side of Fig. 4.

Following the same procedure as in the training phase, the depth images are first smoothed, and the RGB-D features are computed from input RGB and depth images. Finally, the human orientation is estimated from the computed features by using the estimator constructed beforehand.

4 EXPERIMENT

We conducted experiments on human orientation estimation in order to evaluate the effectiveness of the proposed method.

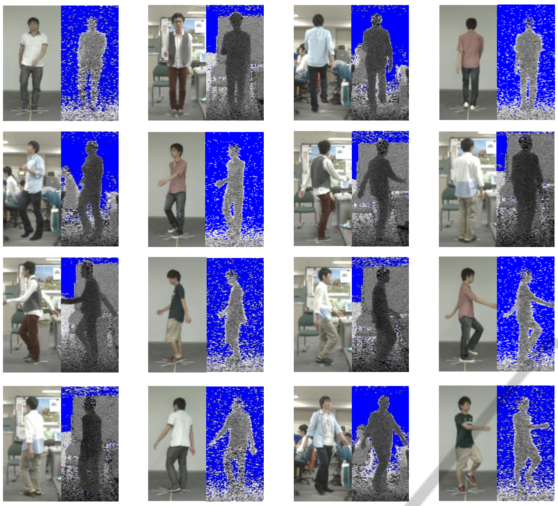


Figure 7: Example of RGB and depth images for the experiment.

4.1 Test Data

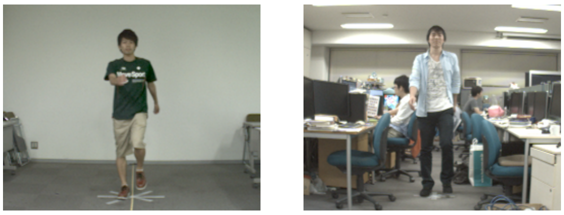
For the experiment, we prepared RGB and depth images captured with the newly available single-chip RGB-ToF camera (Panasonic MN34901TL). The resolution of the captured RGB image was 640×480 pixels and that of the captured depth image was 320×240 pixels. We cropped the human regions from these images manually, and used for the experiment. Example of the prepared images are shown in Fig. 7. We prepared 9,600 images in total, which included eight human orientations and six persons (both standing still and walking).

4.2 Experiments and Results

The accuracy rate of estimating human orientation was used as an evaluation criteria. In the experiment, data of five persons (8,000 images) were used for training and data of another person (1,600 images) were used for evaluation. The experiment was repeated six times with different evaluation data, and then the average correction rate was computed in the 6-fold cross validation manner.

In order to confirm the effectiveness of the proposed RGB-D features in estimating human orientation, we carried out two types of experiments.

First, we carried out an experiment in order to confirm the effectiveness of the proposed DWHOG feature. We compared the proposed method with two current methods; a method using conventional HOG feature and a method using conventional HOG feature and a simple depth feature. Here, as simple depth feature, we used the difference of depth between body sections inspired by Shotton et al.'s method (Shotton et al., 2013). Their method employed the difference of depth between two pixels selected randomly for pose estimation. Their study indicated the effectiveness of the difference of depth features. However, the difference of depth between pixels was vulnerable to noise. Therefore the difference of depth between block regions was used here. This feature should approximately estimate the body inclination. The result of this experiment is shown in Table 1. In addition, the result of each orientation is shown in Table 2, and the confusion matrix of the proposed method (DWHOG) is shown in Table 3.



(a) Image with a simple background.

(b) Image with a complicated background.

Figure 8: Example of images with simple and complicated backgrounds.

Table 1: Experimental results.

Method (Used features)	Accuracy
HOG	64.8 %
HOG + Difference of depth	72.4 %
Depth Weighted HOG (Proposed method)	78.9 %

Next, we carried out an experiment in order to confirm the robustness to a complicated background. Figure 8 shows images with a simple and a complicated backgrounds, respectively. The simple background had fewer texture patterns, and the complicated background had many complex texture patterns. We compared the experimental results using images with complicated backgrounds with those using images with simple backgrounds. The results of this comparison is shown in Table 4. The prepared images included 4,800 images with complicated backgrounds and the same number of images with simple backgrounds. Generally, a human body orientation in a simple background such as in Fig. 8(a) was expected to be able to accurately estimated. In contrast, the estimation accuracy of human orientation in a complicated background such as in Fig. 8(b) was expected to be reduced because of the influence of backgrounds. Therefore, the proposed method was expected to show less reduction in orientation accuracy when compared to other methods.

