Real-Time System for Horizontal Asymmetry Analysis on Facial Expression and Its Visualization

Ryoichi Shimada, Taro Asada, Yasunari Yoshitomi, Masayoshi Tabuse
Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku,Kyoto 606-8522, Japan
E-mail: {yoshitomi, tabuse}@kpu.ac.jp, t_asada@mei.kpu.ac.jp}
http://www2.kpu.ac.jp/ningen/infsys/English_index.html

Abstract

We report on the development of a real-time system for facial expression horizontal asymmetry analysis and visualization. In our system, the image signal input from a webcam is analyzed using OpenCV, and feature parameters (facial expression intensity) are measured separately for the left and right half regions in the mouth area. Our real-time system then draws a graph expressing facial expression intensity changes. The experimental results obtained thus far suggest that this system could be useful.

Keywords: Facial expression analysis, Real-time system, Moving image, Mouth area, Visualization, OpenCV

1. Introduction

In Japan, the average age of the population has been increasing, and this trend is expected to continue into the future. We have therefore been studying ways of applying information technology (IT) to improving the medical treatment provided to elderly people with psychiatric disorders.

One of our current studies is aimed at analyzing the facial expressions of persons in a moving image captured by webcams. In this paper, we propose a real-time facial expression horizontal asymmetry analysis and visualization system. Unlike current and previously used methods[1,2] in which the moving image data are first saved in a personal computer (PC) and then analyzed, the moving image in our method is captured by a webcam and analyzed in real-time via Open Source Computer Vision (OpenCV) image-processing software[3] along with a previously proposed feature parameter (facial expression intensity) [1,2] that is based on the mouth area. The visualization method proposed in this paper is based on our reported method[4] which involves creating a graph from the OpenCV image data and then analyzing the relationship between the facial expression intensity and time. This visualization is performed concurrently with the facial expression intensity measurement.

2. Proposed System and Method

2.1. System overview and outline of the method

In this system, webcam moving imagery captured in real-time is analyzed via the following process. We constructed the modules in our system based on our reported research.[4] First the sizes of faces in the captured image data are standardized, and then analyzed using OpenCV and our proposed facial expression feature parameters.

The frames of the moving images are then changed from RGB image data into YCbCr image data, after which the Y component obtained from each frame in the dynamic image is used for facial expression analysis. The proposed method consists of (1) the extracting the mouth area, (2) measuring the facial expressions feature vectors, and (3) measuring the facial expression intensity. In (1), the mouth area is extracted from the frames by using OpenCV. In (2), for the Y component of the selected frame, facial expressions feature vectors are extracted for each 8×8-pixel section. In (3), the facial expression intensity, which is defined as the norm difference between the facial expression feature vector of the reference and target frames, is measured. These details are explained in the following subsections.

2.2. Mouth area extraction

First, moving image data are changed from RGB to YCbCr image data, after which the face area is extracted from the changed image as a rectangular shape, and the lower 40% portion of the face area is standardized. Next, the mouth area is extracted from that area. The reason the mouth area was selected for facial expression analysis is because it is where the differences between neutral and happy facial expressions appear most distinctly. An example of the moving image frame data and the extracted mouth area image is shown in Fig. 1.

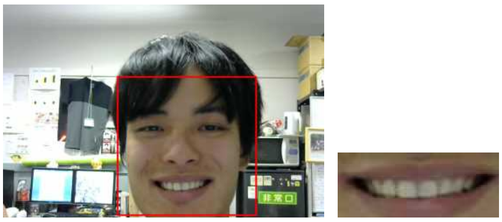

Fig. 1. Moving image frame and extracted mouth area.

In this paper, the mouth area is separated into left and right half regions, and the facial expression intensity of each region is separately measured. An example of a mouth area separated into the left and right half regions is shown in Fig. 2.

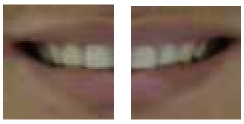

Fig. 2. The mouth area separated in the left and right half regions

2.3. Facial expression intensity measurement

For the Y component of the selected frame, the facial expression feature vector is extracted for the mouth area using a two-dimensional discrete cosine transform (2D-DCT) for each 8×8-pixel section. To measure the feature parameters of the facial expressions, we selected low-frequency components from the 2D-DCT coefficients as the facial expression feature vector elements. However, the direct current component is not included. In total, 15 feature vector elements are obtained. As mentioned above, facial expression intensity is defined the norm difference between the facial expression feature vector of the reference and target frames. The reference frame selection method is explained in next subsection.

