International Journal of New Computer Architectures and their Applications (IJNCAA) 7(4): 112-122
The Society of Digital Information and Wireless Communications, 2017 ISSN 2220-9085 (Online); ISSN 2412-3587 (Print)

An Object Detection Method Using Invariant Feature Based on Local Hue Histogram in Divided Areas of an Object

Tomohiro Kanda
Department of Information Systems Science, Graduate School of Engineering, Soka University
1-236, Tangi-machi, Hachiouji-Shi, Tokyo, Japan 192-8577
e16m5206@soka-u.jp

Kazuo Ikeshiro
Department of Information Systems Science, Graduate School of Engineering, Soka University
1-236, Tangi-machi, Hachiouji-Shi, Tokyo, Japan 192-8577
ikeshiro@soka.ac.jp

Hiroki Imamura
Department of Information Systems Science, Graduate School of Engineering, Soka University
1-236, Tangi-machi, Hachiouji-Shi, Tokyo, Japan 192-8577
imamura@soka.ac.jp

ABSTRACT

In recent years, the decreasing birthrate and aging population are developing, and there is concern about the lack of labor power such as household chores and nursing care at home. Therefore, application of robot technology to the living field is expected. In the living field, Robots that support the lives of people are collectively referred to as life support robots. This robot is required to perform various tasks to support humans. Especially, the object detection task is important when people request the robot to transport and rearrange objects. However, when detecting an object from the camera mounted on the robot, detection becomes difficult because the detection environment is unspecified. Scale Invariant Feature Transform (SIFT) and Color Indexing are widely known as object detection methods using two-dimensional information. However, these methods do not have robustness against all environmental changes. In this research, we focus on the invariant feature of the hue histogram in divided areas of an object and propose a highly accurate object detection method.

KEYWORDS

Life support robot, Cognitive system, Hue histogram, Position of peak and trough, divided areas.

1 INTRODUCTION

In recent years, the decreasing birthrate and aging population are developing, and there is concern about the lack of labor power such as household chores and nursing care at home. Therefore, application of robot technology to the living field is expected. In the living field, robots that support the lives of people are collectively referred to as life support robots [1][2][3]. This robot is required to perform various tasks to support humans. Especially, the object detection task is important when people request the robot to transport and rearrange objects.

(a) A target object (b) An input image

Figure. 1. An example of input data in object detection.

The object detection is technology to detect a target object (Fig. 1 (a)) from an input image (Fig. 1 (b)).

There is a problem which the detection becomes difficult when detecting a target object from the camera mounted on the robot, because differences of visual appearance occur such as the rotation change. Therefore, we consider that there are six necessary properties to detect in domestic environment as follows.

1. Robustness against the rotation change
2. Robustness against the scale change
3. Robustness against the illumination change
4. Robustness against the distortion by perspective projection
5. Robustness against the occlusion
6. Detecting an object which has few textures

Firstly, the robots need the robust detection for the rotation change because the rotation change is occurred in case that an object falls down such as Fig. 1 (b). Secondly, the robots need the robust

detection for the scale change because the scale change is occurred by the position between the robots and an object. Thirdly, the robots need the robust detection for the illumination change because the precision of detection is easy to be affected by illumination condition. Fourthly, the robots need the robust detection for the distortion by perspective projection because the distortion by perspective projection is occurred in case that the robot moves sideways, looks up or looks down. Fifthly, the robots need the robust detection for occlusion because occlusion is occurred between different objects. Finally, the robots need to detect an object which has few textures. As an example, Fig. 2 shows an object which has few textures.

As conventional methods, there are the Scale Invariant Feature Transform (SIFT) [4] and the Color Indexing [5]. Table 1 shows properties of these methods.

To detect a target object from an input image, the SIFT compares local feature amount of a target object with local feature amount of an input image, and extracts correspondence points as shown in Fig. 3. Therefore, as shown item 1-3 and 5 in Table 1, the SIFT has robustness against the rotation change, the scale change, the illumination change and occlusion. However, as shown item 4 in Table 1, the SIFT does not have robustness against the distortion by perspective projection that local feature amount

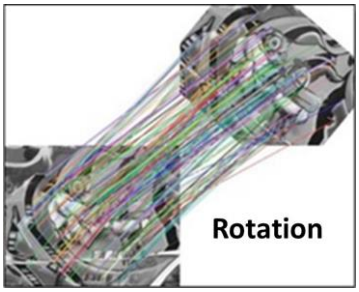


Figure. 3. An example of matching based on the feature of SIFT.

changes. In addition, as shown item 6 in Table 1, the SIFT is difficult to detect an object which has few textures because the SIFT cannot detect feature amount from objects such as shown in Fig. 2 which have few textures.

On the other hand, Swain et al. proposed the Color Indexing. To detect a target object from an input image, the Color Indexing compares three dimensional color histogram of a target object with three dimensional color histogram of a candidate object. As an example, Fig. 4 shows three dimensional color histogram. As shown in Fig. 4, a size of the square in three dimensional color histogram expresses the frequency of each color. The frequency of each color does not change in case that the rotation change, the scale change, the distortion by perspective projection and occlusion are occurred. In addition, the Color Indexing can

Figure. 2. An example of an object with few textures.

