Person following control for a mobile robot based on color invariance corresponding to varying illumination

Shinsuke Oh-hara[1], Kaoru Saito[2], Atsushi Fujimori[1]
[1]Department of Mechanical Engineering, Faculty of Engineering, University of Yamanashi, Kofu, Japan
[2]Mechanical Co. Ltd., Yamanashi, Japan

Article Info	ABSTRACT
	In this paper, we present a method of person following control for a mobile robot using visual information. Color information is often used for object tracking. Color information of objects varies greatly under illumination changing environment. In such conditions, the robot controlled by visual information may lose sight of a person. In this paper, we consider a robust person following method by color invariance and image-based control. Color invariance shows robust features of colored objects in terms of changing illumination conditions. At first, we estimate the lowest positions of both feet of a tracked person through particle filters based on color invariances. Then, we control the velocity of the robot to track the person by using the image-based controller. Experimental results using an actual robot demonstrate the effectiveness of the proposed method. *This is an open access article under the [CC BY-SA](#) license.*

Corresponding Author:

Shinsuke Oh-hara
Department of Mechanical Engineering, Faculty of Engineering, University of Yamanashi
Kofu, Yamanashi, Japan
Email: sohhara@yamanashi.ac.jp

1. INTRODUCTION

Recent robotics technology is expanding not only in special environments such as manufacturing and space but also in various fields such as medical treatment, welfare, and entertainment, because of its remarkable progress [1]–[3]. Examples of service robots include carts that automatically follow customers in supermarkets and that carry nursing care products for caregivers. Such robots require functions that identify the target person and track it automatically.

In order to identify and automatically track a person, it is important to recognize the person and estimate the position and direction, such as the distance between that person and the robot. Some researchers are implementing such recognizing and tracking functions by using distance sensors such as a laser range finder (LRF) [3], [4], and camera [5]–[7]. Okusako *et al.* [3] estimated the foot position of a person by comparing the data obtained from an LRF with the templates consisting of the arrangement data of the foot prepared beforehand. Leigh *et al.* [4] proposed the method for robust detecting a person's leg using an LRF in indoor and outdoor environments. Camera-based person tracking techniques have been actively carried out in the field of image recognition [5]–[7]. Perez *et al.* [5] made a color model of an object from color information of an image captured by a camera and proposed a method for tracking the object based on the model. Rincon *et al.* [6] proposed a method for estimating the position and posture of a person using a particle filter based on image information obtained from a camera. Ahmad *et al.* [7] developed the person tracking algorithm using a convolution neural network for overhead view camera. From the developments in image recognition, camera-based person following robots have been widely studied in many publications, such as monocular camera [2],

[8], [9] stereo camera [10] red green blue, and depth (RGB-D) camera [11]. Moreover, Koide *et al.* [12] developed a person-following robot based on the fusion of camera and LRF.

Nakano *et al.* [2] proposed a method for recognizing the both feet of a target person based on image information and making it follow the robot. In this method, only the feet need to be recognized, so it is not necessary for the camera to be mounted high. Thus, the distance between the tracked person and the robot can be reduced, and safety to the user is ensured by lowering the center of gravity of the robot. It has simplicity of user and environment learning such as inexpensiveness by introducing a commercial camera and identification of a person only by color information of the feet. In this paper, we proposed a method to recognize the person's feet from the image information of a camera.

Nakano *et al.* [2] convert from the image position of the recognized feet to the three-dimensional position information and utilize the position for the control. Such a control method using three-dimensional position obtained from the image information is called the position-based method. The position-based method is easy to implement the control law. However, it is known that three-dimensional restoration from camera images is vulnerable to model errors in camera parameters and is also susceptible to image quantization. If there is such an error in the camera parameters, the position-based method cannot keep the distance to the person properly, and there is a possibility that the tracking fails.

Though person or object recognition based on the color information of the camera image is easy to implement, the apparent color on the screen changes when the illumination brightness of the environment changes, and a large difference occurs from the original color information of the tracking object. This causes the mobile robot to mistakenly lose sight of the person to be tracked.

