

Emotional state recognition using facial expression, voice and physiological signal

Abstract

Emotion recognition is an important aspect of affective computing; one of whose goals is the study and development of behavioural and emotional interactions between humans and animated conversational agents. Discoveries from neurophysiology and neuropsychology,¹ which establish a strong link between emotion, rationality and decision-making, have intensified research in this area for the consideration of emotions in human multimodal interactions, especially in health and robotics areas. This research has given birth to new scientific and technological tracks, largely related to the modeling and emotion recognition.

Keywords: emotion recognition, physiological signal, pattern recognition, affective computing

Volume 4 Issue 3 - 2018

Tahirou DJARA, Matine OUSMANE, Abdoul-Aziz SOBABE, Antoine VIANOU

Laboratory of Electronic Engineering, Telecommunications and Applied data Processing Technology, University of Abomey-Calavi, BENIN

Correspondence: Tahirou DJARA, Laboratory of Electronic Engineering, Telecommunications and Applied data Processing Technology, University of Abomey-Calavi, BENIN, Email csm.djara@gmail.com

Received: October 31, 2017 | **Published:** May 09, 2018

Introduction

Since twenty years, the computer modelling of emotion is a theme increasingly recognized, particularly in the field of human-machine interaction.² The term “emotion” is relatively difficult to define from a scientific point of view. Indeed, the phenomenon of emotion is based at the same time on physical, physiological, mental and behavioural considerations. Thus, many areas such as affective computing and image processing are interested in human emotional dimensions. The growing maturity of the field of emotion recognition is creating new needs in terms of engineering. After a replication phase, during which numerous works have been proposed with recognition systems,³ we are gradually entering an empiricism phase,⁴ where models for the design are developed.³ Most designed systems allow passive recognition of emotions. To define emotion, we base ourselves on Scherer’s theory.⁵ An emotion is characterized by a highly synchronized expression: the whole body (face, limbs, and physiological reactions) reacts in unison and the human emotional expression is clearly multimodal. Indeed, a large number of studies have been carried out in order to define the relationship between emotion and physiological signal. These have allowed highlighting a significant correlation between this type of signal and certain emotional states.

Architecture of emotion recognition systems

The analysis of existing emotion recognition systems reveals decomposition into three levels each fulfilling a specific function. At the Capture level, the information is captured from the real world and in particular from the user through devices (camera, microphone, etc.). This information is then analyzed in the Analysis level, where emotionally relevant characteristics are extracted from the captured data. Finally, the extracted characteristics are interpreted to obtain an emotion.

Physiological activities and emotional induction

There are several physiological activities that can allow the

determination of emotion beyond the face, voice and body gestures:

Electro-myographic activity (EMG)

In particular, EMG makes it possible to measure the electrical activity of the muscles via electrodes placed on the face. Several studies have shown that EMG signals provide an objective measure for the emotion recognition.⁶

Heart rate (ECG)

It defines the number of heartbeats (heartbeats) per unit of time, usually in beats per minute (BPM). It is generally associated with activation of the autonomic nervous system (ANS)⁷ itself related to the emotion treatment.⁸ Thus, the heart rate variation can be associated with different emotions.

Skin temperature (SKT)

The body controls the internal temperature by balancing heat production and heat loss. Heat production is achieved through muscle contraction, metabolic activity and vasoconstriction of the skin blood vessels. The activation of this indicator varies according to the emotion considered and the subjects, which induces a form of complex response making it possible to distinguish different emotions.

Respiratory frequency (FR)

Central nervous system

The central nervous system (CNS) is composed of the brain, cerebellum, brain stem and spinal cord. The brain activities of the CNS play a prominent role in the emotion recognition.