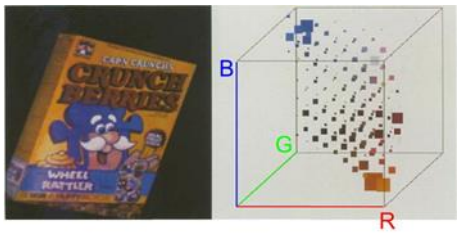


Figure. 4. An example of a three-dimensional color histogram.

Table 1. Properties of conventional methods and the proposed method.

	1. Rotation change	2. Scale change	3. Illumination change	4. Distortion by perspective projection	5. Occlusion	6. An object which has few textures	7. Different objects which have same features of histogram
SIFT	○	○	○	×	○	×	○
Color indexing	○	○	×	○	○	○	○
Tanaka's method	○	○	○	○	○	○	×
Our proposed method	○	○	○	○	○	○	○

detect an object which has few textures by using the color information. Therefore, as shown item 1, 2 and 4-6 in Table 1, the Color Indexing has robustness against the rotation change, the scale change, distortion by perspective projection, occlusion and an object which has few textures. However, as shown item 3 in Table 1, the Color Indexing is difficult to detect in case that illumination is changed, because the Color Indexing uses RGB color system which is easy to be affected by the illumination change.

Tanaka et al. [6] focused on hue. Hue has three merits. Firstly, hue is hardly affected by illumination and shadow. Secondly, hue is invariable against the distortion by perspective projection. Finally, the detection using hue can detect an object which has few textures. In addition, as shown in Fig. 5, the positions of peak and trough of the hue histogram is invariable against the rotation change, the scale change, the illumination change, the distortion by perspective projection and occlusion. In Fig. 5, we define the longitudinal axis as the frequency, and the lateral axis as the hue value. In addition, the circle indicates the position of peak and the square indicates the position of trough. Accordingly, Tanaka et al. proposed the object detection method using invariant feature based on the hue histogram. To detect a target object from an input image, this method extracts the positions of peak and trough of the hue histogram as invariant feature. In addition, this method compares invariant feature of a target object with invariant feature of a candidate object. As mentioned above, the positions of peak and trough of the hue histogram do not change in case that the rotation change, the scale change, the illumination change, the distortion by perspective projection and occlusion are occurred. In addition, this method can detect an object which has few textures by using hue information. Therefore, as shown item 1-6 in Table1, this method has robustness against the rotation change, the scale change, the illumination change, the distortion by perspective projection, occlusion and an object which has few textures. However, this method does not satisfy a property as follows.

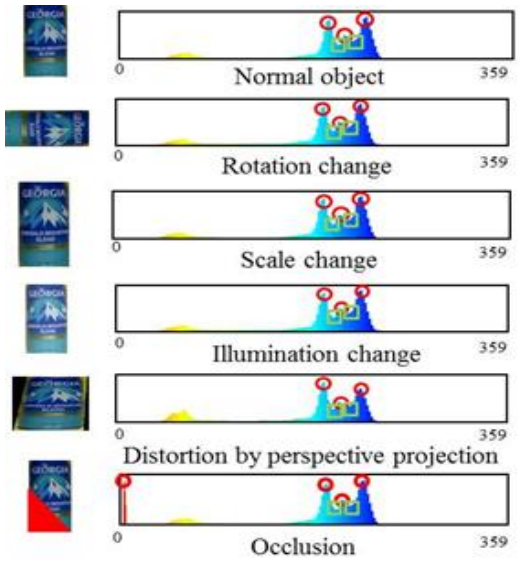


Figure. 5. The invariance of the position of the unevenness of the hue histogram for each change.

7. Distinguishing different objects which have the same features of the hue histogram

This method cannot distinguish objects which have the different textures but have the same positions of peak and trough of the hue histogram. As an example, Fig. 6 (a), (b) shows different objects which have the same features of the hue histogram. Fig. 6 (c), (d) shows the hue histogram of Fig. 6 (a), (b) respectively. As shown item 7 in Table 1, this method cannot distinguish objects (Fig. 6 (a), (b)) as different object because the positions of peak and trough of Fig. 6 (a) equal the positions of peak and trough of Fig. 6 (b) as shown in Fig. 6 (c), (d). As mentioned above, as shown in Table 1, the SIFT, the Color Indexing and this method do not satisfy all seven properties.

Therefore, to satisfy the seven properties for detection, we propose an object detection method using invariant feature based on the local hue histogram in divided areas of an object instead of invariant feature based on the hue histogram in the

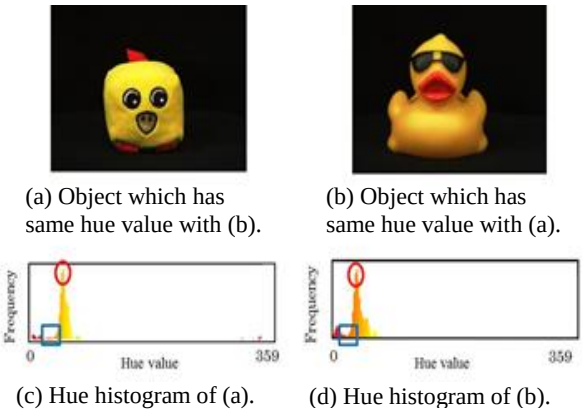
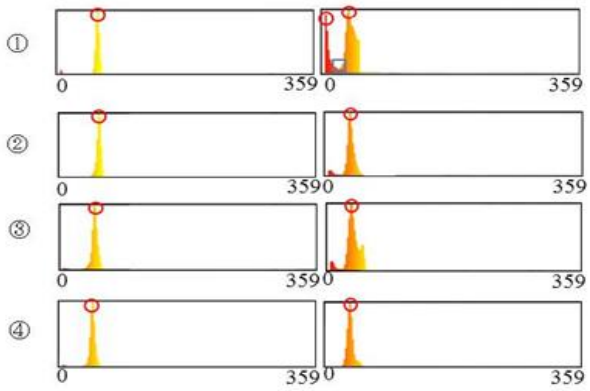


Figure. 6. An example of different objects with a similar hue value.