In this paper, we consider a person tracking control based on an image-based method and color invariance [13]–[15]. The image-based method is one of the visual feedback controls and utilizes the state variables on the camera image plane. By directly controlling the features of the target on the image, it becomes unnecessary to carry out the three-dimensional restoration, and it has the robustness for the model error of the camera parameter [16]–[18]. The color invariance also gives a robust feature quantity against apparent changes such as illumination lightness and shadow. In this paper, we extract the features of a person based on the color invariance and utilize them for tracking control by a mobile robot. The effectiveness of the proposed method is evaluated through indoor experiments. In the experiment, one camera is mounted on a mobile robot to reflect the feet of the tracked person, and both feet are estimated from two independent particle filters from the camera image. Experimental evaluations are carried out by the mobile robot in an indoor environment. We demonstrate the effectiveness of the proposed method by measuring traceable illuminance values using an illuminance sensor.

2.   IMAGE-BASED CONTROL WITH MOBILE ROBOT

This paper presents a model of a two wheel-type mobile robot equipped with a camera and realizes tracking control by image-based method for the robot based on [17].

2.1. Modeling for mobile robots with camera

Figure 1 shows the model of the mobile robot. In this model, it is assumed that the camera is mounted horizontally on the center of the left and right wheels toward the traveling direction of the mobile robot. v denotes the translational speed of the mobile robot and ω denotes the rotational speed. Also, we assume that the camera coordinate system $O' - X'Y'Z'$ with the center of the wheel as the origin. Here, if the camera is mounted horizontally at the center of the wheel, the camera must be placed in a high position in order to view the feet of the person. Therefore, in this paper, we consider that the actual camera is attached to the front of the robot and tilted downward as shown in Figure 2. This camera-coordinate system is assumed to be $O - XYZ$. The image coordinate system for the camera coordinate system $O' - X'Y'Z'$ and $O - XYZ$ is defined as $O'_i - x'y'$ and $O_i - xy$, respectively. The origin of each image coordinate system shall be the image center. The coordinate system $O - XYZ$ of the camera tilted downward can be converted to the camera coordinate system $O' - X'Y'Z'$ on the center of the wheel by applying a translation vector and a rotation matrix. For the sake of simplicity, consider the kinematic model in the horizontal camera coordinate system $O' - X'Y'Z'$. Target point P is assumed to be always on the ground.
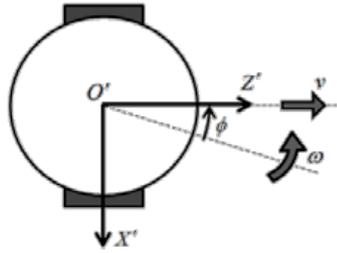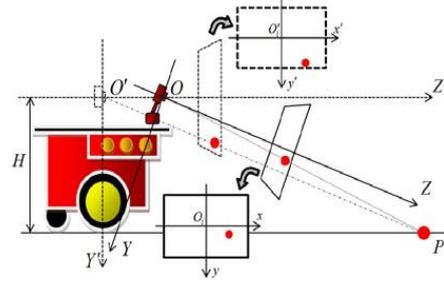
Figure 1. Mobile robot model

Figure 2. A mobile robot equipped with a camera and image planes

The kinematic model of point $P = [X', H, Z']^T$ with respect to the mobile robot is expressed by (1).

$$\frac{d}{dt}\begin{bmatrix} X' \\ H \\ Z' \end{bmatrix} = -V - \Omega \times \begin{bmatrix} X' \\ H \\ Z' \end{bmatrix} \tag{1}$$

Where $V = [0 \quad 0 \quad v]^T$ and $\Omega = [0 \quad -\omega \quad 0]^T$. In addition, since P is on the ground, (1) can be expressed as (2).

$$\frac{d}{dt}\begin{bmatrix} X' \\ Z' \end{bmatrix} = \begin{bmatrix} 0 & Z' \\ -1 & -X' \end{bmatrix}\begin{bmatrix} v \\ \omega \end{bmatrix} \tag{2}$$

In the image-based method, the control is performed using the coordinates on the image. Next, we transform P in the coordinate system $O' - X'Y'Z'$ to the target point on the image plane. The vector $[x' \quad y']^T$ of P on the image plane are obtained by the following.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = f\begin{bmatrix} X'/Z' \\ H/Z' \end{bmatrix} \tag{3}$$

Where f is the focal length of the camera.

