Human Interaction and Emerging Technologies (IHIET 2022), Vol. 68, 2022, 619–624
https://doi.org/10.54941/ahfe1002785

Personal Authentication Method Using Geometrical Features of External Auditory Canal

Yuki Muramoto, Yujie Li, Hideaki Kawano, and Yoshihisa Nakatoh

Kyushu Institute of Technology, 1-1, Sensui-cho, Tobata-ku, Kitakyusyu-shi, Fukuoka Prefecture, Japan

ABSTRACT

Gloves and masks made fingerprint and face recognition difficult in hospitals and factories. In this research, as a precondition for developing an authentication system based on geometrical features inside the ear canal using a camera, we verified whether it is possible to identify individuals from images inside the ear canal using machine learning. The method was to acquire image data of the ear canal using a camera and identify individuals using the image data and a model that has been learned in advance. In this study, because of the small amount of training data, we used a previously trained VGG16 model that was relearned by transfer learning. We used Accuracy, Recall, Precision, and F-measure as evaluation indices and obtained a high evaluation of 0.989 for all results.

Keywords: Image processing, Personal authentication, Biometric identification

INTRODUCTION

The number of workers in Japan's health and welfare industry gradually increases. In addition, more healthcare workers were expected to be needed due to the coronavirus. In line with this, there is a need to develop a medical environment in which many healthcare workers can work efficiently and smoothly. In recent years, there have been many moves in the medical field to reduce various burdens on medical personnel through IT, such as the storage of patient information in electronic media like electronic medical records and the digitization of medical questionnaires. This has led to the development of devices that use biometric authentication, such as fingerprint recognition, to enhance security to protect personal information such as electronic medical records.A biometric authentication system was a system that identifies an individual using physical characteristics such as human eyes and ears and behavioral characteristics. Because biometric authentication systems used the unique characteristics of each individual as a pass, they were complicated to impersonate or forge and were used in many fields as a robust security measure.The types of biometrics currently in use were as follows. Fingerprint authentication was an authentication method that identifies and compares fingerprints engraved on a finger. Facial recognition was an authentication method that compares facial features. Otoacoustic authentication was an

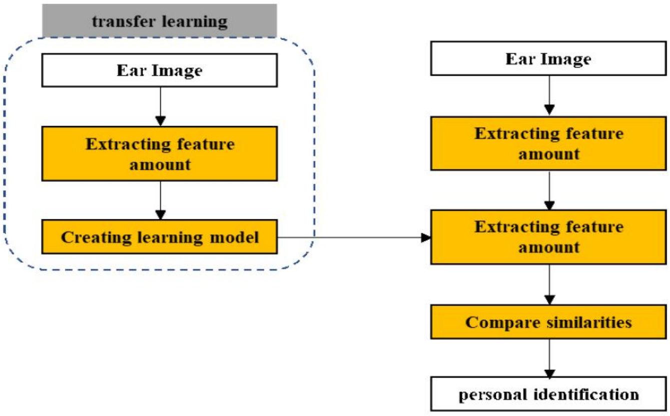

Figure 1: Proposed method.

authentication method that sends a sound into the ear and compares the features of the echo sound. In addition, "multimodal biometric authentication" was attracting attention as superior to conventional authentication systems. This multimodal biometric authentication was a revolutionary authentication system that combines two or more biometric methods and was expected to improve performance significantly. However, people wear masks and gloves when working in hospitals and factories, making fingerprint and face recognition challenging to match. Therefore, we investigated an authentication system that uses the ear as in otoacoustic authentication. The following was an overview of the system. In the current Japanese medical field, information technology was advancing, and biometric authentication was being considered to strengthen security to protect that information. However, there was a situation in which conventional biometric authentication was too cumbersome due to masks, hats, and gloves in hospitals. Therefore, developing a personal authentication system based on a different approach other than acoustic features used in the ear was expected to complement multiple authentication methods and provide high authentication accuracy. The main goal of our research was to develop a new ear-based authentication system. We proposed a biometric authentication system that used geometrical (image) features within the ear canal.

