The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Privacy-Preserving Resident Monitoring System with Ultra Low-Resolution Imaging and the Examination of Its Ease of Installation

Takumi Kimura [*1] Shogo Murakami [*1] Ikuko Egushi Yairi [*1]

[*1] Graduate School of Science and Technology, Sophia University

Monitoring systems using infrared array sensors allow monitoring of residents while protecting their privacy. However, since such a sensor is vulnerable to subtle movements, accuracy of posture classification is low, and limits the locations and methods available for installation. This study proposes a posture classification method with higher accuracy. Over 93% accuracy was achieved in posture classification by RGB conversion of infrared array sensor images and successfully decreased loss due to displacement by DCNN. Additionally, this research considers methods to create artificially simulated data for postural-behavioral study. To check the validity of this method, postures of 3 subjects were examined using a classifier with studied simulation data. Finally, simulation environments with different sensor altitudes and angles were created to examine the ease of installation for the proposed method. As a result, the experiments showed that accuracy was highest at approximately 90% when the sensor was located 50cm below the height of the target and when the tilt angle was within ±2°.

1. Introduction

Resident monitoring systems are useful in detecting abnormal conditions of residents. However, as every move is under inspection, privacy issues arise. As a solution, usage of infrared array sensors has been proposed to preserve privacy as well as to avoid physical burden on the target. [Okada 13][高木 16][楠亀 17]. Such sensors can be placed in various places as they solely rely on temperature data obtained from the infrared sensor to detect the target. Spatial information and light measurements received from the sensors are used to identify the posture and location of the target and to observe their changes.

It is known that sensor installation angle is a factor for decreased classification accuracy, but the analysis on its effects are yet insufficient. Acquisition of learning data for machine learning is key in improving classification accuracy. To solve the aforementioned tasks, this paper assesses the classification accuracy by the single 8x8 infrared array sensor, proposes the methods of artificially creating simulated data to study postures for machine learning, and analyzes the effects caused by the angle of the installed infrared array sensor.

2. A resident monitoring system using ultra low-resolution infrared array sensor imaging

2.1 Posture classification system

To examine posture classification accuracy, a data collection device including an infrared array sensor was developed. A diagram of the device is shown in Figure 2. This device composes of a Raspberry Pi3 Model B mounted with a Grid-EYE (AMG8833) sensor. The Grid-EYE will output an 8x8 pixel image data of the surface temperature for objects detected in the observation space. Temperatures between 0°C -80°C can be detected with a step increment of 0.25°C.

Deep Convolution Neural Network (DCNN), is an effective method for high accuracy image recognition. It is used in this study to examine the obtained infrared image data for use in the production of a posture classifier. Figure 3 shows the structure of the DCNN used.

2.2 Posture classification experiment of a subject using DCNN

A posture classification experiment was conducted on 3 subjects to check the operation of the posture classification system and to evaluate the performance of the DCNN. The first two experiments were conducted in a 9.5m^2 Japanese-style room at roughly 13°C room temperature. Here, the data of a male subject of age 24 and height of 170cm, and a female subject of age 20 and a height of 160cm was obtained. The experiment for the third subject was done in a 20m^2 room with a room temperature of roughly 11°C. The subject was male, of age 22, and was 170cm in height. For all three scenarios, the sensor was placed 140cm from ground level such that the entire body of each subject could be observed.

The subjects were stationed 1-3m away from the sensor and were told to stand, sit, or lie down within the view of the sensor. A total of 14983 frames worth of data were obtained. 10% of the above data were randomly selected and studied by the classifier, then were used to classify the remaining 90%. The result of posture classification, evaluation of accuracy, and recall ratio are shown in Table 1.

Considering practical use, a resident monitoring system needs to be able to detect instantaneous dangerous incidents such as slips and therefore it is desirable for the F-measure to be above 90%. Experimental results showed an F-measure of roughly 87%. It is probable that directly inputting 8x8 temperature data to the DCNN is insufficient for image feature extraction.

However, since DCNN is known to have high accuracy object detection for colored images, focus was placed on converting data

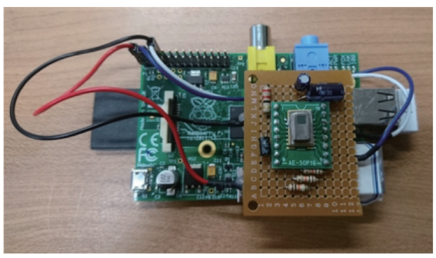

Fig. 1 A posture recognition device equipped with an 8×8 infrared array sensor.

