Facial Expression Recognition Using Infrared Rays Image Processing

5 N－6

YASUNARI YOSHITOMI,† SHUICHI KIMURA,†† EIZOU HIRA†† and SHIGEYUKI TOMITA†

†Department of Computer Science and Systems Engineering, Faculty of Engineering, Miyazaki University,††Mechanics Department, Miyazaki Prefectural Industrial Research Institute

1. Introduction

The present stage of computer vision is far from goal of humanlike capability,especially from the point of understanding human feeling or mind through facial expression. One reason for the difficulty is that the input image from ordinary camera dose not have a big difference between neutral and smiling faces, for example, from the point of gray level distribution. For this reason we try to use an image describing thermal distribution of face, with Infrared Rays (IR), instead of ordinary visible ray. As the first step, the facial expression selected here is neutral or smile.

2. Image Acquisition and Analysis System

In this study, thermal images of human faces were produced by Thermal Video System (Nippon Avionics Co.,Ltd,TVS-3000) with IR. The principle of thermal image generation comes from the well-known law by Stefan and Boltzmann, which is expressed by the following equation.

$$W = \epsilon \cdot \sigma \cdot T^4 \quad (1)$$

,where W is radiant emittance (W/cm^2),ϵ is emissivity,σ is Stefan-Boltzmann constant ($=5.6705\times10^{-12}W/cm^2K^4$), TisTemperature(K).

For human skin, ϵ is estimated at 0.98 to 0.99[1]. In this study, however, the approximate value of 1 was used as ϵ for human skin. The input images were produced under the operating condition which presented lower-gray-level for colder part and higher-gray-level for hotter part. Each face had thermal range by 5 to 10 K with each spatial distribution. The input images were stored on fixed disks and processed by a personal computer (NEC PC-H98S model 8).

keywords:computer vision, image understanding, facial expression, infrared rays

3. Approach

3.1 Recognition Algorithm

The algorithm for recognizing smiling and neutral faces is as follows.

1) A filter for extracting upper reference region and lower characteristic region for a typical set of smiling and neutral faces is produced with manual operation. The filter, which is a binary image with 1 for upper reference region and lower characteristic region, 0 for other region, is basically used for every subject of recognition. The upper skin region which has little chance between smiling and neutral faces is selected as a reference. On the other hand, the lower skin region containing mouth and its surrounding which area becames bigger in smiling is also selected as a characteristic part. The central region containing nose and eyes is included in neither reference nor characteristic regions since some persons wear glasses which prevent IR from penetrating them.

2) The input-gray-image is transformed into binary image with Otsu's segmentation method[2].

3)The logical product for the binary image and the filter is performed, followed by erasing smaller areas which have no meaning for recognizing smiling face.

4)The relative area (S/S_0) of the characteristic part, where S, S_0 are the area of the characteristic part, the reference part respectively, is measured, followed by the normalization that the relative area (S/S_0) is multiplied by a coefficient (C) which transforms the smallest relative area for each male neutral face in a series of scenes into unit, viz., 1. The normalized value $C(S/S_0)$ is hereinafter referred as SF (Smiling Factor).

5)The measured SF is compared with a threshold. The threshold which is the lowest

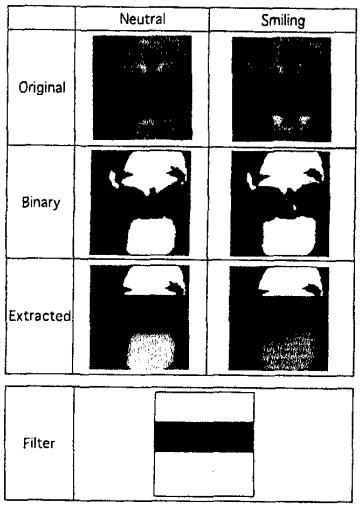

Fig. 1 Examples of the filter, original, binary and extracted images of the main male.

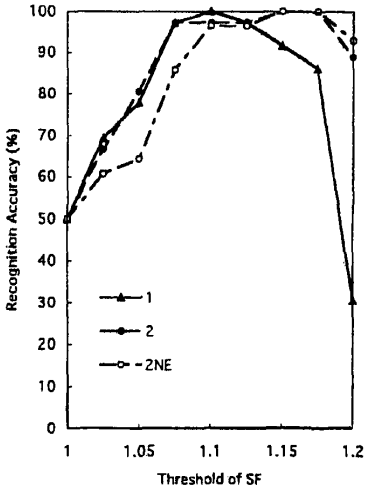

Fig. 2 Relationship between the threshold of SF value and the recognition accuracy of neutral and smiling faces for the main male. 1,2 and 2NE indicate the first, second and second (naked eyes) image-recording-day, respectively.

SF value judged to be from smiling face is decided as an experimental parameter. Therefore, the measured SF is judged to be from smiling face, if it is not smaller than the threshold.

Fig.1 shows examples of the filter, original, binary and extracted images. The extracted regions consist of the reference and characteristic parts which are described with white and gray, respectively.

3.2 Experimental and Discussion

Image sequences of smiling or neutral faces were collected. We assembled a file of 114 pictures for the main male (100) and the second male (14). The filter produced for the main male was also used for the second male face.

Fig.2 shows the relationship between the threshold of SF value and the recognition accuracy of smiling and neutral faces, for the main male. The best threshold for recognizing smiling and neutral faces hardly depended on glasses. However, the observation day influenced the relationship between threshold of SF and recognition accuracy of smiling and neutral faces. It might be mainly caused by change of mental situation. Nevertheless,the highest recognition rate with each best threshold of SF was 100% for every sequence of image with or without glasses. Moreover, the recognition rate with the best threshold of SF, 1.10 to 1.125, was higher than 96% for all images.

For the second male, the highest recognition rate with the best threshold of SF was 86%.

The imperfect recognition is considered to be mainly caused by smaller difference between smiling and neutral faces of the second male, compared with the main male. However, the recognition rate with the best threshold of SF, 1.10, was higher than 86% for all images of both the main and second males.

The future system will be also able to distinguish more precisely one person from others because the thermal image of face has more fruitful information than ordinary image with visible ray.

4. Conclusion

A method is presented for recognition of facial expression. The method is based on 2-dimensional detection of temperature distribution of face, using Infrared Rays. The area ratio of hotter region consisting of mouth and its surrounding is measured to identify smiling face, using image processing. The area ratio becomes bigger in smiling. The method can be applied to both cases with and without glasses. The recognition accuracy for smiling and neutral faces reached to 100%. The right recognition rate hardly depended on scene.

References

1) Kuno H.:Sekigaisen Kougaku,pp.22,pp.45,IEICE,Tokyo (1994). (in Japanese).
2) Otsu N.:A Threshold Selection Method from Gray-Level Histograms, Trans. IEEE, Vol. SMC-9, No. 1, pp. 62–66 (1979).