Estimation of Learners' Subjective Difficulty in e-Learning Using Thermal Image Processing

Yuki Yoshimitsu and Masayoshi Tabuse
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University
1-5 Hangi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan
E-mail: y_yoshimitsu@mei.kpu.ac.jp, tabuse@kpu.ac.jp

Abstract

In recent years, e-learning has been utilized as a learning system in many schools. In a conventional class, a teacher teaches learners face-to-face, so that a teacher observes facial expressions and posture change of learners and estimates learners' subjective difficulty of the lecture. On the other hand, e-learning is usually utilized in the self-study. A learner learns a teaching material using a computer by oneself, so that a teacher can't observe the state of the learner. It is difficult for a teacher to estimate learners' subjective difficulty. Therefore, we propose a learners' subjective difficulty estimation system. This system captures the face of a learner in e-learning with a thermal camera. It extracts the face region, measures the temperature changes of the nose region and the forehead region of the learner, and estimates the subjective difficulty of the learner based on the temperature changes.

Keywords: learners' subjective difficulty, thermal image, e-learning, temperature changes.

1. Introduction

In recent years, e-learning has been utilized as a learning system in many schools. In a conventional class, a teacher teaches learners face-to-face, so that a teacher observes facial expressions and posture change of learners and estimates learners' subjective difficulty of the lecture. On the other hand, e-learning is usually utilized in the self-study. A learner learns a teaching material using a computer by oneself, so that a teacher can't observe the state of the learner. It is difficult for a teacher to estimate learners' subjective difficulty. Therefore, we estimate a learners' subjective difficulty. Because the thermal image of the human face changes depending on a bloodstream change with the automatic nerve activity, the change of the thermal image of the human face, especially the nose region, reflect human physiological psychological conditions.[1]

Our proposed system captures the face of a learner in e-learning with a thermal camera. It extracts the face region, measures the temperature changes of the nose region and the forehead region of the learner, and estimates the subjective difficulty of the learner based on the temperature changes.

2. Thermal Image

Thermal image is obtained according to the Stefan-Boltzmann law which describes the power radiated from a black body in terms of its temperature, expressed as $W = \varepsilon\sigma T^4$, where ε is emissivity, σ is the Stefan–Boltzmann constant (=5.6705×10^{-12} W/cm^{2}K^4), and T is the temperature (K). For human skin, ε is estimated as 0.98 to 0.99.[2] The human face area is easily extracted from an image by using the value of 1 for ε when a range of skin temperatures is selected to produce a thermal image. Fig. 1 shows example of thermal image of a male. In our study thermal image is expressed in gray image with 256 gray levels where gray value 0 and 255 indicate temperatures beyond the setup minimum and maximum ones respectively.

Fig. 1. Thermal image of a male.

3. Extraction of face region

When the setup temperatures are chosen so that the human skin temperature might enter between the minimum temperature and the maximum temperature, the human region are expressed by gray values between 0 and 255 and the lower temperature background are expressed by 0. We first operate binary image processing to thermal image and obtain binary image, shown in Fig. 2. The human region has value 255 and the background has 0. We apply labeling to the binary image and have regions with more than 2000 pixels, shown in Fig. 3. Next we scan labeling image from $y=0$ to $y=H$ (height of thermal image) for each x and when the coordinate (x, y) is in the labeling region, we set 1 on 2 dimensional array $M(x, y)$ from there to $y=H$. We scan labeling image from $x=0$ to $x=W$ (width of thermal image) and from $x=W$ to $x=0$ for each y. When the coordinate (x, y) is in the labeling region, we add 1 on $M(x, y)$ from there to right and from there to left, respectively, shown in Fig. 4.

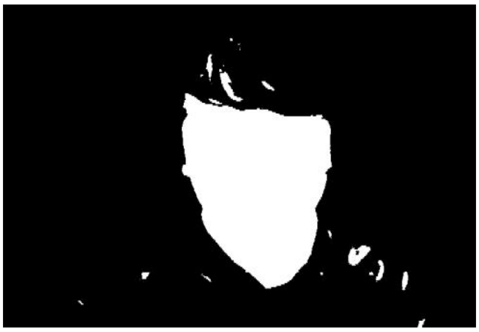

Fig. 2. Binary image.

Fig. 3. Labeling.

Fig. 4. 2 dimensional array $M(x,y)$. White, light gray, gray and black colors indicate the area with the values of 3, 2, 1 and 0.