2.4. Reference frame selection

In this subsection, we propose a new method for automatically selecting reference frames. This method was adopted because, in our new system, the moving image captured by webcam is analyzed in real-time. This means we cannot use our previous method,[1,2] which was designed for the non-real-time analysis.

In this method, the first 10 continuous frames of mouth area data successfully extracted after the webcam recording begins are treated as reference frame candidates. Then, beginning with the first reference frame candidate, the facial expression intensity of all candidate frames is measured, and the sum of all facial expression intensities is calculated. The candidate frame with the minimum value is selected as the reference frame.

2.5. Facial expression intensity visualization

Facial expression intensity is measured using our previously discussed method,[4] in which a straight line is drawn on a graph image prepared using OpenCV.

3. Experiment

The experiment was performed on a Dell OPTIPLEX 780 PC equipped with Intel Core 2 Duo E8400 3.0 GHz central processing units (CPUs) and 4.0 GB of random access memory (RAM). The Microsoft Windows 7 Professional operating system (OS) was installed on the PC and Microsoft Visual C++ 2008 Express Edition was used as the development language.

The newly proposed system display is shown in Fig. 3. We performed evaluation experiments for two test subjects (hereafter Subjects A and B) under the three conditions listed below. Furthermore, the message “SMILE, PLEASE!” was flashed on the display screen at points 20 and 40 s from the system start.

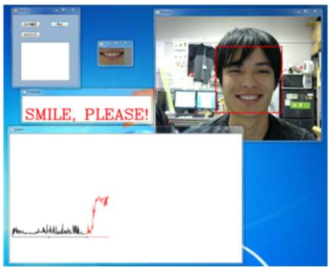

Fig. 3. Display while new system is in operation.

1. Subjects maintained neutral faces for about 60 s.
2. Subjects smiled when the above message was displayed.
3. Subjects raised the right side of their mouth when the above message was first displayed, and raised the left side of their mouth when the above message was displayed a second time.

The experimental results are shown in the following figures (Figs. 4 to 16). Arrows in each graph indicate the max point of the facial expression intensity.

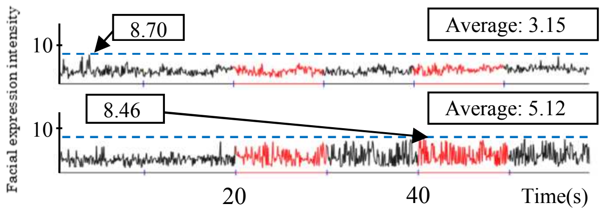

Fig. 4. Facial expression intensity of the whole mouth under experimental condition 1 (upper: Subject A, lower: Subject B).

Fig. 5. Reference frame under experimental condition 1 (left: Subject A, right: Subject B).

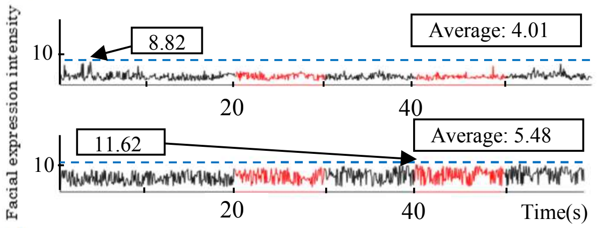

Fig. 6. Facial expression intensity of the right half region of the mouth under experimental condition 1 (upper: Subject A, lower: Subject B).

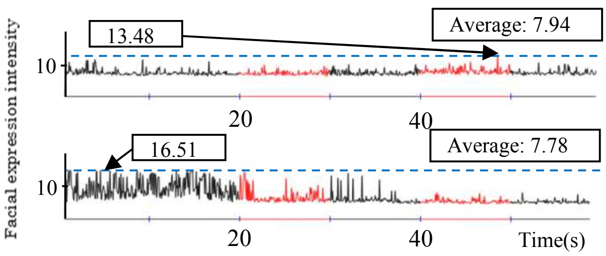

Fig. 7. Facial expression intensity of the left half region under experimental condition 1 (upper: Subject A, lower: Subject B).