Substituting the time-differentiated (3) into (2), the motion of the target vector on the image plane can be expressed by (4).

$$\dot{s} = Ju \tag{4}$$

Where,

$$s = \begin{bmatrix} x' \\ y' \end{bmatrix}, J = \begin{bmatrix} \dfrac{x'y'}{fH} & \dfrac{f^2 + x'^2}{f} \\ \dfrac{y'^2}{fH} & \dfrac{x'y'}{fH} \end{bmatrix}, u = \begin{bmatrix} v \\ \omega \end{bmatrix}$$

The vector s is the controlled variable. In this paper, we refer to the variable s as the image feature point. The matrix J containing the components of s is called the image Jacobian matrix and is a regular matrix unless y' is 0.

2.2. Image-based tracking control

In this paper, we control the mobile robot so that the image feature point on the image plane coincides with the target point. Let s^* be the target value for the tilted camera image coordinates system $O_i - xy$. The control error e can be expressed as (5).

$$e = s - s^* \tag{5}$$

Then, let e' be the e error converted to the horizontal camera coordinate system $O'_i - x'y'$. When e' is differentiated with respect to time, we can obtain (6).

$$\dot{e}' = Ju \tag{6}$$

The matrix J is a regular matrix as long as $y' \neq 0$. That is, if P is not at the same height as the camera, J is regular. In our problem setting, if the camera tilted down can recognize the target P in the image, J is always regular and the inverse matrix exists. Therefore, the following control law is used for the speed control of the mobile robot.

$$u = -\lambda J^{-1} e' = \begin{bmatrix} \lambda_v & 0 \\ 0 & \lambda_\omega \end{bmatrix} J^{-1} e' \tag{7}$$

Where λ_v and λ_ω are positive feedback gains related to speed and angular speed, respectively. From the above, the closed-loop system of the mobile robot is expressed as follows:

$$\dot{e}' = -\lambda e' \tag{8}$$

It can be seen that it can be stabilized by the control law of (7). In this paper, we detect the lower end positions of both feet of a person from the image of the camera and then determine the average of the two positions as the image feature point s. The image Jacobian matrix J is calculated from the point s in real-time, and the person tracking is realized by the control law of. (7).

3. FEET RECOGNITION AND TRACKING OF PERSON BASED ON COLOR INVARIANCE
3.1. Color invariance
The color invariance [13] is robust to apparent changes such as shadows and illumination lightness. By deriving spectral derivatives from the physical model of the reflected light spectrum of Kulbelka-Munk, two invariances, H and C, are calculated. In addition, the invariances H and C are calculated from the linear transformations of RGB color information of the images and the differential coefficients of the reflected spectra as (9) and (10).

$$H = \frac{0.30R + 0.04G - 0.35B}{0.34R - 0.60G + 0.17B} \tag{9}$$

$$C = \frac{0.30R + 0.04G - 0.35B}{0.06R + 0.63G + 0.27B} \tag{10}$$

The invariance H is the amount due solely to the color of the object's surface, and C is the amount with respect to non-glossy objects. If the values of H and C are constant, they represent a plane of RGB space. Since the values of H and C can be infinite, we use the following inverse trigonometric functions for H and C to finite them.

$$\begin{cases} \theta_H = \tan^{-1} H \\ \theta_C = \tan^{-1} C \end{cases} \tag{11}$$

