Vision Based Body Dither Measurement for Estimating Human Emotion Parameters

Sangin Park, Deajune Ko, Mincheol Whang, and Eui Chul Lee*

Department of Computer Science & Department of Emotion Engineering,
Sangmyung University, Seoul, Republic of Korea
ini0630@naver.com, kodeajune@gmail.com, {whang,eclee}@smu.ac.kr

Abstract. In this paper, we propose a new body dither analyzing method in order to estimating various kinds of intention and emotion of human. In previous researches for quantitatively measuring human intention and emotion, many kinds of physiological sensors such as ECG, PPG, GSR, SKT, and EEG have been adopted. However, these sensor based methods may supply inconvenience caused by sensor attachment to user. Also, therefrom caused negative emotion can be a noise factor in terms of measuring particular emotion. To solve these problems, we focus on facial dither by analyzing successive image frames captured from conventional webcam. For that, face region is firstly detected from the captured upper body image. Then, the amount of facial movement is calculated by subtracting adjacency two image frames. Since the calculated successive values of facial movement has the form of 1D temporal signal, all of conventional temporal signal processing methods can be used to analysis that. Results of feasibility test by inducing positive and negative emotions showed that more facial movement when inducing positive emotion was occurred compared with the case of negative emotion.

Keywords: Body dither measurement, Emotion recognition, Image subtraction.

1 Introduction

In previous researches for quantitatively measuring human intention and emotion, many kinds of physiological sensors (ECG (electrocardiography), PPG (photoplethysmography), GSR (galvanic skin response), SKT (skin temperature), and EEG (electroencephalography)) based methods have been adopted [1]. Conventionally, in case of using ECG or PPG, heart rate can be measured by analyzing successive pulse to pulse intervals [2]. Also, amplitude levels are analyzed in order to respectively measure skin response and skin temperature in case of using GSR and SKT. Especially, EEG data can be interpreted by using various kinds of method at time domain or frequency domain. However, the above mentioned methods may supply inconvenience caused by sensor attachment to user. Also, therefrom caused negative emotion can be a noise factor in terms of measuring particular emotion.

* Corresponding author.

M. Kurosu (Ed.): Human-Computer Interaction, Part V, HCII 2013, LNCS 8008, pp. 346–352, 2013.

Recently, some camera vision based physiological data acquisition methods were proposed which is instead of the above mentioned conventional physiological sensors. The Cardio-Cam was proposed for measuring human heart rate by ICA (Independent Component Analysis) based color channel analyses without any sensor attachment [3]. For the same purpose, many smartphone applications have been released which can real-timely measure heart rate by using built-in backside camera and white illuminator. In these applications, they used a concept that the brightness levels of successive images were changed because the amount of illuminative reflection is continuously and regularly changed according as the blood flow. Even though the mentioned methods are meaningful in terms of no sensor attachment, they only can decode heart rate.

To solve these problems, we focus on various kinds of dither of human body. In our analyses, we define the hierarchical model of human body as shown in figure 1 which can clarify dependency (or independency) of each part's dither. For convenience in this paper, only facial dither is analyzed.