PROPOSED METHOD

The flow of the proposed method is shown in Figure 1. In this study, a camera was used to capture the inside of a human ear and acquire image data. Features were extracted from the acquired image data, and the features were compared with a training model that was previously trained using images of the human ear canal. Personal identification was performed based on the results. The final goal was to develop a device that can be easily combined with other authentication systems by using a camera built into the earphone, thereby improving authentication accuracy without increasing the burden on the user. Since the amount of training data we prepared on our own was small, we used a VGG16 model that has already been trained and re-trained by

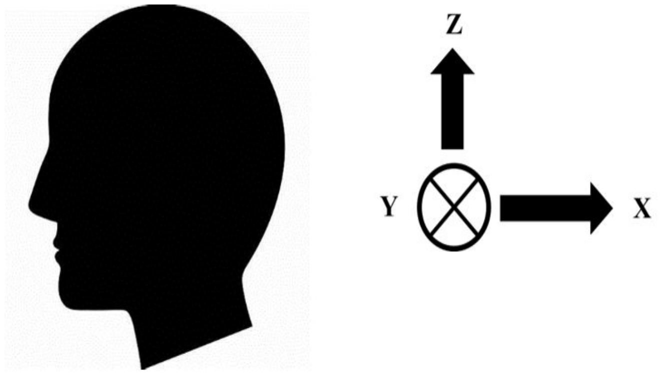

Figure 2: Coordinate system.

transfer learning (Simonyan et al., 2014), (Yusuke et al., 2013). The VGG16 model was a convolutional neural network consisting of 16 layers with parameters pre-trained on the ImageNet dataset (Ciresan et al., 2011), (Deng et al., 2009). The advantage of transfer learning is that high discrimination accuracy can be achieved even with a small amount of data prepared by the researcher because it was based on a learned model. In addition, it can be applied to a wide range of domains and tasks that can be transferred.

EXPERIMENTAL METHOD

Shooting Method

Figure 2 showed the coordinate system at the time of insertion. As shown in Figure 2, in this study, the camera angle was adjusted to capture images of a human profile with the direction from the nose to the back of the head as the X-axis, the direction from the front ear to the opposite ear as the Y-axis, and the direction from the neck to the top of the head as the Z-axis. The camera was used to take pictures with a device called EARPICK.

Thirteen subjects attempted to take pictures from five different angles with their left and right ears. First, the case in which the camera was inserted into the ear parallel to the Y-axis, with the ear hole position as the origin, was defined as the "normal shooting" case. Next, the subjects were asked to take photographs with the ear tilted 10 to 20 degrees in a positive direction along the X-axis and similarly tilted in a negative direction on the same axis. Finally, they were asked to take pictures with the ear tilted 10 to 20 degrees positively along the Z-axis and similarly tilted in a negative direction on the same axis. The reason for the ambiguity of 10 to 20 degrees is that the width of the ear can be moved depending on the size of the individual's ear canal, so this time the camera was moved to the very edge of focus to take pictures from a variety of angles. There is a difference.

Questionnaire by Experiment

As a preliminary step to machine learning, a questionnaire was used to verify how human cognitive abilities can identify images of the inner ear canal. The

Table 1. Questionnaire results.

Combination (when registered: when answered correctly)	Percentage of correct answers (max: min) (%)
(Right side: Right side)	72.0(86.7: 53.3)
(Left side: Left side)	78.7(100: 53.3)
(Right side: Left side)	42.7(73.3: 20.0)
(Left side: Right side)	61.3(86.7: 20.0)
Average of 4 patterns	63.7(80.0: 55.0)

questionnaire consisted of a question. Participants were asked to select one image of the external auditory canal and ten alternatives that they thought were identical to the image. The patterns of questions are shown below.

Pattern 1: Select the image of the right ear from the image of the right ear.
Pattern 2: Select the image of the left ear from the image of the right ear.
Pattern 3: Select the left ear image from the left ear image.
Pattern 4: Select the right ear image from the left ear image.

Experiment by Machine Learning

In the transfer learning, the number of classes was classified into 26 classes with left and right ears for 13 people, and 600 to 800 image data per class were randomly assigned to training and validation at a ratio of 4:1. The number of epochs was performed ten times, and the training model was used to evaluate the individuality of the geometrical features within the ear canal.