We operate the raster scan on $M(x, y)$ from $(0,0)$ to (W, H) and find the top coordinate of the human region (x_{human}, y_{human}) where the value of $M(x, y)$ has 3. We measure the maximum width $W_{\max}$ of the human region from y_{human} to *y*-coordinate of the center of mass of the largest labeling region. From y_{human} to *y*-coordinate of the bottom of the largest labeling region we calculate the frequency distribution of the width of the region which have 2 or 3 and determine the width of the maximum frequency within $1.2W_{\max}$ as the facial width W_{face} and x-coordinate of the left and right boundaries of the human region as left and right boundary of the face region x_{left} and x_{right}, respectively. We find *y*-coordinate of $0.8W_{face}$ width of the human region downward from y_{human} and the top y-coordinate of the face region y_{top} is $0.2W_{face}$ above from there. The bottom of the face region y_{bottom} is W_{face} downward from y_{top}, shown in Fig. 5. The upper left and the lower right coordinates of the face region are (x_{left}, y_{top}) and (x_{right}, y_{botoom}). The height of face region is $H_{face} = W_{face}$.

Fig. 5. Face region.

Finally we operate a rotation correction to the thermal image. We measure midpoints of the left and right boundaries of binary image in the face region from $y_{top} + 0.45H_{face}$ to $y_{top} + 0.75H_{face}$. We estimate a roll angle θ of the face using the least-square method (Fig. 6.) and rotate the thermal image by $-\theta$ around the center of the face region. We normalize the face region with the size of 400×400 pixels.

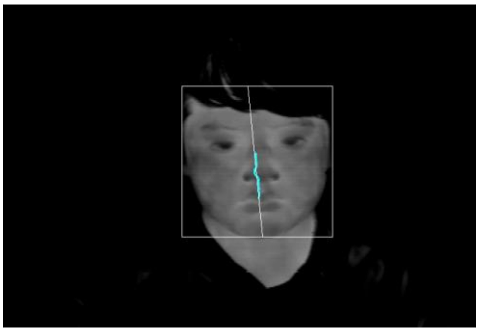

Fig. 6. Estimation of a roll angle.

4. Template matching, the nose and forehead regions

We utilize the template matching to detect the nose region correctly. We first click the tip of the nose and obtain a template. The size of a template is $H_{tmp} \times W_{tmp} = 140 \times 280$ pixels and the upper left coordinate is $(x_{click} - 0.5W_{tmp}, y_{click} - 0.85H_{tmp})$ where (x_{click}, y_{click}) is a coordinate of the tip of the nose in the face region. The nose region is $(x_{tmp} + 0.35W_{tmp}, y_{tmp} + 0.2H_{tmp})$ - $(x_{tmp} + 0.65W_{tmp}, y_{tmp} + H_{tmp})$ where (x_{tmp}, y_{tmp}) is the coordinate of the tip of the nose. The forehead region is $(x_{tmp} - 0.35W_{tmp}, y_{tmp} - 40)$ - $(x_{tmp} + 0.65W_{tmp}, y_{tmp})$, shown in Fig. 7.

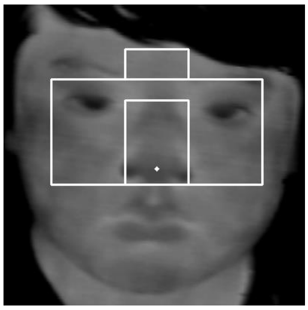

Fig. 7. Template, nose region and forehead region.

5. Experiments

5.1. *Condition*

The thermal image produced by the infrared thermal image camera (Nippon Avionics TVS-700) is output as a composite video signal, which is input to the PC (Dell Optiplex 9010, CPU: Intel Core i7-3770 3.4GHz, main memory: 8GB and OS: Microsoft Windows7 Professional 64bit) through the video capture device (I-O DATA GV-USB2/HQ). The thermal image is expressed in gray image with 256 gray levels and the setup minimum and maximum temperatures are 29℃ and 39℃, respectively. The thermal image size is 704× 480 pixels. We used Microsoft Visual C++ 2013 Express as a programming language and OpenCV 2.4.11 as image processing library. Subject A is a male in his 20s.