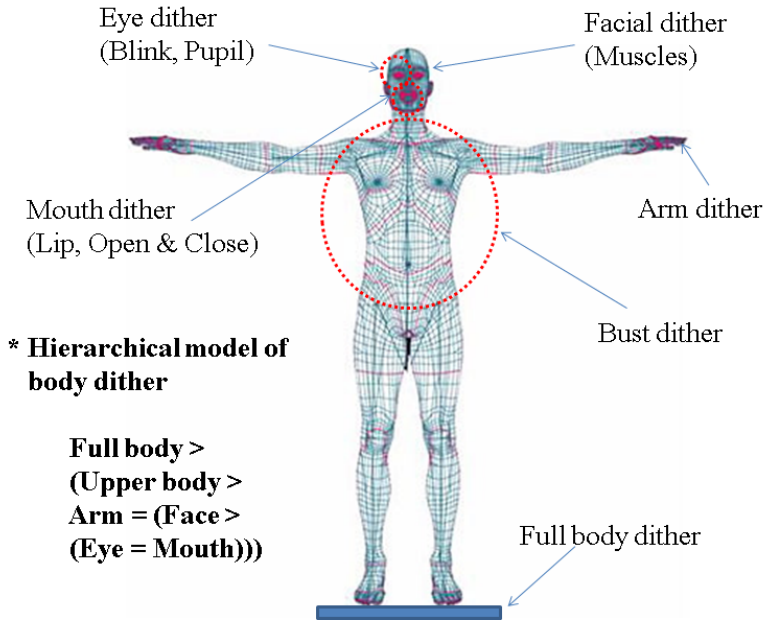

Fig. 1. Definition of various dithers of human body and hierarchical model

To confirm the feasibility of dither analyzing based emotion measurement, facial dither is analyzed in this paper. In captured upper body image, face region is firstly detected by using Adaboost (adaptive boosting) method. Then, the amount of facial movement is calculated by subtracting successive two images. At result of comparing two groups which respectively inducing positive and negative emotions during 10 minutes, more dither was occurred in case of positive emotion group than the case of negative emotion.

2 Proposed Method

Firstly, face region is detected in the 1st frame of upper body image by using OpenCV Adaboost face detector. Face detections of every frames do not needed because our propose method extracts the amount of facial motion by subtracting pixels of same position of two successive image frames. The Adaboost method uses a strong classifier generated by combining simple weak classifiers to detect face on an input image [4]. Although this algorithm takes much training time, it has advantages such as rapid time required for detection and good detecting performance. It took 29ms per an image in average to detect facial region. Examples of face detection in our experimental images are shown in red rectangles of Fig. 2.

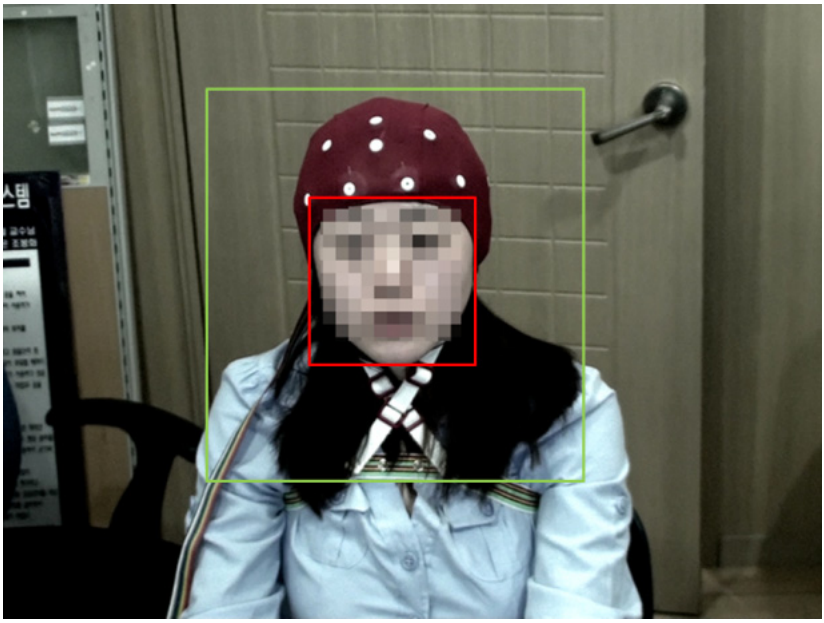

Fig. 2. Face detection results using Adaboost. (Red: Detected facial region, Green: Defined candidate region of facial dither).

After face detection shown in Fig. 2, the candidate region for subtracting image to calculate facial dither is defined by expanding 80 pixels directed to four direction of facial region rectangle as shown in green rectangles of Fig. 2.

To measure the amount of dither, the camera vision analysis program was implemented. In dither analyses, the captured color image is converted to gray level one because color component is not important in terms of estimating motion. The average amount of facial dither (M) can be calculated as following equation.

$$\mathrm{M} = \frac{1}{WH}\sum_{j=y}^{H}\sum_{i=x}^{W}|I_n(i,j) - I_{n-1}(i,j)| \tag{1}$$

In equation (1), W and H are the horizontal and the vertical length of the facial dither candidate region respectively. And $I_n(i, j)$ means pixel value of ith column and jth row of nth image frame. Example of successive image frames and their dither extraction result is shown in Fig. 3.

After all, since the extracted continuous M values in equation (1) generate 1D temporal signal as shown in figure 4, it can be analyzed by using same way of conventional signal analyses methods. Although previous background subtraction methods for object detection have problem of continuously changed background or complex background modeling, our proposed is independent upon background changes because the proposed method uses only the latest two image frames.