Table 2: Experimental results of each orientation.

Method (Used features)	Accuracy							
	0°	45°	90°	135°	180°	225°	270°	315°
HOG	69.2 %	66.4 %	68.5 %	64.9 %	50.0 %	62.0 %	92.5 %	44.9 %
HOG + Difference of depth	63.8 %	77.1 %	68.6 %	73.3 %	61.7 %	75.5 %	95.5 %	64.2 %
Depth Weighted HOG (Proposed method)	69.6 %	77.7 %	79.3 %	91.8 %	66.1 %	78.6 %	94.3 %	73.9 %

Table 3: Confusion matrix of the proposed method (DWHOG).

		Correct orientation							
		0°	45°	90°	135°	180°	225°	270°	315°
Estimated orientation	0°	**69.6 %**	0.8 %	0 %	20.6 %	0.9 %	6.9 %	0.9 %	0.3 %
	45°	1.4 %	**77.7 %**	13.7 %	7.3 %	0 %	0 %	0 %	0 %
	90°	0 %	20.5 %	**79.3 %**	0.1 %	0 %	0 %	0.1 %	0 %
	135°	0.8 %	1.9 %	2.5 %	**91.8 %**	2.8 %	0.3 %	0 %	0 %
	180°	0.2 %	0 %	0 %	28.2 %	**66.1 %**	4.9 %	0.7 %	0 %
	225°	1.3 %	0 %	0 %	9.9 %	1.7 %	**78.6 %**	8.5 %	0 %
	270°	0 %	0 %	0 %	0.5 %	0 %	0.4 %	**94.3 %**	4.8 %
	315°	1.6 %	0 %	0 %	5.3 %	0 %	7.9 %	11.3 %	**73.8 %**

Table 4: Complicated backgrounds vs. Simple backgrounds.

Method (Used features)	Accuracy	
	Complicated backgrounds	Simple backgrounds
HOG	54.7 %	74.9 %
HOG + Difference of depth	52.5 %	92.4 %
Depth Weighted HOG (Proposed method)	63.9 %	93.9 %

4.3 Discussion

As shown in Table 1 and Table 4, the proposed method achieved the highest accuracy of the three methods. In addition, the degradation of the accuracy between the experiment using images with simple backgrounds and those with complicated backgrounds was smaller than the method using the HOG and the difference of depth features. This shows that the proposed method's accuracy was least affected by background textures.

Focusing on the accuracy of each orientation in Table 2, the accuracy at 135° and 315° were particularly improved by the proposed method compared to the method using the simple HOG feature. In order to distinguish a human body from an oblique view point, the use of the human body shape is effective. However, when there is a large intensity gradient in the background, the HOG feature cannot correctly represent the shape of a human body. The DWHOG feature solved this problem by emphasizing the human body outline, and improved the estimation accuracy from an oblique view point.

Sample images where human orientation was correctly estimated by the proposed method are shown in Fig. 9, and sample images where human orientation estimation was incorrect are shown in Fig. 10. When a person swings his/her arm largely, the proposed method may fail to estimate the orientation. The proposed method weighted the outline of the human body. However, it was not appropriate to weight the body parts with large movement such as arms and

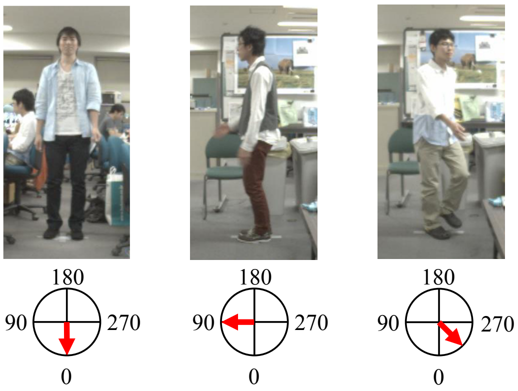


Figure 9: Example of images where human orientation was correctly estimated.