(a) Divided area of Fig. 6 (a). (b) Divided area of Fig. 6 (b).



(c) Local hue histograms of each divided area of Fig.7 (a). (d) Local hue histograms of each divided area of Fig.7 (b).

Figure. 7. An example of local hue histograms in divided areas.

whole object area. As our approaches, firstly, the proposed method divides an object area into plural areas. As an example, Fig. 7 (a), (b) shows divided objects. Secondly, the proposed method extracts hue from each divided area and generates local hue histograms. As an example, Fig. 7 (c), (d) shows the local hue histograms of Fig. 7 (a), (b). Thirdly, the proposed method extracts the positions of peak and trough from each local hue histogram as invariant feature. As shown in Fig. 7 (c), (d), the positions of peak and trough of the local hue histogram changes for each divided area by dividing an object area. Finally, the proposed method compares invariant feature of the target object with invariant feature of the candidate object every each local hue histograms. Thereby, the proposed method can distinguish different objects which have the same feature of hue histogram. Furthermore, the proposed method has robustness against the rotation change, the scale change, the illumination change, the distortion by perspective projection, occlusion and an object which has few textures as shown in Table 1.

2 PROPOSED METHOD

2.1 Flow of Proposed Method

We show the flow of the proposed method in fig.8. We represent each processing in the next sections.

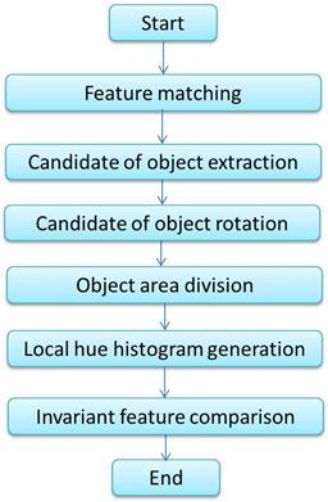


Figure. 8. Flow of the proposed method.

2.2 Feature Matching

Firstly, the proposed method extracts feature points from the registered image (Fig. 9 (a)) and the input image (Fig. 9 (b)) based on the feature of the BRISK [7]. As an example, we set the registered image which rotated 90 degrees to the input image in this time. Finally, the proposed method performs matching based on the BRISK descriptor between the registered image and the input image.

2.3 Candidate of Object Extraction

Firstly, as shown in Fig. 10, the proposed method calculates a vector **A** from the most matched feature point of BRISK to the most far away endpoint on the object area in the registered image, and defines a norm of vector **A** as the maximum distance dm. Secondly, the proposed method calculates vector **R** from the most matched feature point of BRISK to the second most matched feature point of BRISK in the registered image and vector **I** from the most matched feature point of BRISK to the second most matched feature point of BRISK in the input image, and calculates the ratio of the size of the registered image to the size of the input image by using

(a) Registered image. (b) Input image.

Figure. 9. Sample data for overview.

(a) Registered image. (b) Input image.

Figure. 10. Candidate area extraction.

$$r = \|\mathbf{R}\| \,/\, \|\mathbf{I}\| \tag{1}$$

Where r is the ratio, $\|\mathbf{R}\|$ is the norm of vector **R**, $\|\mathbf{I}\|$ is the norm of vector **I**. Thirdly, the proposed method describes a circle with the center is the most matched feature point of BRISK and radius is $\|\mathbf{A}\|$ in the registered image. In addition, the proposed method describes a circle with the center is the most matched feature point of BRISK and radius is $\|\mathbf{B}\|$ in the input image. Vector **B** is calculated by using

$$\mathbf{B} = \mathbf{A} \,/\, r \tag{2}$$

Where **B** is vector **B**, **A** is vector **A**, r is the calculated ratio. Fourthly, the proposed method extracts the contour of the circle in the registered image and the input image. Finally, the proposed method trims the registered image and the input image with in the extracted contours. Thereby, it is possible to accurately extract the candidate area even when scale is different or occlusion occurs as shown in Fig. 10.

2.4 Candidate Object Rotation

As shown in Fig. 11, in case that the input image rotates, the feature amount of the hue histogram in the same region changes. Therefore, it is necessary to rotate the input image. Firstly, the proposed method extracts feature points from the trimmed registered image and the trimmed input image, and performs matching based on the BRISK descriptor between the trimmed registered image and the trimmed input image. Secondly, since the proposed method rotates the trimmed input image, it is necessary to calculate a homography matrix describing the relationship between corresponding points as shown in equation (3).