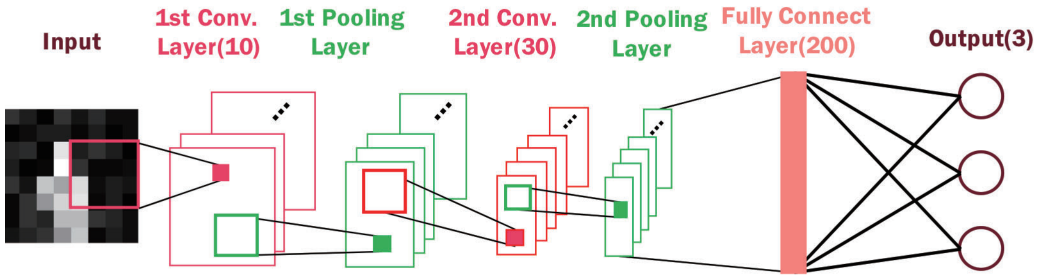

Fig. 2 Illustration of DCNN structure used in 2.2 experiment.

from the infrared array sensor into RGB and inputting them to the DCNN to improve the accuracy of the classifier. [Simonyan 15] This was done by mapping temperature data into a color space and using that as a reference to convert images to RGB. Data were then divided into R, G, and B, and were separately inputted to the DCNN. The classification results after RGB conversion, evaluation of accuracy, and recall ratio are shown in Table 2. In comparison to Table 1, the number of misclassifications decreased, and F-measure was above 90% for all postures. From such results, it could be concluded that RGB conversion of data led to less number of misclassifications. RGB converted data were used for DCNN input hereafter.

3. Learning data generator construction and evaluation of classification accuracy

3.1 Learning data generation procedures

Using the DCNN posture classifying method previously evaluated in 2.2, methods to generate simulation learning data without subject-based experiments were made. The infrared images handled in this research are for a room temperature distribution represented by 8x8 image pixels. Therefore, it can be theorized that placing a human model in the view of the sensor would require less effort while outputting results like those of the subject based experiments. For such a reason, a Unity program including physics engines and functions was used to simulate this environment and was used to produce learning data. This learning data generator was made to arbitrarily set the height and physique of the human model, the tilt and altitude of the sensor, and the size of the room. As for physique, the human model was composed of several body parts including the head, arms, and torso, with each part having its own temperature distribution that could also be arbitrarily changed. For this research, the radiant heat distribution obtained from a real environment experiment was used to set parameters for each body part. As for posture, the human model had 3 types of postures namely, “stand”, “sit” or "lie down”.

After running the learning data generator, ray tracing was performed from the sensor. Whenever the ray hit the human model, the radiant heat information of the human model on the incident spot was recorded. When the ray did not hit the model, the preset radiant heat or temperature of the background was recorded. Simulation by the learning data generator is depicted in Figure 3.

3.2 Learning method evaluation using simulated learning data

DCNN with studied simulation data were used to classify real environment data. The real data used were the same data as those used in 2.2. Simulation data collection was done by placing a human model randomly 1m-3m away from the sensor with three postures— either “stand”, “sit” or "lie down”. Room temperatures

Table 1 Posture classification results of 8x8 infrared array sensor images by DCNN.

Posture classification results

Recognition	Answer		
	Stand	Sit	Lie down
Stand	3582	266	15
Sit	496	5243	435
Lie down	185	590	4171

Posture classification evaluations

	Stand	Sit	Lie down
Accuracy	0.92726	0.84921	0.84331
Recall ratio	0.84025	0.85965	0.90262
F-measure	0.88161	0.85440	0.87196

Table 2 Action classification result of 8x8 RGB converted images of infrared array sensor.

