INTRODUCTION

The elderly population of Japan exceeded 25% of its total population in 2013, and thus, Japan has become a super-aged society [1]. There is therefore a need to develop monitoring support systems that can help observe the state of the elderly even from a remote location. Nowadays, video chat is widely used as a support tool to monitor elderly people remotely. Systems using cameras such as Ramrock eye’s [2] and Mimamori Cube [3] are provided by Ramrock Co., Ltd. A system using sensors such as the skeleton tracking sensor / fall detection system [4] provided by the corporation Network21 is also provided. Because images captured by the camera during video chat contains information about the interiors of the room, the potential leakage of private or confidential information is of great concern. Therefore, in order to improve the security and quality of life, it is necessary to protect privacy in using video chat to monitor elderly people. To that end, it will suffice to extract only a person from the camera image in real time and replace background information with a predetermined image. In addition, background replacement can improve the extraction accuracy of a person by using a monochromatic background such as a chrome frame processing. When assuming use in daily life, it is necessary to extract people from the various objects located in the living space, without using background papers. In this paper, we propose a background replacement method for privacy protection in the elderly care system with regard to person extraction during video chat.

DATA ACQUISITION ENVIRONMENT AND IMAGE DATA USED

Figure 1 and 2 show the data acquisition environment and the image data used. There were 15 subjects (men and women in their twenties) and the data was acquired using a web camera (Logicool HD Pro Webcam C920). As an indoor environment was used, photographs were taken with five different backgrounds under ordinary fluorescent lights (illuminance 400 to 1000 lx): background A (a bookshelf, piece of furniture, and refrigerator), background B (a piece of furniture and small articles like books), background C (a chair and blinds), background D (clothing and furniture) and background E (bed). We also acquired data from an environment involving the daily living conditions of the elderly. Furthermore, we used a landscape image [7] as the background replacement image.

The data used in the study was acquired in accordance with the ethical regulations regarding human subjects at Akita University, Japan.

BACKGROUND REPLACEMENT METHOD

The conventional method of performing background replacement is by using a mask obtained through background subtraction [5]. However, considering the use of the system in daily life, it is necessary to extract person information from images that include various objects present in their living space without using background papers [6].

In this paper, we therefore propose a background replacement method for person information extraction and privacy protection for the elderly through video chat.

Keywords: Elderly care, support for monitoring elderly people, privacy protection
Person Region Extraction and Background Replacement in Images for Privacy Protection

In this study, (i) the outermost contour extraction process, (ii) human region extraction, and (iii) shadow removal process were performed after background subtraction processing. The outline of the proposed method is shown in Figure 3.

3.1 Outermost contour extraction process

Even after subjecting the image to background subtraction and dilation erosion processing [5], black pixels were found in the human region. Therefore, the outermost contour extraction process was applied to remove black pixels in the human region. Specifically, by using the contour extraction process [8], the outermost contour of concatenated white pixels was acquired. By replacing the inside of the acquired contour with white pixels, the loss of the human region was reduced, and a new background difference mask was obtained. Figure 4 shows the processing result.

3.2 Human region extraction

The outermost contour extraction process cannot eliminate the dotted white pixels in the background area. Therefore, the human region extraction was performed to remove the white pixels from the background area. Specifically, the front face, profile, and upper body were detected from the portrait image using Haar-like features [8] with respect to the background image, and the coordinates of their centers were acquired. Those features, whose center coordinates were within the outermost contour, were extracted as human region labels. By filling the inside of the extracted labels with white pixels, the noise in the background area was removed, and a mask of the human region was obtained. The processing result is shown in Figure 5.

3.3 Shadow removal process

In the case of the background environment involving elderly people, the shadow of the person cast on the wall or bed caused the human region to be excessively extracted. Therefore, shadow removal processing was applied to remove shadows. Specifically, variance, one of the texture features [5] defined in, was calculated using the portrait image and the background image.
image. Then, the difference between the texture features amounts calculated for both images was acquired, and dilation erosion processing [5] was sequentially applied. Further, contour extraction processing [8] was performed. Finally, we compared the results of contour extraction with the human region mask. Specifically, a white pixel in the human region mask with a corresponding black pixel in the contour extraction processing result was output as a black pixel in the final result. The processing result is shown in Figure 6.

4 EXPERIMENTAL RESULTS AND DISCUSSION

To verify the usefulness of the proposed method, three types of experiments were conducted as follows:

i. In daily life, because a person does not always face the front, background replacement was carried out using the front face and profile of the subject.

ii. Background replacement was carried out under the condition that two subjects are present in the same image.

iii. Background replacement was carried out in an environment considering the living conditions and care of the elderly.

Table 1 shows the number of subjects and image data used in each experiment.

| Experiment | Number of subjects | Number of image data |
|------------|--------------------|----------------------|
| i          | 15                 | 86                   |
| ii         | 4                  | 6 (Two subjects in one image) |
| iii        | 14                 | 84                   |

4.1 Subjective evaluation

The results of experiments i and ii are shown in Figure 7 and 8. Table 2 lists the number of people detected and the number of successful background replacements by the proposed method in experiment i. The background environment includes miscellaneous objects such as electric appliances, chairs, and blinds. The aim of the experiment was to replace the background in case of environments that we come across daily. Compared to the conventional method, the proposed method shows minor loss of the human region and noise in the background area.