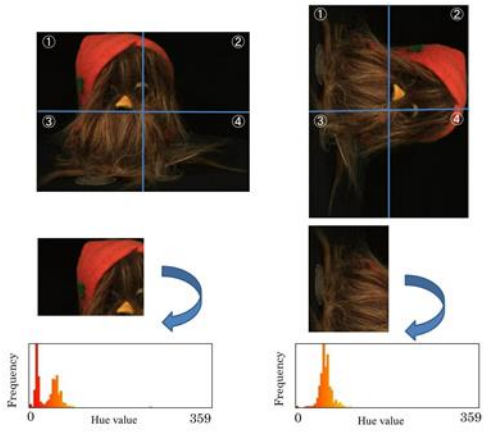


Figure. 11. An example of a case that the target object in the input image is rotating.

$$\begin{pmatrix} sx_2 \\ sy_2 \\ s \end{pmatrix} = H \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ 1 \end{pmatrix}, \tag{3}$$

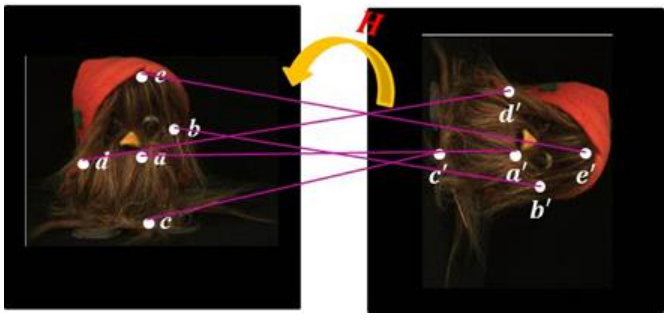
Where H is the homography matrix, (x_1, y_1) are the coordinates before the rotation, (x_2, y_2) are the coordinates after the rotation, s is a coefficient. Therefore, the proposed method calculates a homography matrix by using Direct Linear Transform algorithm. Finally, the proposed method rotates the trimmed input image by calculated the homography matrix as shown in Fig. 12.

2.5 Object Area Division

To generate the local hue histogram, the proposed method divides the trimmed registered image and the rotated input image into plural areas. Here, we define the number of division as four as an example.

2.6 Local Hue Histogram Generation

Firstly, the proposed method extracts hue from each divided area and generates local hue histograms. The hue value of the generated histogram is represented from 0 to 359. Fig. 13 shows the divided areas and local hue histograms. We define the vertical axis as the frequency, and the horizontal axis as the hue value. Here, we pay attention to item ① of Fig. 13. Fig. 14 (a) shows an



(a) Registered image. (b) Input image.

Figure. 12. Rotation processing.

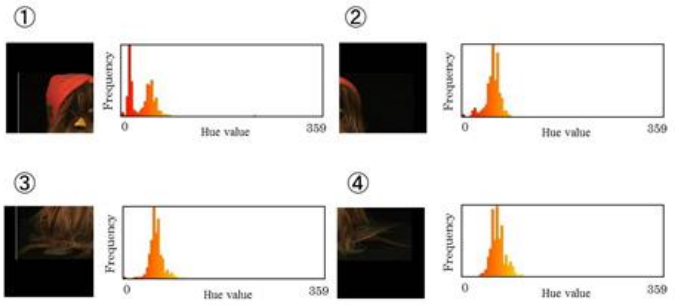
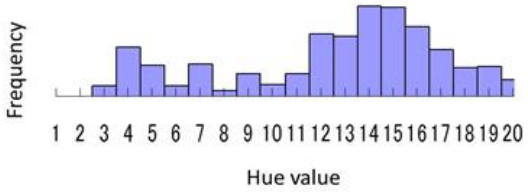


Figure. 13. An example of divided areas and local hue histograms.

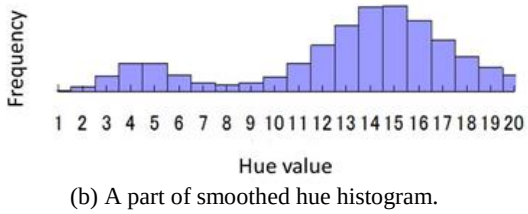
expanded hue histogram of item ① of Fig. 13. In Fig. 14 (a), because there are small irregularities at 7 and 9 of hue value, the feature amount becomes unstable by detecting the positions of peak and trough of the hue histogram in this state. Secondly, to eliminate those small irregularities, the proposed method smooths the hue histogram by using Gaussian filter. Fig. 14 (b) shows a smoothed hue histogram of Fig. 14 (a). In Fig. 14 (b), we can see that small irregularities of the hue histogram are omitted and the characteristic positions of peak and trough of the hue histogram are remained. Finally, the proposed method extracts the positions of peak of the smoothed hue histogram from each divided areas of the registered image and the input image by using

$$(H_{x-1} < H_x) \wedge (H_x > H_{x+1}), \quad (4)$$

and the positions of trough of the smoothed hue histogram from each divided areas of the registered image and the input image by using



(a) A part of hue histogram.



(b) A part of smoothed hue histogram.