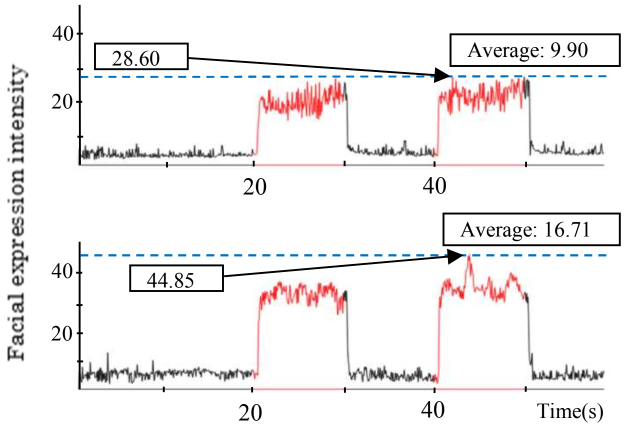

Fig. 8. Facial expression intensity of the whole mouth under experimental condition 2 (upper: Subject A, lower: Subject B).

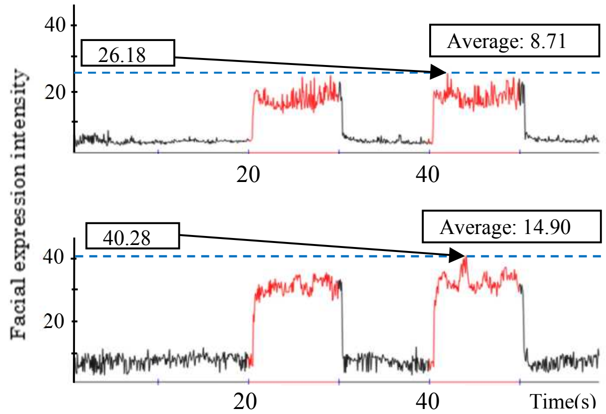

Fig. 9. Facial expression intensity of the right half of the mouth under experimental condition 2 (upper: Subject A, lower: Subject B)

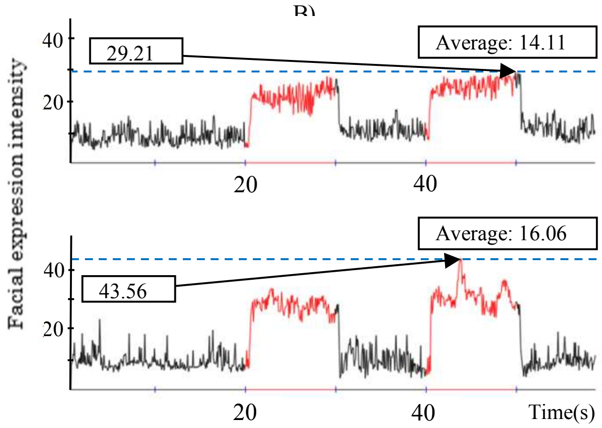

Fig. 10. Facial expression intensity of the left half region under experimental condition 2 (upper: Subject A, lower: Subject B).

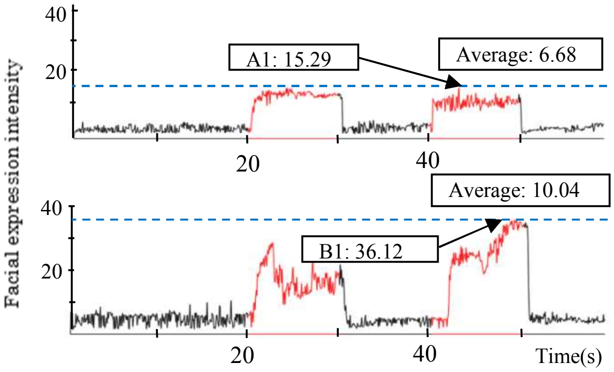

Fig. 11. Facial expression intensity of the whole mouth under experimental condition 3 (upper: Subject A, lower: Subject B).

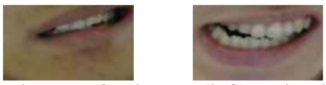

Fig. 12. Mouth area of point A1 (left) and point B1 (right).