Acquisition and processing of physiological signals

The physiological activity is characterized by the calculation of several characteristics from the recorded signals. Once the acquisition of physiological signals is done, it is important to define

a methodology that allows the acquired signals to be translated into a specific emotion. Several works in the emotion recognition have been carried out using these methods⁶ based on statistical values as well as the construction of relevant indicator vectors. Each physiological signal (EEG, ECG, etc.) is designated by the discrete variable X . X_n represents the value of the n th sample of the raw signal, where $n = 1, \dots, N$ and N is the total number of samples corresponding to T seconds of signal recording. Assuming that each measured signal is generated by a Gaussian process, with independent samples and identically distributed. The two physiological functions that can be used to characterize a raw physiological signal are the mean and the standard deviation (Eq.1 and Eq.2):

$$\mu_x = \frac{1}{T} \sum X(t) = \bar{X}(t) \quad (1)$$

$$\sigma_x = \sqrt{\frac{1}{T} \sum_{t=1}^T (X(t) - \mu_x)^2} \quad (2)$$

In order to evaluate the trend of an X -signal on a test, the derived average (Eq.3), the normalized first derivative (Eq.4), the second derivative (Eq.5) and the normalized second derivative of the signal (Eq.6) can also be calculated:

$$\delta_x = \frac{1}{T-1} \sum_{t=1}^{T-1} |X(t+1) - X(t)| \quad (3)$$

$$\bar{\delta}_x = \frac{1}{T-1} \sum_{t=1}^{T-1} |\bar{X}(t+1) - \bar{X}(t)| = \frac{\delta_x}{\sigma_x} \quad (4)$$

$$\gamma_x = \frac{1}{T-2} \sum_{t=1}^{T-2} |X(t+2) - X(t)| \quad (5)$$

$$\bar{\gamma}_x = \frac{1}{T-2} \sum_{t=1}^{T-2} |\bar{X}(t+2) - \bar{X}(t)| = \frac{\gamma_x}{\sigma_x} \quad (6)$$

Finally, the maximum (Eq.7) and minimum (Eq.8) of a signal can also provide relevant information.

$$\min_x = \min_x x(n) \quad (7)$$

$$\max_x = \max_x x(n) \quad (8)$$

These characteristics are very general and can be applied to a wide range of physiological signals (EEG, EMG, ECG, RED, etc.). Using these characteristics, we obtain a characteristic vector Y of 8 values for each sample. This vector can cover and expand a statistical series typically measured in the literature.⁶

$$X = \left[\mu_x \ \sigma_x \ \delta_x \ \bar{\delta}_x \ \gamma_x \ \bar{\gamma}_x \ \min_x \ \max_x \right]$$

Conclusion

After the extraction of the desired characteristics, it is necessary to identify the corresponding emotion. This treatment is usually done by a classifier. A classifier is a system that groups similar data into a single class. It is able to make the correspondence between the calculated

parameters and the emotions. There are several classification methods. These include: Support Vector Machine (SVM), Bayesian Naïve Classification, Logistic Regression. It is therefore necessary to evaluate the performance of a classifier. The criterion used to evaluate this performance is the rate of good classification expressed by the following formula:

$$tbc = \frac{\text{Number of item identified correctly}}{\text{total element Number}}$$

Acknowledgements

None.

Conflict of interest

The author declares there is no conflict of interest.

References

- Allain P, Francois E. Epreuve de reconnaissance d'émotions en situation dynamique. Technical report, Laboratoire de Psychologie, University of Angers, France, 2010. p. 56–105.
- Picard RW. Affective Computing. Massachusetts Institute of Technology, USA, 1997. p. 1–167.
- Jaimes A, Sebe N. Multimodal human-computer interaction: A survey. *Comput Vis Image Underst.* 2007;108(1-2):116–134.
- Clay A, Couture N, Nigay L. Engineering affective computing: a unifying software architecture. In Proceedings of the 3rd IEEE International Conference on Affective Computing and Intelligent Interaction and Workshops; 2009. p. 1–6.
- Pantic M, Sebe N, Cohn JF, et al. Affective multimodal human-computer interaction. In MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia. New York, USA: ACM; 2005. p. 669–676.
- Hamza Hamdi. Plate-forme multimodale pour la reconnaissance d'émotions via l'analyse de signaux physiologiques: Application à la simulation d'entretiens d'embauche. *HAL.* 2014.
- Brownley KA, Hurwitz BE, Schneiderman N. Cardiovascular Psychophysiology, Cambridge University Press. 2nd ed. APA PsycNET. 2000. p. 224–264.
- Yoo SK, Lee CK, Park JY, et al. Neural network based emotion estimation using heart rate variability and skin resistance. *Advances in Natural Computation.* 2005;3612:818–824.
- Gaudelus B. Photographies TREF classées par émotion et par ordre d'intensité d'expression croissant; 2015.