Figure. 14. Smoothing processing.

$$(H_{x-1} > H_x) \wedge (H_x < H_{x+1}), \quad (5)$$

and registers them as invariant feature. Where H_x is the hue value which is focused on, H_{x-1} is the hue value before one of H_x. H_{x+1} is the hue value after one of H_x. And then, the extracted positions of peak and trough of the registered image are expressed by using

$$\left\{{p_1}_{(a)}{}^{(\alpha)}, {p_2}_{(a)}{}^{(\alpha)}, \cdots\right\} \in {P_{(a)}}^{(\alpha)}, \quad (6)$$

$$\left\{{t_1}_{(a)}{}^{(\alpha)}, {t_2}_{(a)}{}^{(\alpha)}, \cdots\right\} \in {T_{(a)}}^{(\alpha)}, \quad (7)$$

Where ${P_{(a)}}^{(\alpha)}$ is a set of peak position of a smoothed hue histogram in a divided area α of a. ${p_1}_{(a)}{}^{(\alpha)}, {p_2}_{(a)}{}^{(\alpha)}, \cdots$ which are each peak positions of a smoothed hue histogram, ${T_{(a)}}^{(\alpha)}$ is a set of trough position of a smoothed hue histogram in a divided area α of a. ${t_1}_{(a)}{}^{(\alpha)}, {t_2}_{(a)}{}^{(\alpha)}, \cdots$ which are each trough positions of a smoothed hue histogram. In addition, the extracted positions of peak and trough of the input image are expressed by using

$$\left\{hp_1{}^{(\beta)}, hp_2{}^{(\beta)}, \cdots\right\} \in hP^{(\beta)}, \quad (8)$$

$$\left\{ht_1{}^{(\beta)}, ht_2{}^{(\beta)}, \cdots\right\} \in hT^{(\beta)}, \quad (9)$$

Where $hP^{(\beta)}$ is a set of peak position of a smoothed hue histogram in a divided area β, $hp_1{}^{(\beta)}, hp_2{}^{(\beta)}, \cdots$ which are each peak positions of a smoothed hue histogram, $hT^{(\beta)}$ is a set of trough position of a smoothed hue histogram in a divided area β, $ht_1{}^{(\beta)}, ht_2{}^{(\beta)}, \cdots$ which are each trough positions of a smoothed hue histogram.

2.7 Invariant feature comparison

Firstly, the proposed method calculates difference values between invariant feature of the registered image and invariant feature of the input image by using

$$Dp_1 = \min_{\substack{1 \le y \le n \\ 1 \le \alpha, \beta \le i \\ 1 \le a \le m}} \left| {p_1}_{(a)}{}^{(\alpha)} - hp_y{}^{(\beta)} \right|, \quad (10)$$

$$Dp_2 = \min_{\substack{1 \le y \le n \\ 1 \le \alpha, \beta \le i \\ 1 \le a \le m}} \left| {p_2}_{(a)}{}^{(\alpha)} - hp_y{}^{(\beta)} \right|, \quad (11)$$

$$\vdots$$

$$Dt_1 = \min_{\substack{1\le z\le l \\ 1\le \alpha,\beta\le i \\ 1\le a\le m}} \left| {t_1}_{(a)}{}^{(\alpha)} - ht_z{}^{(\beta)} \right|, \quad (12)$$

$$Dt_2 = \min_{\substack{1\le z\le l \\ 1\le \alpha,\beta\le i \\ 1\le a\le m}} \left| {t_2}_{(a)}{}^{(\alpha)} - ht_z{}^{(\beta)} \right|, \quad (13)$$

$$\vdots$$

$$DP = \sum_{\alpha=1}^{f} Dp_\alpha, \quad (14)$$

$$DT = \sum_{b=1}^{k} Dt_b, \quad (15)$$

$$D = DP + DT, \quad (16)$$

Where n is the number of peak of a smoothed hue histogram in a divided area β , i is the number of division, m is the number of registered image, l is the number of trough of a smoothed hue histogram in a divided area β . $Dp_1, Dp_2, \cdots$ which are the smallest difference values between $hP^{(\beta)}$ to ${p_1}_{(a)}{}^{(\alpha)}, {p_2}_{(a)}{}^{(\alpha)}, \cdots, {p_1}_{(a)}{}^{(\alpha)}, {p_2}_{(a)}{}^{(\alpha)}, \cdots$ which are each peak positions of a smoothed hue histogram in a divided area α of α, $Dt_1, Dt_2, \cdots$ are the smallest difference values between $hT^{(\beta)}$ to ${t_1}_{(a)}{}^{(\alpha)}, {t_2}_{(a)}{}^{(\alpha)}, \cdots, {t_1}_{(a)}{}^{(\alpha)}, {t_2}_{(a)}{}^{(\alpha)}, \cdots$ which are each trough positions of a smoothed hue histogram in a divided area α of α, f is the number of peak of a smoothed hue histogram in a divided area α of α, k is number of trough of a smoothed hue histogram in a divided area α of α, DP is the total value of difference value of peak position, DT is the total value of difference value of trough position, D is the total difference value. As an example, Fig. 15 shows that the comparison of the hue histogram of a registered image and the hue histogram of an input image. As shown in Fig. 15, the proposed method compares the positions of peak and trough of a smoothed hue histogram of a registered image with the positions of peak and trough of a smoothed hue histogram of an input image, and calculates difference values. In addition, the proposed method registers a peak and a trough having the smallest difference value as the nearest peak and trough. Finally, the proposed method detects the object which has smallest D.