EXPERIMENTAL RESULTS

Questionnaire by Experiment

Table 1 showed the tabulation of the questionnaire results. The results showed that when the ear at registration and the ear at evaluation was the same, the discrimination performance ranged from 72-78% as the average response rate for all respondents. On the other hand, when the left and right ears at the time of registration and at the time of evaluation were different, the discrimination performance was significantly lower. However, the correct response rate exceeding 50% suggests that some parts are similar, so it will be essential to determine how much this affects the machine learning process.

Experiment by Machine Learning

Table 2 showed the results of the machine learning evaluation. The evaluation metrics were Accuracy, Recall, Precision, and F-measure. The odd-numbered classes were the image data of the left ear of each person, and the even-numbered classes were classified as the data of the right ear. They were also sorted one by one from the top (left and right ears of the same person in class 1 and class 2). Table 2 showed that the prediction results were high, with a 98.9% correct response rate. This confirmed a high degree of individuality in the shape (image) characteristics of an individual's ears. In addition, the same person's left and right ears also showed highly accurate discrimination

Table 2. Shows the results of the machine learning evaluation.

class	Recall	Precision	F-measure
1	0.969	1.000	0.984
2	1.000	1.000	1.000
3	0.975	1.000	0.988
4	1.000	0.961	0.980
5	0.991	0.995	0.993
6	1.000	0.996	0.998
7	0.973	1.000	0.986
8	0.988	0.992	0.990
9	0.979	0.995	0.987
10	1.000	0.991	0.995
11	1.000	1.000	1.000
12	1.000	1.000	1.000
13	0.993	0.974	0.983
14	0.945	1.000	0.972
15	1.000	0.990	0.995
16	1.000	1.000	1.000
17	0.996	1.000	0.998
18	1.000	0.995	0.997
19	1.000	0.946	0.972
20	0.971	1.000	0.985
21	0.974	0.977	0.976
22	1.000	0.988	0.994
23	1.000	0.962	0.981
24	0.979	1.000	0.989
25	0.987	1.000	0.993
26	0.995	0.957	0.976
Accuracy		0.989	

results for each. From these results, it could be said that a personal authentication system using both ears can be used for personal identification with higher accuracy than using only one ear. However, some images were judged to be similar, even though they were considered to be completely different by human recognition ability. This is thought to be due to dark images or white-outs.

CONCLUSION

In this study, we proposed a personal authentication system based on machine learning used image data of the ear canal captured by a camera, and evaluated the system through questionnaires and experiments. The results showed that the questionnaire indicated that human recognition ability was sufficient to recognize the ear canal and eardrum. In addition, a high discrimination rate of approximately 98.9% was achieved in learning, confirming that the human ear canal's image characteristics vary among individuals. It was also confirmed that the left and right ears of the same person had different characteristics. On the other hand, some of the images were recognized as not similar by human recognition ability. In the future, the insertion angle and length of the camera should be strictly considered. In addition, experiments that used photographs and third-party data under actual usage conditions were also necessary.

REFERENCES

AVNET Corporation (November 12, 2018). Biometric identification, overview of six major technologies. AVNET Corporation Website: https://www.avnet.com/wps/portal/japan?locale=ja-JP >

D Ciresan, U Meier, J Masci, L Gambardella, and J Schmidhuber (2011). Flexible, High Performance Convolutional Neural Networks for Image Classification. IJCAI.

J Deng, W Dong, R Socher, L Li, K Li, and F Li (2009). ImageNet: A large-scale hierarchical image database. CVPR.

K Simonyan, and A Zisserman (2014) .Two-stream convolutional networks for action recognition in videos. NIPS.

Ministry of Health (October 21, 2020). Labour and Welfare, Trends in the diffusion of electronic medical record systems, Ministry of Health Website: https://www.mhlw.go.jp/index.html

Mynavi (2019) "Deep Learning and VGG16", in:Learn while you make! Development Deep Learning by Pytorch, O Yutaro. pp. 6–17

NEC. biometric identification. NEC Website: https://jpn.nec.com/

Statistics Bureau (2021). Ministry of Internal Affairs and Communications, Japan, Labor Force Survey, Statistics Bureau Website: https://www.stat.go.jp/data/roudou/sokuhou/tsuki/index.html

U Yusuke, and Y Takayoshi (2017). Research Trend of Convolutional Neural Networks. Information and Communication Engineers.