In experiment i, we were successful in detecting the face in 84 of the 86 images acquired and in creating 72 shadow masks. The results indicate that face and upper body detection enables the removal of noise areas even when the person is not facing the front, and this helps in creating a good shadow removal mask, as shown in Figure 7.

In experiment ii, the results indicate that the mask for two persons was created successfully in all the six images acquired, as shown in Figure 8. This suggests that the proposed method can replace a background containing miscellaneous objects even when multiple persons are present.

Figure 9 and 10 show the results of experiment iii. Table 3 lists the number of people detected and the number of successful shadow removals by using the proposed method in experiment iii. The aim of the experiment was to replace the background for an environment with a bed and pillows, where the elderly are cared for. Unlike the conventional method, the proposed method can remove shadows cast on the wall.

Table 2 shows the number of subjects and image data used in each experiment.

In experiment iii, we successfully detected the face in all the 84 images acquired. By applying the shadow removal process, we succeeded in removing the shadows cast on the wall in 83 of the 84 images acquired. The results suggest that person detection is possible even if the face and upper body are tilted. Furthermore, by using the shadow removal process, it is possible to create a good shadow removal mask to remove shadows cast on the wall. Even if the position of the bedding and its color changes, the proposed method is able to adapt by updating the background.

The outermost contour extraction process and human region extraction process can extract the human region, and the shadow removal process is able to reduce the influence of shadows in comparison with the conventional method.

4.2 Quantitative evaluation

In order to calculate the accuracy of background replacement, a human region is extracted by visual inspection, a correct mask (Figure 11) is created, and the matching rates of the foreground (human region) background regions and the overall matching rate are calculated. Tables 4 to 6 show the average concordance rates calculated for each experiment and subject.

In experiment i, the $\bar{C}$ of the proposed method is 98.2%, which is 8.4% higher than that of the conventional method, indicating that noise in the background region has been removed. The overall coincidence rate $\bar{C}$ was obtained as the proposed method was able to replace the background with an accuracy of 98.9%, which is 2.9% higher than the conventional method.

In Experiment ii, as well as in Experiment i, the proposed method showed higher values than the conventional method at all concordance ratios.

In Experiment iii, the $\bar{C}$ of the proposed method is 92.8% which is 10.5% higher than that of the conventional method, indicating that the noise in the human region is removed. The overall coincidence rate $\bar{C}$ was obtained as the proposed method was able to replace the background with an accuracy of 93.9%, which is 8.2% higher than that of the conventional method.

5 CONCLUSIONS

In this paper, for the purpose of privacy protection, we propose a background replacement method considering the daily living
conditions of the elderly. The conclusions are as follows:
1. The proposed method can achieve effective background replacement for both the front face and profile.
2. Person detection is possible even when the face and upper body are tilted.
3. The shadow removal process enables effective removal of shadows cast on the wall.
4. It is revealed that the proposed method is useful for background replacement when assuming the living environment and care for elderly people.
Table 4  Average concordance rate in Experiment i

| Method          | $\bar{C}_F$ (%) | $\bar{C}_B$ (%) | $\bar{C}$ (%) |
|-----------------|-----------------|-----------------|--------------|
| Proposed method | 99.2            | 98.2            | 98.9         |
| Conventional    |                 |                 | 96.0         |

Table 5  Average concordance rate in Experiment ii

| Method          | $\bar{C}_F$ (%) | $\bar{C}_B$ (%) | $\bar{C}$ (%) |
|-----------------|-----------------|-----------------|--------------|
| Proposed method | 99.4            | 97.6            | 98.9         |
| Conventional    |                 | 89.9            | 96.4         |

Table 6  Average concordance rate in Experiment iii

| Method          | $\bar{C}_F$ (%) | $\bar{C}_B$ (%) | $\bar{C}$ (%) |
|-----------------|-----------------|-----------------|--------------|
| Proposed method | 92.8            | 97.7            | 93.9         |
| Conventional    | 82.3            | 97.4            | 85.7         |

References

[1] Statistics Bureau, Ministry of Internal Affairs and Communications, “Population of the elderly,” http://www.stat.go.jp/data/topics/topi971.htm (accessed 2017.05.02)

[2] “Dementia senile wandering sensing equipment Ramrock eye’s” http://www.ramrock.co.jp/ramrocksys/ram

[3] “Mimamori Cube” https://ramrock-eyes.jp/ Ramrock Co., Ltd. (accessed 2018.03.31)

[4] “Watch over / fall detection system for lonely elderly (Network21 Co., Ltd.)” http://robotcare.jp/?page_id = 2320 (accessed 2018.03.31)

[5] Takagi M.; Shimoda H., “Handbook of Image Analysis” University of Tokyo Press (2004)

[6] Murakami T.; Kageyama Y.; Nishida M., “Uncomfortable Feeling Reduction of Background Replacement on the Basis of Features of Color Temperature” Journal of Japan Society for Fuzzy Theory and Intelligent Informatics 25, 4, 806-818 (2013) (in Japanese)

[7] Yun T., “Yun Free Stack Photos,” http://www.yunphoto.net/ (accessed 2017.05.02)

[8] Bradski G.; Kaehler A.; Matsuda K., “OpenCV Image processing and recognition using computer vision library” O’Reilly (2009)