Figure 3 shows the image obtained by converting the original one when the light is turned off with H and C and displaying it on a grayscale. Figure. 3(a) shows the image captured in an illuminated room. We can clearly recognize the color and the pattern on the cover of the object. On the other hand, Figure. 3(b) shows the image taken in the room with lights off. It is difficult to recognize the color and the pattern from the image. The image converted by H in Figure. 3(c) can be seen that the characteristics of the object become clearer. On the other hand, C shown in Figure. 3(d) shows a small change in the whole image, which obscures the characteristics of the object from the image converted by H. This suggests that the use of C together may reduce the recognition of the target, so it is effective to use H proactively. However, there is a problem that H cannot distinguish its opposite color for a given color, and Kobayashi *et al.* [14] have made the opposite color correspond to a different value by dividing the value of H based on C. In this paper, the following invariance H' is obtained by using the invariant $C_0 = C \approx -0.01$ as the standard in the same way as [14].
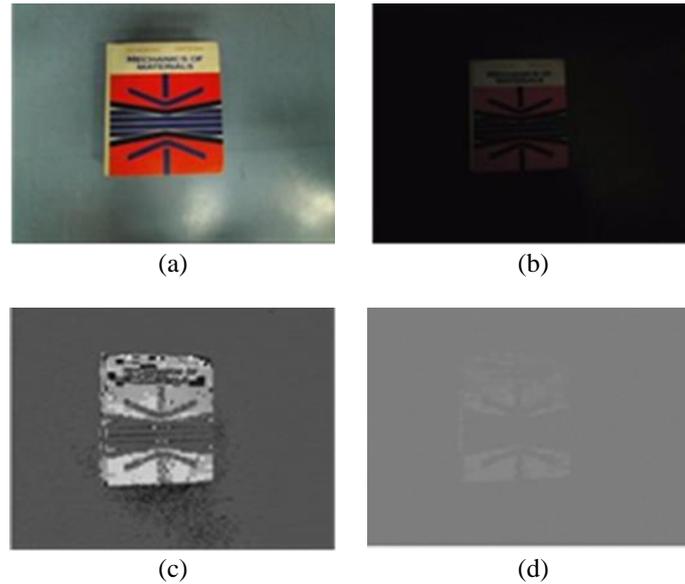
Figure 3. Examples of transferred images using color invariance are (a) original image with light, (b) original image without light, (c) color invariance H, (d) color invariance C

$$\theta_{H'} = \begin{cases} \frac{1}{2}(\theta_H - \pi) \ (\theta_c \geq \theta_{C_0}, \theta_H \geq 0) \\ \frac{1}{2}\theta_H \ (\theta_c \geq \theta_{C_0}, \theta_H < 0) \\ \frac{1}{2}\theta_H \ (\theta_c < \theta_{C_0}, \theta_H \geq 0) \\ \frac{1}{2}(\theta_H + \pi) \ (\theta_c < \theta_{C_0}, \theta_H < 0) \end{cases} \tag{12}$$

In this paper, the invariance H' of (12) is utilized as a feature quantity of the foot of a tracking person. There is a hue as an index to express the change of color, and it shows a robust property for illumination lightness and so on. Using this property, Ri and Fujimoto [19] proposed a person tracking method that fuses hue and location. However, a comparative study by Matsumoto *et al.* [15] shows that the invariance H' is more robust to illumination changes than the hue. Therefore, in this paper, we use the color invariance H' for person tracking.

3.2. Person tracking based on particle filter

Figure 4 shows a block diagram of person tracking in this paper. In this paper, based on the method of [2] we implement independent two-particle filters on both feet of a tracked person. The following shows the procedure of the person tracking by using a camera.
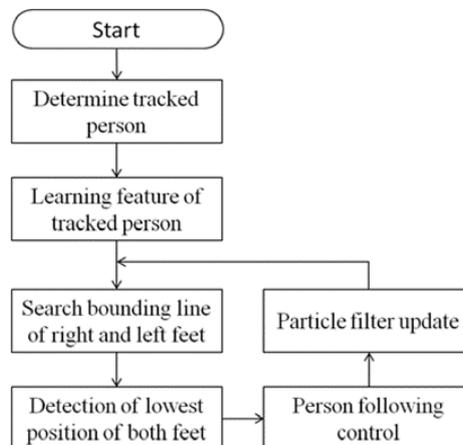

Figure 4. Block diagram of the proposed method

3.2.1. Determination of a tracked person

First, a mobile robot is kept on standby, and when a person intrudes in front of the camera, a tracking person is decided by background difference.

3.2.2. Learning feature of the tracked person

The color information obtained from the image of the tracked person is converted into invariances H and C, and the invariance H' is calculated based on (12). We also create and normalize a histogram of the invariance H' to take advantage of this for the likelihood of particle filters. Along with this, the particle filters are initialized.

3.2.3. Search bounding line for both feet

Based on the method of [2] the likelihood of the particles of each particle filter is projected and accumulated in the x-axis direction as (13).

$$Acc(x) = \sum_{n=1}^{N} w_l^{(n)} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{\left(x-x_l^{(n)}\right)^2}{2\sigma}\right) + \sum_{n=1}^{N} w_r^{(n)} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{\left(x-x_r^{(n)}\right)^2}{2\sigma}\right) \qquad (13)$$

Where N is the number of the particle, $x^{(n)}$ is the state and $w^{(n)}$ is the corresponding likelihood of the n-th particle of each particle filter, respectively. The subscripts l and r denote for the left and right foot, respectively, First, we find the two centers of gravity of all particles for each particle filter. Then, the x-value where $Acc(x)$ is the minimum between the x-values of each center of gravity is used as the dividing line.