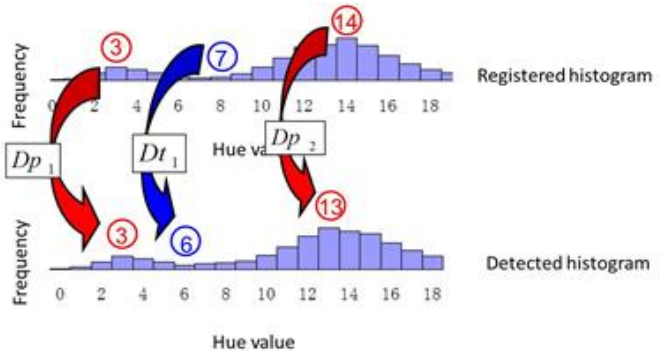


Figure. 15. An example of matching based on the positions of the unevenness of the hue histogram.

3 EXPERIMENT

3.1 Experiment Relating to the Number of Division of an Image

1) Experiment Overview: In this experiment, to determine the optimum number of division, changing the number of division from 1 piece to 25 pieces, the proposed method detected a registered object from an input image. Furthermore, we calculated the correct answer ratio by using

$$A = \frac{c}{z} \times 100\ [\%]\ , \quad (15)$$

and defined the number having the highest correct answer ratio as the number of division. Where A is the correct answer ratio, c is the number which the proposed method could correctly detect objects, z is the number of the registered objects. We selected 5 images in normal object (Fig. 16 (a)), 5 images in object which has few textures (Fig. 16 (b)), and 5 images in object which has same hue value (Fig. 16 (c)), totaling 15 images of objects from Amsterdam Library of Object Images [8] as the test image. In the object which has same hue value, we selected the object which has hue values of yellow and red.

2) Experimental Results: Fig. 17 shows the result of this experiment. As a result, a correct answer ratio increased as number of division increased. In addition, the correct answer ratio became the highest in case that the number of division is four and continued maintaining afterwards. However, we carried out the experiment on occluded objects, because we could not determine the most suitable number of division in this result. Fig. 18 shows the

(a) Normal object. (b) Object with few textures. (c) Object which has same hue value.

Figure. 16. Features of each object.

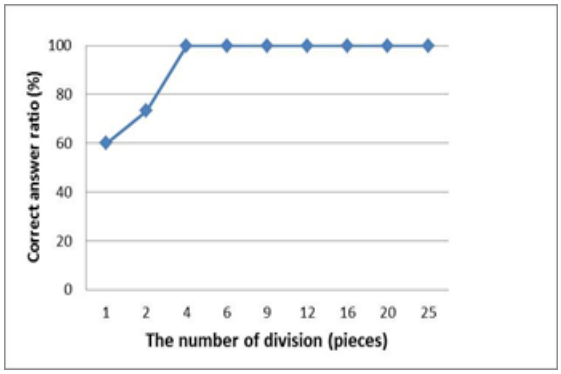


Figure. 17. The correct answer ratio of the experiment for the number of division.

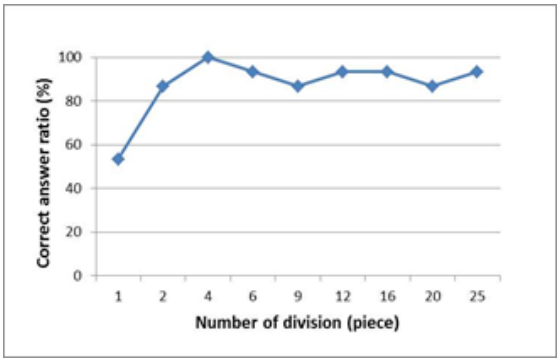


Figure. 18. The correct answer ratio of the experiment for the number of division in occluded objects.

result of the experiment of the occluded objects. As a result, the correct answer ratio became highest in case that the number of division is four. However, the correct answer ratio repeated increase and decrease between 80% and 100%. From these results, we defined the number of division as four pieces in this paper.

3) Discussions: We understood that the optimal number of division was difference according to objects from these experimental results. Therefore, we saw that it is necessary to statistically calculate the optimal parameter to increase the correct answer ratio. We understood that the optimal number of division was difference according to objects from these experimental results. Therefore, we saw that it is necessary to statistically calculate the optimal parameter to increase the correct answer ratio.

3.2 Experiment of Robustness against All Difference of Visual Appearance

1) Experiment Overview: In this experiment, to show the robustness against seven properties mentioned above of the proposed method, we carried out the experiment while giving objects changes, and compared the correct answer ratio of the proposed method with the correct answer ratio of conventional methods. We used objects same as experiment of the number of division as the inspection object. In addition, we calculated the correct answer ratio by using equation (17). Here, we show setting of each method. In the proposed method, we defined the number of division as four pieces and σ of Gaussian filter as 1.0. In the Tanaka's method, we defined σ of Gaussian filter as 1.0. In SIFT, we defined the threshold of correspondence points as 70.0. In addition, we defined the object which has correspondence points more than ten points and most correspondence points as a detection result. In the Color Indexing, we defined 255 gradation which was divided into four as each axis of the histogram.