5.2. *Mental workload*

Subjects perform a mental calculation as a mental workload.[1,3] Fig. 8 shows a mental calculation system which gives questions of mental calculation of double-, triple- or quadruple-digits. The system displays a number and the subject memorizes it and clicks the enter button. Then the system displays a different number, the subject adds it to the memorized number, input the answer and clicks the enter button. The system records the number of the answers, the number of the correct answers and answer time for each question.

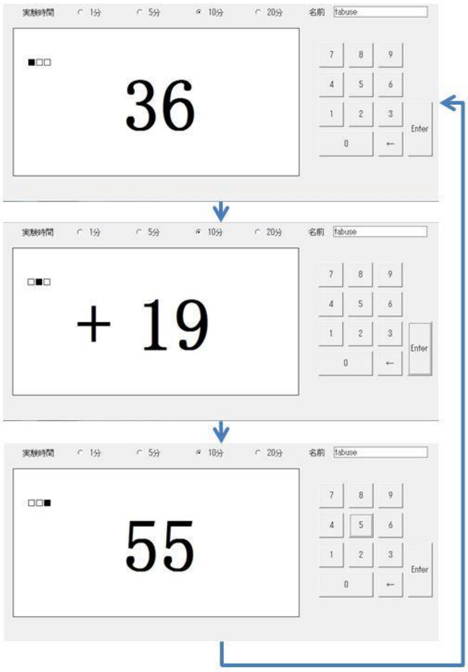

Fig. 8. Mental calculation system.

5.3. *Results*

The change of temperatures of the face region, the nose region and the forehead region of subject A with quadruple-digits mental calculation are shown in Fig. 9, 10, 11. Answer time for each question and time series variation of correct and incorrect answers are shown in Fig. 12, Fig. 13.

6. Conclusion

We measure the change of temperatures of the face region, the nose region and the forehead region of a subject while doing mental calculations.

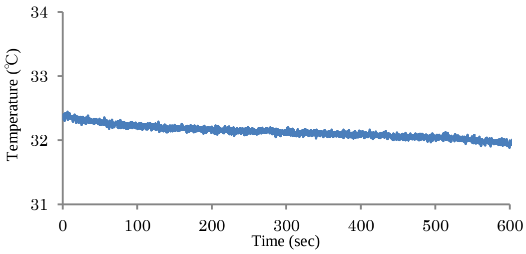

Fig. 9. Average temperature of the face region.

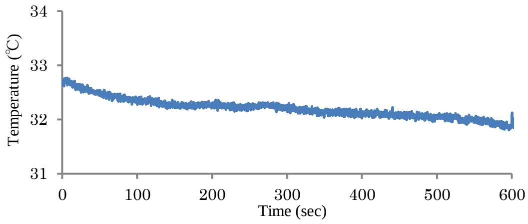

Fig. 10. Average temperature of the nose region.

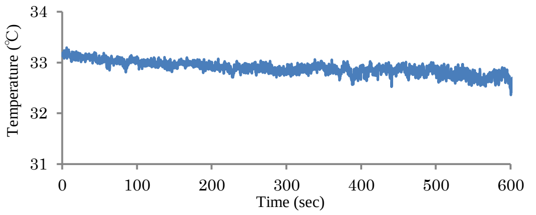

Fig. 11. Average temperature of the forehead region.

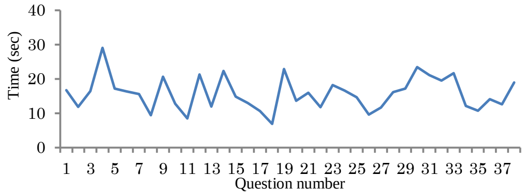

Fig. 12. Answer time.

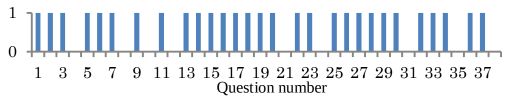

Fig. 13. Time series variation of correct and incorrect answers (1: correct answer, 0: incorrect answer).

References

1. T. Mizuno, S. Nomura, A. Nozawa, H. Asano and H. Ide, Evaluation of the effect of intermittent mental work-load by nasal skin temperature (in Japanese), IEICE J93-D(4) 535-543 (2010).
2. H. Kuno, *Infrared rays engineering* (in Japanese). (IEICE, Tokyo, 1994).
3. S. Kunimasa, K. Seo, H. Shimoda and H. Ishii, The inference method of cognitive working states while performing mental tasks based on performance-cognitive load model (in Japanese), The Transactions of Human Interface Society 17(4) 395-410 (2015).