3.2.4. Searching for the lowest points of both feet

We determine the lowest positions of both feet on an image by utilizing both foot dividing lines and the likelihood of particles. The x-coordinate of the center of gravity of the particles for each filter is used as the x-coordinate of the lowest position. In this system, the likelihood of a particle with a negative y-coordinate trend is high because the foot is projected above the image plane. Therefore, we examine the likelihood of each particle from the upper part of the image plane, if a particle is in Table the set threshold, we regard the y-coordinate of the particle as the lowest point of the foot. Figure 5 shows the result of searching the lowest point of both feet obtained by this method. The green dots are the bottom point of both feet, the bars on both sides of the image indicate the magnitude of the likelihood in the y-axis direction of each particle, and the bars on the bottom indicate the magnitude of the likelihood in the x-axis direction. Blue is the right foot, and red is the likelihood of the left foot.

Figure 5. Example of the feature points extracted by the proposed method

3.2.5. Tracking control by an image-based method

The average value of the obtained lowest positions of both feet is set as the feature s, and s is converted into the coordinates of the horizontal camera image coordinate system $O' - X'Y'Z'$. The state variable is applied to the control law in (7).

3.2.6. Particle filter update

In the particle filter update, the resampling of the particles is performed. The process returns to the two-leg dividing line search in the procedure (3.2.3). In this paper, we assume a constant velocity motion model

as a model of the particle filter. The mobile robot is controlled to follow the person by repeating these procedures.

The particle number of each foot particle filter in this paper is 1200. The number of bins in the histogram for the invariance H' is 1000, and the threshold of the likelihood for determining the lowest point is set to 10^{-5}. The target value for person tracking control is set at the center of the tilted camera image, that is, the origin. The gain λ_v and λ_ω in the controller of the (7) are set to 0.6 and 1.8, respectively.

In the tracking procedure (3.2.4), Nakano *et al.* [2] extracted the lower edges from the position of the center of gravity of the particles, then determined the lowest points by edge detection. However, in this method, the foot image is blurred due to the fast movement of the foot and the change of the illumination brightness, and it becomes difficult to find the lowest point by the edge detection. Therefore, the point is extracted at a position different from the actual foot, and it becomes impossible to keep the distance between the person and the robot properly. On the other hand, the likelihood-based method in this paper makes it possible to find the endpoints even for blurred images.

4. EXPERIMENTAL EVALUATIONS

Pioneer 3DX [20] was fitted with a universal serial bus (USB) ELECOM Corporation Universidad Católica San Antonio de Murcia (UCAM)-C 0220 to reflect the feet of a person, see Figure 6. Dell latitude 3550 (4 GB) laptops and Microsoft visual C++ 2010 express were used to implement the proposed method, and OpenCV-2.1 was used to process images. The proposed particle filters were implemented based on the condensation algorithm [21] in OpenCV. The illuminance sensor Graphtec GS-LXUV was mounted to measure the ambient illumination during experiments.

Figure 6. Mobile robot pioneer 3DX

The experiments were carried out in an indoor environment. The person went straight to the hallway where the lights were turned off when the person started walking. Figure 7 shows an example of the results of measurement using an illuminance meter in an experiment. The experimental environment changes from a maximum illuminance of 27.4 [lx] to a minimum illuminance of 1.8 [lx].

Figure 7. Illumination condition in the experimental environment

The proposed method is compared with the RGB color information method for tracking a person's feet. The method of Nakano et *al.* [2] is based on RGB color information. We show the process of tracking a person by using the onboard camera when using the RGB color-based method and the proposed method, respectively. The lightness of each image is processed to make it easier to see the experiment. Note that the actual images are darker than them.