2) Experiment Results: In the experiment of scale change and rotation change, we only indicate the result of the proposed method without comparing the proposed method with the conventional methods, because both the proposed method and the conventional methods have robustness against the rotation change and the scale change.

a) The scale change: We carried out the experiment while increasing the scale level from 0.8 times to 1.2 times every 0.2. As a result of this experiment, the correct answer ratio was 100% in each scale level.

b) The rotation change: We carried out experiment while increasing the rotation degree from 0 degrees to 180 degrees every 90 degrees. As a result of this experiment, the correct answer ratio was 100% in each rotation degree.

We indicate the result of the experiment about robustness against the illumination change, the distortion by perspective projection and occlusion. In other expriments, we indicate the result of the proposed method and the conventional methods.

c) The illumination change: We carried out the

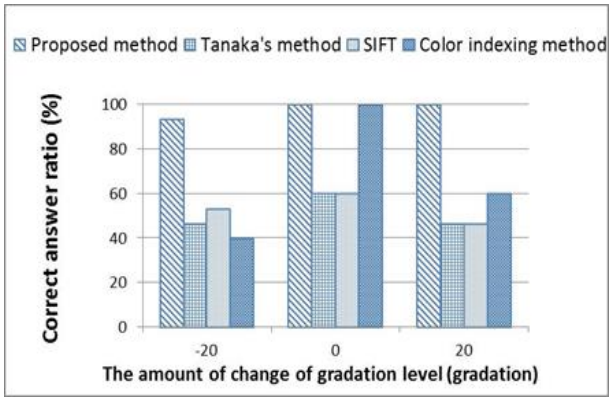


Figure. 19. The correct answer ratio of the experiment about the illumination change.

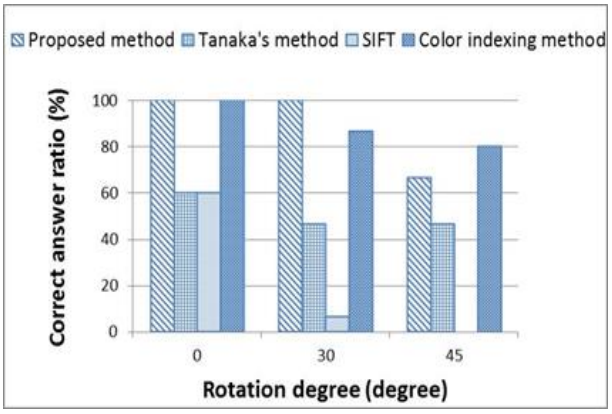
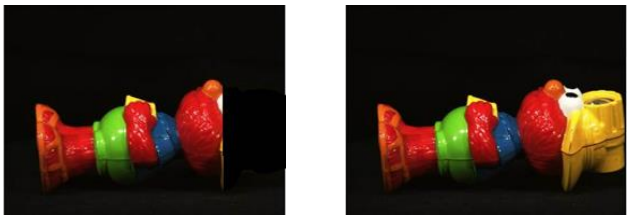


Figure. 20. The correct answer ratio of the experiment about the distortion by perspective projection.

experiment while increasing the gradation levels from -20 gradations to +20 gradations every 20 gradations. Fig. 19 shows the result of the experiment about the illumination change. As shown in Fig. 19, the correct answer ratio of the proposed method was the highest ratio, whereas the correct answer ratio of the Color Indexing was the lowest ratio as we expected.

d) The distortion by perspective projection: We carried out the experiment for the rotation in 30 degrees and 45 degrees. Fig. 20 shows the result of the experiment about the distortion by perspective projection. As shown in Fig. 20, the correct answer ratio of the proposed method was the highest ratio, whereas the correct answer ratio of the SIFT was the lowest ratio as we expected.

e) The occlusion: We carried out the experiment while increasing occlusion ratio from 10 percent to 40 percent every 10 percent. As an example, we show an occluded object in Fig. 21 (a). Fig. 22 shows the result of the experiment about occlusion. As shown in Fig. 22, the correct answer ratio of the proposed method was higher than the correct answer ratio of conventional methods as we



(a) Occluded object. (b) Original image of the occluded object.

Figure. 21. An example of the occluded object.

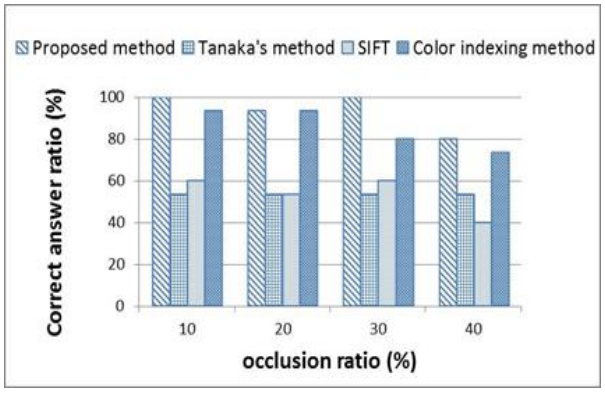


Figure. 22. The correct answer ratio of the experiment about occlusion.

expected.

In addition, we focus on the feature of the each object. Fig. 23 shows the correct answer ratio of the each object. As shown in Fig. 23, the proposed method and the Color Indexing obtained a high correct answer ratio in all features, whereas the SIFT obtained a low correct answer ratio in the object which has few texture. In addition, the Tanaka's method got a low correct answer ratio in the object having the same hue value and the normal object.