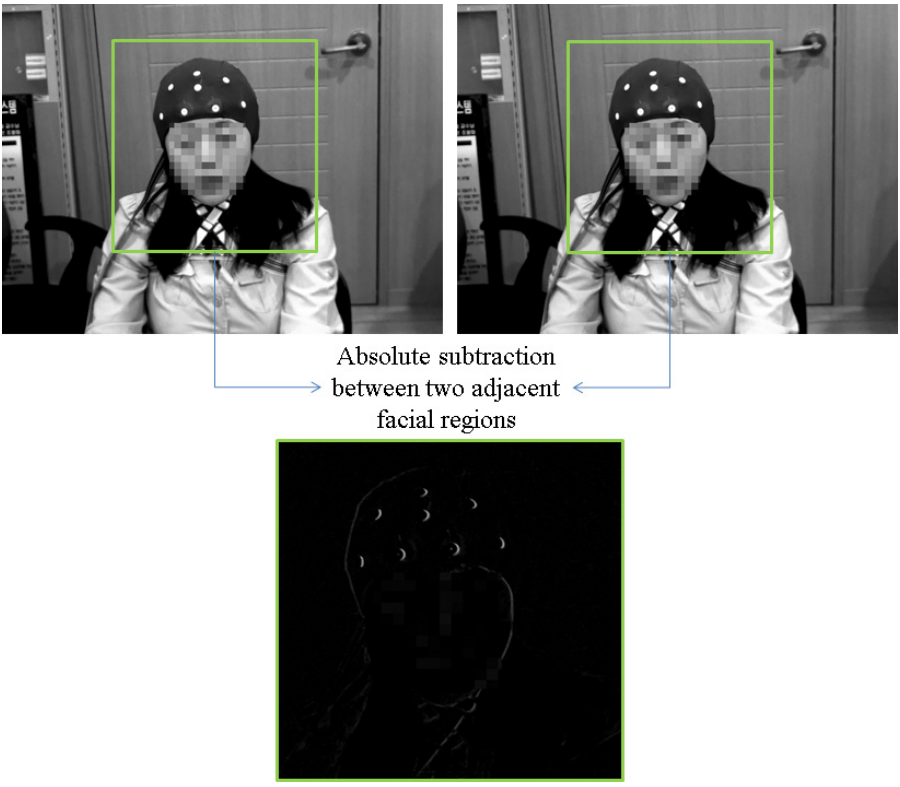

Fig. 3. Example of successive image frames (left and middle) and their extraction result (right)

The proposed method can be used to any part of human body mentioned in figure 1. If the previous implemented particular body part detection method is combined, their dither and muscle movement can be analyzed. For example, if Adaboost (Adaptive boosting) based face detection method is used, only face part dither can be analyzed. Fig. 5 shows many kinds of dither detecting results by using the program. In the figures, bright regions are regions in dither.