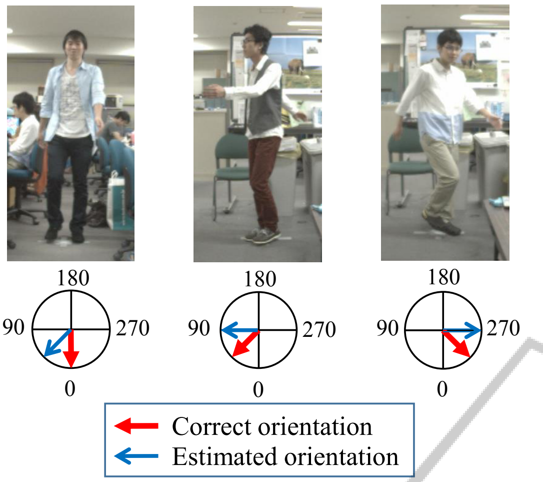


Figure 10: Example of images where human orientation was incorrectly estimated.

legs with a large weight, and these body parts should be weighted with a small weight. In order to solve this problem, it is necessary to investigate a method of weighting adapted to estimating human orientation based on the selection of body part.

5 CONCLUSION

This paper proposed a method for estimating human orientation using coaxial RGB and depth images. We utilized a newly available single-chip RGB-ToF camera in order to use coaxial RGB and depth features. This paper is the first research on human orientation estimation using this camera, and we propose a novel combination of RGB and depth features (Depth Weighted HOG). We experimentally confirmed the effectiveness of the combination of RGB and depth features. Our future work will include development of a more effective feature to combine RGB and depth information considering high motion body areas such as arms and legs.

ACKNOWLEDGEMENTS

Parts of this research were supported by JST, Nagoya University COI and MEXT, Grant-in-Aid for Scientific Research.

REFERENCES

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698.

Chang, C.-C. and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27. Software available at http://www.csie.ntu.edu.tw/∼cjlin/libsvm.

Chen, X. and Koskela, M. (2014). Using appearance-based hand features for dynamic RGB-D gesture recognition. In *Proceedings of the 22nd International Conference on Pattern Recognition*, pages 411–416.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 886–893.

Gandhi, T. and Trivedi, M. M. (2008). Image based estimation of pedestrian orientation for improving path prediction. In *Proceedings of 2008 IEEE Intelligent Vehicles Symposium*, pages 506–511.

Hu, W., Tan, T., Wang, L., and Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on System, Man, and Cybernetics, Part C: Applications and Reviews*, 34(3):334–352.

Hu, Y., Cao, L., Lv, F., Yan, S., Gong, Y., and Huang, T. S. (2009). Action detection in complex scenes with spatial and temporal ambiguities. In *Proceedings of 12th IEEE International Conference on Computer Vision*, pages 128–135.

Liu, W., Zhang, Y., Tang, S., Tang, J., Hong, R., and Li, J. (2013). Accurate estimation of human body orientation from RGB-D sensors. *IEEE Transaction on Cybernetics*, 43(5):1442–1452.

Pestschnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M., and Toyama, K. (2004). Digital photography with flash and no-flash image pairs. *ACM Transaction on Graphics*, 23(3):664–672.

Shotton, J., Girshick, R., Fitzgibbon, A., Sharp, T., Cook, M., Finocchio, M., Moore, R., Kohli, P., Criminisi, A., Kipman, A., and Blake, A. (2013). Efficient human pose estimation from single depth images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2821–2840.

Straka, M., Hauswiesner, S., Rüther, M., and Bischof, H. (2011). Skeletal graph based human pose estimation in real-time. In *Proceedings of the 22nd British Machine Vision Conference*, pages 69.1–69.12.

Weinrich, C., Vollmer, C., and Gross, H.-M. (2012). Estimation of human upper body orientation for mobile robotics using an SVM decision tree on monocular images. In *Proceedings of 2012 IEEE/RS International Conference on Intelligent Robots and Systems (IROS)*, pages 2147–2152.

Yang, Q., Ahuja, N., Yang, R., Tan, K.-H., Davis, J., Culbertson, B., Apostolopoulos, J., and Wang, G. (2013). Fusion of median and bilateral filtering for range image upsampling. *IEEE Transactions on Image Processing*, 22(12):4841–4852.