Posture classification results

Recognition	Answer		
	Stand	Sit	Lie down
Stand	3929	137	20
Sit	217	5620	113
Lie down	121	336	4490

Posture classification evaluations

	Stand	Sit	Lie down
Accuracy	0.96158	0.94454	0.90762
Recall ratio	0.92079	0.92237	0.97123
F-measure	0.94074	0.93332	0.93835

and sizes were set according to real data. The human model also corresponded with subjects from the experiment, and the learning data generator was run for 30 minutes outputting 50000 sample data for each simulated subject. Finally, the posture classifier was used to analyze the obtained data. Table 3 shows the posture classification and evaluation results. Results show that average F-measure was relatively low— generally under 80%. From observing misclassified examples, it was hypothesized that such classification accuracy loss occurred due to the existence of high-temperature pixels in the background. Therefore, background elimination was conducted by analyzing the temperature difference in each data, estimating and eliminating background parts, and leaving only the human model in the data. The background elimination process is shown in Figure 4. This elimination method was applied to both simulated learning data and real data used for classification and evaluation. Then, classification was repeated for the second time. Table 4 shows the classification and evaluation results. Results showed that all classification average F-measures increased from under 80% to over 90% accuracy. From such results, it was safe to say that background elimination method was effective when using simulated learning data to classify real data.

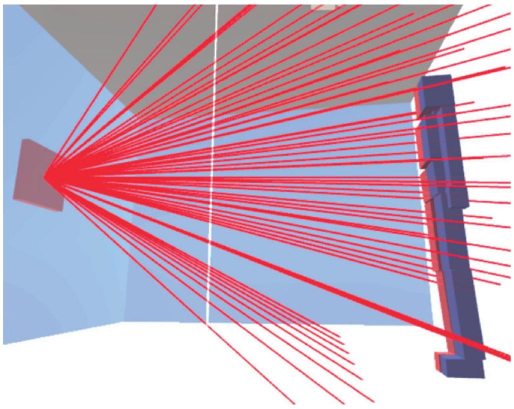

Fig. 3 The learning data generator simulating 8x8 infrared array sensing of a human in a room.

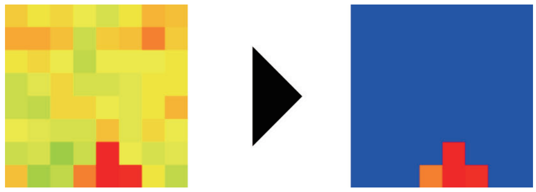

Fig. 4 Process of background elimination algorithm.

4. Evaluation on classification accuracy effects due to the system installation condition

So far, this research had fixed the sensor at a height of 140cm to successfully classify the postures with over 90% accuracy. However, it was never tested to confirm the range of heights this high accuracy rate can be sustained. In implementation, it is highly likely that the sensor will be slightly displaced or tilted from external factors. Therefore, the effect and degree of these factors against classification accuracy were tested, and system installation conditions were considered.

4.1 Evaluation on installation height

Using the learning data generator, the sensor height was changed from 50-170cm with a step increment of 10cm, and the human model height was set at either 170cm, 160cm, or 150cm. Then, for each condition, an evaluation of classification accuracy was performed. The results are shown in Figure 6. For all models, the accuracy peaked when the sensor height was 50cm below the model height.

4.2 Evaluation on sensor tilt

Next, the degree of effect on accuracy by the sensor angle were inspected. The human model height was set at 170cm, and sensor installation height was set at 120cm. Tilt range was set from -5° to 5° with increments of 1°. As for learning data, the preceding data without any tilt was used. Results are summarized in Figure 7. Results show that when the sensor is tilted upwards to 2°, the F-measure is over 90%, but when the tilt reaches 5°, the F-measure is decreased to roughly 86%. When the sensor is tilted down, F-measure is sustained at a high value until -1°. However, a tilt of over -2° drastically decreases the F-measure until under 80% at -

Table 3 Action classification results of the real data by learning with the simulated data.

Posture classification results

Recognition	Answer: Stand	Answer: Sit	Answer: Lie down
Stand	3306	537	216
Sit	959	5129	1274
Lie down	5	427	3130

Posture classification evaluations

	Stand	Sit	Lie down
Accuracy	0.81449	0.69669	0.87872
Recall ratio	0.77424	0.84179	0.67749
F-measure	0.79385	0.76239	0.76509

Table 4 Effects of background image removal on action classification of the real data by learning with the simulated data.