3) Discussions:

a) Proposed method: The correct answer ratio was higher or nearly equal than the conventional methods. However, the correct answer ratio decreased a little in case of 45 degrees of the distortion by perspective projection. The reason for this, since the local feature amount of the image was changed by perspective projection, the matching by the BRISK feature could not be performed well, and errors occurred on the posture estimation for the registered image and the input

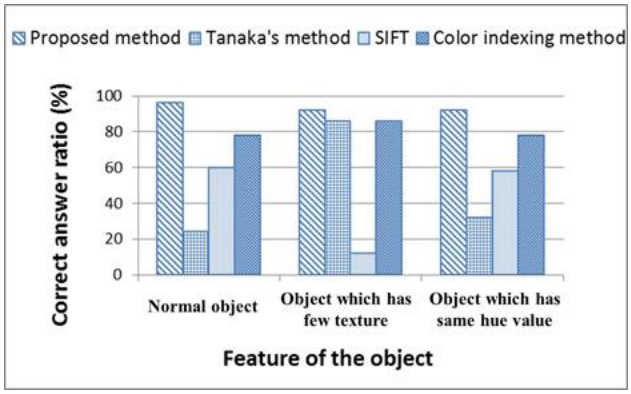


Figure. 23. The correct answer ratio for each feature of the object.

image. Thereby, it is thought that errors occurred on the position of the unevenness of the hue histogram within the divided areas, and the recognition ratio decreased.The correct answer ratio was higher or nearly equal than the conventional methods. However, the correct answer ratio decreased a little in case of 45 degrees of the distortion by perspective projection. The reason for this, since the local feature amount of the image was changed by perspective projection, the matching by the BRISK feature could not be performed well, and errors occurred on the posture estimation for the registered image and the input image. Thereby, it is thought that errors occurred on the position of the unevenness of the hue histogram within the divided areas, and the recognition ratio decreased.

b) Tanaka's method: The correct answer ratio was less than 60% in all changes. However, we see that there are few differences by comparing the correct answer ratio in the object which a change is given and the correct answer ratio in the object which a change is not given. Therefore, there seems to be other reasons why the correct answer ratio of Tanaka's method decreased. So, we focus on the feature of the each object. As shown in Fig. 23, the correct answer ratio was low in the object having the same hue value and the normal object. The reason for this, since Tanaka's method uses the position of the unevenness of the hue histogram in whole object area as feature descriptor, the Tanaka's method can not distinguish different objects having the same hue value. Furthermore, when objects have plural pieces of color information, they have the same hue value with high probability, therefore it is thought that erroneous recognition occurred in a normal object as shown in Fig. 16 (a) and the correct answer ratio decreased.

c) SIFT: The correct answer ratio was less than 60% in all changes. Especially, the correct answer ratio decreased on the distortion by perspective projection. The reason for this, the local feature amount changes by the distortion by perspective projection, therefore it is thought that the SIFT descriptor changes and the correct answer ratio decreased. In addition, we focus on the feature of the each object. As shown in Fig. 23, the correct answer ratio was low in the object which has few textures. The reason for this, the SIFT descriptor uses the gradient information of the object, therefore it is thought that the SIFT could not detect the feature amount on the object which has few edges such as the object which has few textures, and the correct answer ratio decreased.

d) Color Indexing: The correct answer ratio decreased on the illumination change. The reason for this, the RGB color system which is used for the Color Indexing is easy to be affected by the illumination change, therefore it is thought that the value of the three-dimensional color histogram changed and the correct answer ratio decreased.

4 CONCLUSION

In this study, we proposed the object detection method using invariant feature based on the local hue histogram in divided areas of an object for the human support robot. To show that the proposed method satisfies all seven properties as follows,

1. Robustness against the rotation change
2. Robustness against the scale change
3. Robustness against the illumination change
4. Robustness against the distortion by perspective projection
5. Robustness against the occlusion
6. Detecting an object which has few textures
7. Detecting different objects which have the same features of the hue histogram

We carried out experiments. As a result, we could show that the proposed method satisfies all seven properties. However, the detection accuracy is limited, because the proposed method uses only two-dimensional information. Therefore, in the future, we aim to improve the ability for detection by using three-dimensional information such as the shape information.

REFERENCES

[1] S. Sugano, T. Sugaiwa, and H. Iwata, "Vision System for Life Support Human-Symbiotic-Robot," The Robotics Society of Japan, vol. 27, pp. 596–599, 2009.

[2] T. Odashima, M. Onishi, K. Thara, T. Mukai, S. Hirano, Z. W. Luo, and S. Hosoe, "Development and evaluation of a human-interactive robot platform "RI-MAN"," The Robotics Society of Japan, vol. 25, pp.554–565, 2007.

[3] Y. jia, H. wang, P. Sturmer, and N. Xi, "Human/robot interaction for human support system by using a mobile manipulator," ROBIO, pp. 190–195, 2010.

[4] D. G. Lower, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision,60, pp.91-110, 2007.

[5] M. j. Swain, and D. H. Ballard, "Color Indexing," IJCV, vol. 7, pp.11–32, 1991.

[6] K. Tanaka, Y. Hagiwara, and H. Imamura, "Object Dtection in Image Using Feature of Invariant based on Histogram of Hue," IEICE, pp.187–194, 2011.

[7] Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In Computer Vision (ICCV), 2011 IEEE International Conference on (pp. 2548-2555). IEEE.

[8] http://aloi.science.uva.nl/