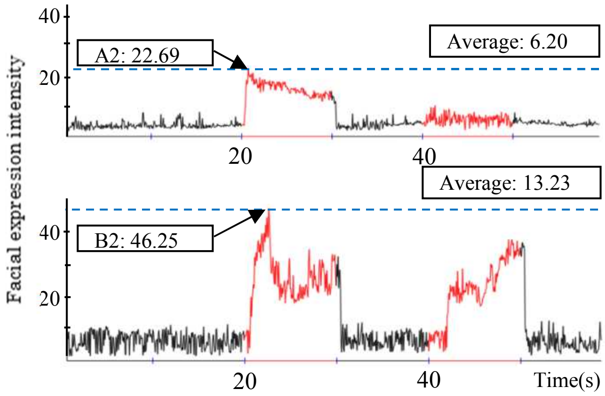

Fig. 13. Facial expression intensity of the right half of the mouth under experimental condition 3 (upper: Subject A, lower: Subject B).

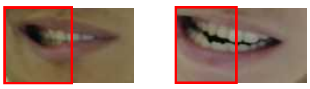

Fig. 14. Right half mouth area of point A2 (left) and point B2 (right).

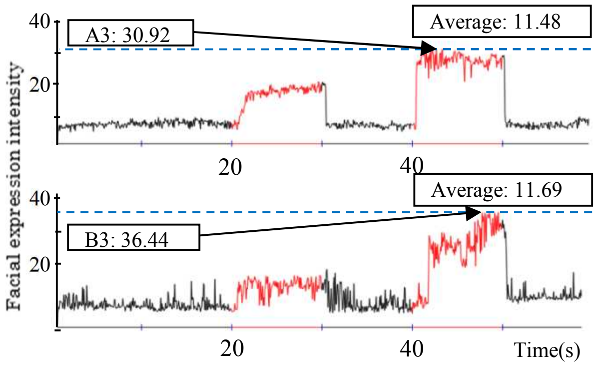

Fig. 15. Facial expression intensity of the left half region under experimental condition 3 (upper: Subject A, lower: Subject B).

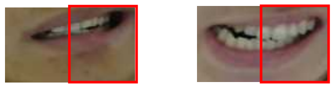

Fig. 16. Left half mouth area of point A3 (left) and point B3 (right).

The experimental results obtained thus far suggest that this system could be useful for situations where real-time reaction to facial expressions is one of the important factors. Upon examining the experimental results, some problems were discovered. Facial expression intensities on the left half of the subject's faces during neutral face recording were always higher than those of the right half, and the reason for this discrepancy is unclear.

4. Conclusion

Herein, we proposed a real-time system for facial expression horizontal asymmetry analysis and visualization and performed a number of evaluation experiments. It is believed that this system could prove useful in the treatment of psychiatric ailments such as depressive disorder and dementia, in addition to providing treatment options for persons with half side body paralysis.

Acknowledgements

The authors would like to thank Professor J. Narumoto of Kyoto Prefectural University of Medicine for his valuable support and helpful advice in the course of this research. We would also like to thank Mr. K. Nishimura, a student of the Graduate School of Kyoto Prefectural University, for his cooperation in the experiments. This research was supported by COI STREAM of the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

1. T. Asada, Y. Yoshitomi, R. Kato, M. Tabuse, and J. Narumoto, Quantitative evaluation of facial expressions and movements of persons while using video phone, *J. Robotics, Networking and Artif. Life* **2**(2) (2015) 111-114.
2. T. Asada, Y. Yoshitomi, R. Kato, M. Tabuse and J. Narumoto, Analysis of facial expressions robust against small imperfections in mouth-part area extraction from face-images of persons while using video phone (in Japanese), in *Proc. of Human Interface Symposium 2015* (Japan, Hakodate, 2015), pp.187-190.
3. Open CV. http://opencv.org/ Accessed 1 December 2015.
4. R. Shimada, T. Asada, Y. Yoshitomi, M. Tabuse, and J. Narumoto, A real-time system for facial expression analysis and its visualization (in Japanese), in *Proc. of Human Interface Symposium 2016* (Japan, Tokyo, 2016), pp.643-646.