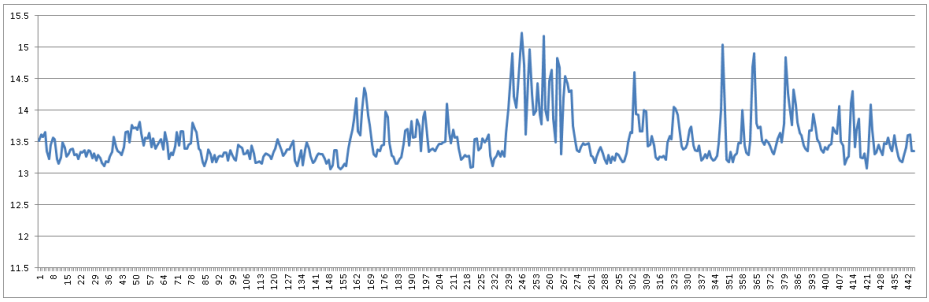

Fig. 4. Example of the successive value of facial dither

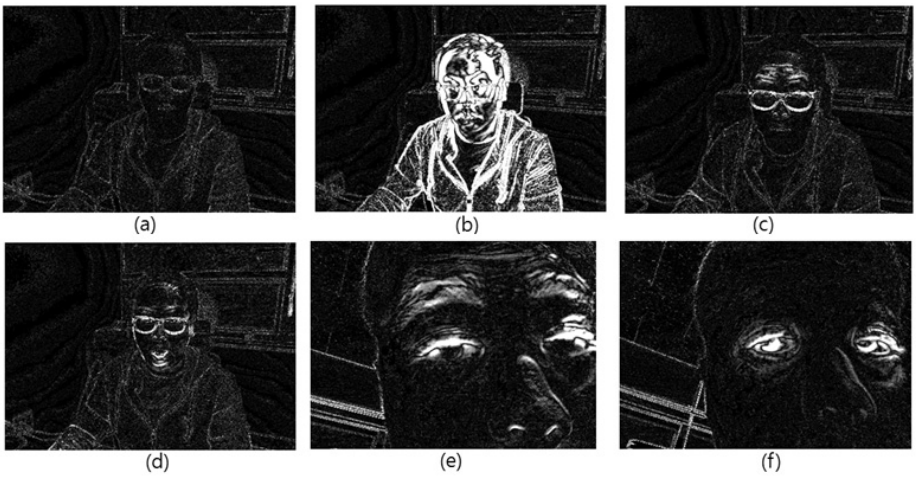

Fig. 5. Dither detection results. (a) No dither. (b) Bust dither. (c) Upper facial muscle movement. (d) Mouth movement. (e) Eye blink. (f) Changing gaze direction.

3 Experimental Result

To confirm the feasibility of our proposed method, a test for classification between positive and negative emotions was performed. For that, 10 persons were participated in which each subject heard announcement which causing positive and negative emotions during 10 minutes. During hearing the announcement, upper body image was captured by using a conventionally used webcam as resolution of 640 pixels by 480 pixels and 15 frames per second. Consequently, our method analyzed the amount of dither at 15Hz frequency band.

The average result of the dither amounts for two groups is shown in Fig. 6. According to this result, we found that the amount of dither for inducing positive emotion was greater than the case of negative emotion. Also, the difference between two cases for causing two contrary emotions was statistically significant in terms of *t*-test based average difference validation [5].

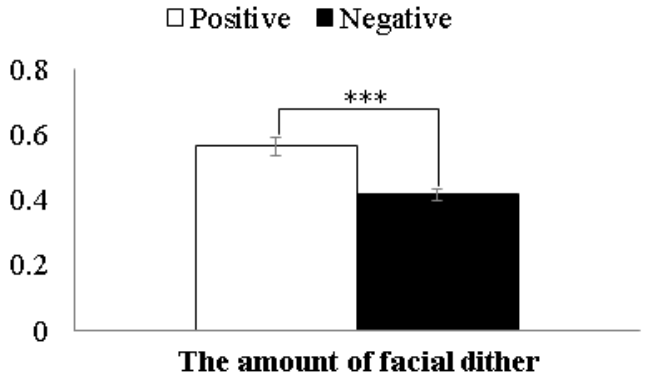

Fig. 6. The average amount of dither for two groups such as inducing positive and negative emotions (***: statistically significant at the confidence level of 99%)

4 Conclusion

In this paper, we proposed a new body dither analyzing method in order to estimating various kinds of intention and emotion of human. To solve problems of previously performed bio-signal based emotion measuring methods, we focused on facial dither by analyzing successive image frames captured from conventional webcam. After that, the amount of facial dither was measured by subtracting adjacency two image frames. Because the measured successive values of facial movement were the form of 1D temporal signal, all of conventional temporal signal processing methods might be used. Results of feasibility test by inducing positive and negative emotions showed that more facial movement when inducing positive emotion was occurred compared with the case of negative emotion.

In future works, we will experimentally validate connectivity between each part's dither of human body and various kinds of conventional physiological responses. For example, we will analyze correlation between pulse to pulse interval and the amount of bust dither after acquiring both ECG signal and bust dither for specific visual stimulus. Also, our dither analysis will be performed in terms of various frequency bands.

Acknowledgements. This work was supported by the by the Global Frontier R&D Program on <Human-centered Interaction for Coexistence> funded by the National Research Foundation of Korea grant funded by the Korean Government (MEST) (2012-055701).

References

1. Wu, N., Jiang, H., Yang, G.: Emotion recognition based on physiological signals. In: Zhang, H., Hussain, A., Liu, D., Wang, Z. (eds.) BICS 2012. LNCS, vol. 7366, pp. 311–320. Springer, Heidelberg (2012)
2. Chang, F.C., Chang, C.K., Chiu, C.C., Hsu, S.F., Lin, Y.D.: Variations of HRV analysis in different approaches. In: Proc. of Computers in Cardiology, pp. 17–20 (October 2007)
3. http://www.livescience.com/15469-cardiocam-mirror-mit-siggraph.html (accessed on March 17, 2013)
4. Han, P., Liao, J.: Face detection based on Adaboost. In: Proc. of Apperceiving Computing and Intelligence Analysis, pp. 337–340 (October 2009)
5. Johnson, D.H.: The insignificance of statistical significance testing. The Journal of Wildlife Management 63(3), 763–772 (1999)