Figure 8 shows the RGB color-based method in which particles are dispersed in the shadows of the feet as the illumination intensity decreases. In Figure 8(a), the robot specified the target person based on the background difference and determined the lowest positions. In Figures 8(b) and (c), the robot recognized the feet of the person and followed it since the lighting in the hallway was bright. The particles began to disperse even toward the shadow of the foot from Figure 8(d), and the distance between the person and the robot began to increase. Because the color of the person's pants was dark blue, the particles in the shaded area were given likelihood. In Figure 8(e), the shadows of both feet were extracted as the lower endpoints, even though the feet were captured by the camera. The robot lost sight of the feet and could not track the person in Figure 8(f). The illuminance in Figure 8(f) was 7.2 [lx], where the foot could not be recognized and tracking became impossible from the images shown in Figure 8(f) and later.

In the proposed method shown in Figure 9, the particles did not disperse in the shadows of the feet even in situations where both feet were difficult to recognize due to the darkness. Figure 9 shows that the color invariance is robust against shadows. In Figure 9(a), the robot specified the target person by the background difference and determined the lowest positions. In Figure 9(b), it was equivalent to the brightness when tracking fails using a color information-based method. In the proposed method, a group of highly likelihood particles was distributed around both feet, and the lower endpoints of both feet were accurately detected. From Figures 9(c), (d), and (e), the robot recognized both feet and tracked the person, even though the person was walking into the darker corridor. In Figure 9(f), the lower endpoints of both feet were correctly extracted in the darkest place. The illuminance at Figure 9(f) was 2.0 [lx], which was in a darker environment than in the case of color information.

In this experiment, a color information-based method and the proposed method were performed 50 times each, and the results were measured by an illuminance meter. Table 1 shows the mean and standard deviation of minimum illumination where the robot could follow the person.
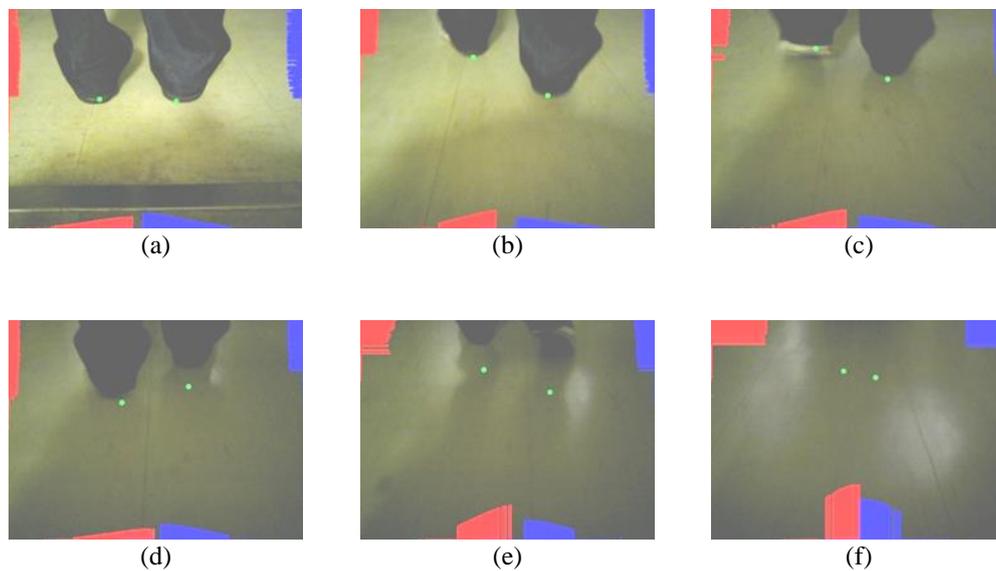