Posture classification results

Recognition	Answer: Stand	Answer: Sit	Answer: Lie down
Stand	4130	0	0
Sit	201	5665	585
Lie down	0	217	4185

Posture classification evaluations

	Stand	Sit	Lie down
Accuracy	1.00000	0.87816	0.95070
Recall ratio	0.95359	0.96311	0.87736
F-measure	0.97624	0.91867	0.91256

5°. Since this research takes privacy preservation as a serious consideration, the number of image pixels used are very low. Therefore, the posture classification relied heavily on high-temperature distributions per row, and a one-row difference gave an extensive impact on classification results.

The high-temperature distribution change caused by the sensor tilt was assumed to be the major cause of the decrease in classification accuracy.

5. Evaluation on classification accuracy effects due to the system installation conditions

In chapter 4, the influence on classification by installation height and sensor tilt were investigated to consider potential external effects upon real implementation. As a result, it was found that installation height yielded highest accuracy at height 50cm below the height of the target, and that sensor tilt largely impacted the classification accuracy.

With these results in mind, real implementation is further considered. Assuming the wall were to be perpendicular, the sensing device could be installed 50cm below the target height after the height of the target is measured. On the other hand, if a wall installation is difficult, then there may be a need to call a specialist to install the device.

A solution to this installation problem could be to create a personalized classifier for the target by inputting the target's room information data into the learning model. Although this method requires meticulous interview on the house conditions, by utilizing the learning data generator one can simulate and obtain data corresponding to the target room and apply the resident monitoring system

6. Conclusions

This paper justified that an infrared array sensor resident monitoring system using an infrared array sensor image with 8x8 pixels would output over 90% accuracy for posture action classification of the target. Noise analysis was performed on a tilt and it was concluded that approximately 90% accuracy was sustained for tilt angle within ±2° by extending the classifier. In addition, posture pattern learning data simulation was taken into consideration, and through comparison against real data, a high accuracy classifier construction was achieved. As for further research, an improvement in simulated learning data, learning data generator, and learning algorithm will be continued for application in a real environment.

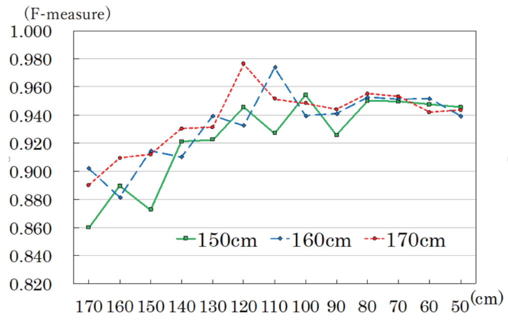

Fig. 5 Accuracy for various sensor altitudes.

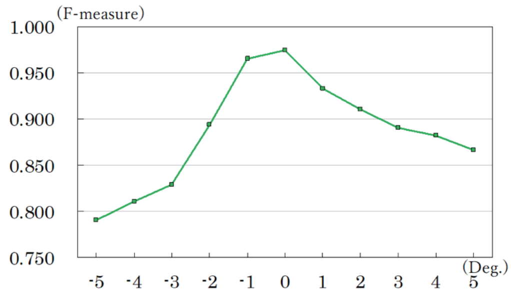

Fig. 6 Accuracy for vertical sensor tilt.

Acknowledgements

We would like to show our gratitude to Yuki Kato, Motoharu Sakurai, and all the supporters for their participated and assistance. This research was conducted under the support of 23rd and 24th year of Heisei period research grant from the Support Center for Advanced Telecommunications Technology Research, Grant-in-Aid for Scientific Research B(17H01946), and Grant-in-Aid for challenging Exploratory Research (16K12537).

References

[Okada 13] R. Okada and I. Yairi, "An indoor human behavior gathering system toward future support for visually impaired people," Proc. of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, no. 36, Washington, Oct. 2013.

[高木 16] 高木，高橋，大塚，"温度センサを用いた高齢者の見守り," 第8回データ工学と情報マネジメントに関するフォーラム (DEIM 2016), P6-1，福岡，2016.

[楠亀 17] 楠亀，米田，式井，シラワン，野坂，久保，"サーモカメラによる非接触温冷感センシング，"Panasonic Technical Journal, pp.118-122, vol.63, no.2, Nov. 2017.

[Simonyan 15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," International Conference on Learning Representations (ICLR), May 2015.