| (a) | (b) | (c) |
| (d) | (e) | (f) |

Figure 8. Scene of tracking based on RGB color information are (a) T=0.0 [s], (b) T=6.8 [s], (c) T=12.2 [s], (d) T=15.8 [s], (e) T=23.5 [s], (f) T=28.7 [s]
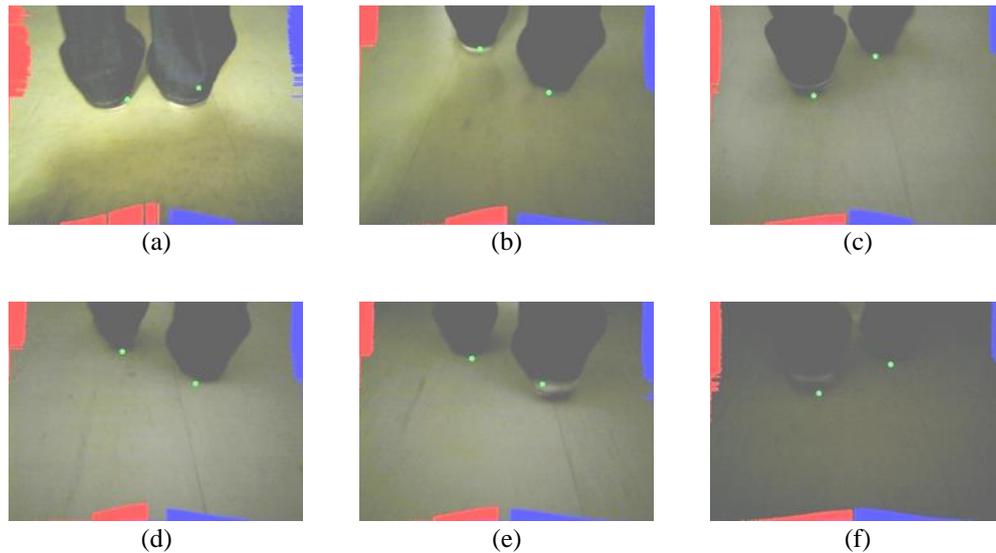
Figure 9. Scene of tracking based on color invariance are (a) T=0.0 [s], (b) T=10.3 [s], (c) T=40.9[s], (d) T=47.5 [s], (e) T=54.7 [s], (f) T=58.3 [s]

Table 1. Minimum illumination

	Minimum illumination [lx]
RGB color space	7.2 ± 0.4
Color invariance	2.0 ± 0.1

5.   CONCLUSION

In this paper, we consider the recognition of the person's feet and tracking control by a camera mounted on a mobile robot under illumination variation. In this paper, we proposed an image-based tracking control method, which is robust visual feedback control. Furthermore, in order to correspond to the change of the lighting environment, the color information of the image is converted to the color invariant, and the person tracking method using it is proposed. The effectiveness of the proposed method was evaluated by experiments with an actual mobile robot. The measurement of the illuminance value by the illuminance sensor was also carried out in the experiment, and it was shown that the person could be tracked even in the environment in which the lighting brightness was low by the proposed technique.

REFERENCES

[1]   N. Matsuhira, H. Ogawa, and T. Yoshimi, "Life support robots that coexist in harmony with people (in Japanese)," *TOSHIBA Rev.*, vol. 60, no. 7, pp. 112--115, 2005.

[2]   H. Nakano, Y. Shimowaki, T. Yamanaka, and M. Watanabe, "Person following system for the autonomous mobile robot by independent tracking of left and right feet (in Japanese)," *J. Robot. Soc. Japan*, vol. 25, no. 5, pp. 707–716, 2007, doi: 10.7210/jrsj.25.707.

[3]   S. Okusako and S. Sakane, "Human tracking with a mobile robot using a laser range-finder (in Japanese)," *J. Robot. Soc. Japan*, vol. 24, no. 5, pp. 605–613, 2006, doi: 10.7210/jrsj.24.605.

[4]   A. Leigh, J. Pineau, N. Olmedo, and H. Zhang, "Person tracking and following with 2D laser scanners," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 726–733, doi: 10.1109/ICRA.2015.7139259.

[5]   P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color based probabilistic tracking," in *European Conference on Computer Vision*, 2002, pp. 661–675.

[6]   J. M. del Rincón, D. Makris, C. O. Uruñuela, and J.-C. Nebel, "Tracking human position and lower body parts using kalman and particle filters constrained by human biomechanics," *IEEE Trans. Syst. Man, Cybern. Part B*, vol. 41, no. 1, pp. 26–37, Feb. 2011, doi: 10.1109/TSMCB.2010.2044041.

[7]   M. Ahmad, I. Ahmed, F. A. Khan, F. Qayum, and H. Aljuaid, "Convolutional neural network based person tracking using overhead views," *Int. J. Distrib. Sens. Networks*, vol. 16, no. 6, p. 155014772093473, Jun. 2020, doi: 10.1177/1550147720934738.

[8]   A. E. Guevara, A. Hoak, J. T. Bernal, and H. Medeiros, "Vision-based self-contained target following robot using bayesian data fusion," in *International Symposium on Visual Computing*, 2016, pp. 846–857.

[9]   K. Koide, J. Miura, and E. Menegatti, "Monocular person tracking and identification with on-line deep feature selection for person

Person following control for a mobile robot based on color invariance corresponding … (Shinsuke Oh-hara)

following robots," *Rob. Auton. Syst.*, vol. 124, p. 103348, Feb. 2020, doi: 10.1016/j.robot.2019.103348.

[10] B. X. Chen, R. Sahdev, and J. K. Tsotsos, "Person following robot using selected online ada-boosting camera," in *2017 14th Conference on Computer and Robot Vision (CRV)*, 2017, pp. 48–55, doi: 10.1109/CRV.2017.55.

[11] M. Munaro and E. Menegatti, "Fast RGB-D people tracking for service robots," *Auton. Robots*, vol. 37, no. 3, pp. 227–242, Oct. 2014, doi: 10.1007/s10514-014-9385-0.

[12] K. Koide and J. Miura, "Identification of a specific person using color, height, and gait features for a person following robot," *Rob. Auton. Syst.*, vol. 84, pp. 76–87, Oct. 2016, doi: 10.1016/j.robot.2016.07.004.

[13] J.-M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 12, pp. 1338–1350, 2001, doi: 10.1109/34.977559.

[14] M. Kobayasshi and K. Kameyama, "A composite illumination invariant color feature and its application to partial image matching," *IEICE Trans. Inf. Syst.*, vol. E95.D, no. 10, pp. 2522–2532, 2012, doi: 10.1587/transinf.E95.D.2522.

[15] W. Matsumoto and K. Kameyama, "Comparison of hue feature and illumination invariant color feature h' in partial image retrieval (in Japanese)," *IEICE Tech. Report; IEICE Tech. Rep*, vol. 112, no. 335, pp. 41–46, 2012.

[16] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Autom.*, vol. 12, no. 5, pp. 651–670, 1996, doi: 10.1109/70.538972.

[17] K. Kurashiki, J. Nagata, and T. Fukao, "Image-based Stabilizing Control of a Nonholonomic Mobile robot not to go beyond targets (in Japanese)," *Trans. Inst. Syst. Control Inf. Eng.*, vol. 25, no. 4, pp. 75–83, 2012, doi: 10.5687/iscie.25.75.

[18] Y. Ri and H. Fujimoto, "Image Based Visual Servo Application on Video Tracking with Monocular Camera Based on Phase Correlation Method," in *IEEJ International Workshop on Sensing, Actuation, Motion Control, and Optimization*, 2017.

[19] G. M. Y. Isobe and K. Umeda, "Robust target detection for human following with a mobile robot under varying illumination environments (in Japanese)," in *Proceedings of 2014 JSPE Autumn Conference*, 2014, pp. 163–164, doi: https://doi.org/10.11522/pscjspe.2014A.0_163.

[20] ActivMedia Inc., "Pioneer 3 operations manual, version 2," in *ActivMedia Inc.*, 2015.

[21] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," *Int. J. Comput. Vis.*, vol. 29, no. 1, pp. 5–28, 1998.

BIOGRAPHIES OF AUTHORS

Shinsuke Oh-hara 🆔 📷 SC P revised the Ph.D. degree in mechanical engineering from Osaka University in Japan in 2006. Since 2009, he has been working at University of Yamanashi as an Assistant. His research interrests include visual feedback control and mobile robot control. He can be contacted at email: sohhara@yamanashi.ac.jp

Kaoru Saito 🆔 📷 SC P revised the Master degree in mechanical engineering form University of Yamanashi (Japan) in 2016. Since April 2016, he has been working at Autec Mechanical Co. Ltd. He can be contacted at email: ah.r.sora@gmail.com

Atsushi Fujimori 🆔 📷 SC P received the Ph.D. degree in aeronautics from Nagoya University in Japan in 1989. Since 1990, He had worked at Shizuoka University in Japan as an Associate Professor. Since 2006, He have been working at University of Yamanashi as a Professor. His current research interests are in fields of system identification, flight control and mobile robots. He can be contacted at email: afujimori@yamanashi.ac.jp