

The emotional state through visual expression, auditory expression and physiological representation

Benyoussef Abdellaoui^{1*}, *Aniss Moumen*¹, *Younes El Bouzekri El Idrissi*¹, and *Ahmed Remaida*¹

¹ENSA Kenitra, Engineering Science Laboratory, Ibn Tofaïl University, Kenitra, Morocco

Abstract. As emotional content reflects human behaviour, automatic emotion recognition is a topic of growing interest. During the communication of an emotional message, the use of physiological signals and facial expressions gives several advantages that can be expected to understand a person's personality and psychopathology better and determine human communication and human-machine interaction. In this article, we will present some notions about identifying the emotional state through visual expression, auditory expression and physiological representation, and the techniques used to measure emotions.

1. Emotional state

This work was made a literature review concerning emotion recognition methods; we have analyzed in-depth this work [1], our choice was arbitrary, but we judged that it covers and treats our problem well.

The emotional state of a person can be known through his behaviour, the visual expression as a facial and auditory expression estimated by the modulation of the vocal signal, the physiological representation by the analysis of the activities of the autonomic nervous system (ANS). The emotional state helps in decision making, assists creativity and manages human cognition and human-machine communication. The data captured is usually very much related to the measurement tool, such as the quality of the video, the quality of the lighting, the pose and size of the face on the video and the noises in a voice recording [1].

Emotions play an implicit role in the communication process compared to the explicit message spread by the linguistic context. As emotional content incarnates human behaviour, the automatic recognition of emotions is a subject that inspires a growing interest. The element to be recognized is complicated and subtle, producing diverse manifestations that depend on many factors (social and cultural context, speaker's personality, etc.) [1].

Emotion measurement is, in general, a very complex topic because it needs to involve several classical measurement methods: facial expression verification, speech verification and physiological signal analysis. This paper aims to merge some emotion recognition methods. We will discuss here the literature review of physiological signals and facial expressions. Other techniques can be used but are not considered in this work. We also quote

the reason for this fusion and some constraints related to each method, and some open questions that can be answered. Each technique and some open questions remain to be researched.

The importance of facial expression comes from the fact that the emotional message is communicated through this channel, according to Mehrabian, and the rest of message is communicated through the linguistic channel and paralanguage. Thus, facial expressions or voices can be camouflaged but in general physiological tests are difficult to manipulate. The envisaged solution is then to merge the processing of physiological sensors and facial data.

1.1. Notions on emotions

According to the Larousse dictionary, the word emotion comes from the old French word “émouvoir”, which designates the transient connotative response of great force, usually caused by excitation from the environment, that emotion is its channel of reaction with its environment. We distinguish several classifications of emotions. We quote here the six so-called basic emotions other than the normal state: joy, sadness, anger, fear, disgust and surprise [2], which are observed in people who are blind from birth, which proves that they are genetically transmitted without being mimic gestures, the following table represents a description of these emotions [2].

Table 1: The description of the basic emotions.

| The emotion | Description of the emotion |
|--------------------|--|
| The joy [3] | The mouth is open to show the smile, the edges of the eyes narrow, the cheekbones and the corners of the lips go up. |
| Sadness [4] | The mouth closed, the lips turned downwards, the eyelids are lowered, and the look is lost. |
| The surprise [5] | The mouth is round, the eyes are wide open, and the eyebrows are raised. |
| The fear [6] | Eyebrows are up and close together, upper eyelids are lifted, the lower eyelids are tight, and lips have drawn towards the ears. |
| The anger [7] | Eyebrows are down and together, lips are pursed, and the look is frank and snarling. |
| The disgust [8] | The nose is turned up, and the lower lips are down. |
| Contempt [9] | A single corner of the lips points to one side. |

Secondary emotions (sympathy, gratitude, admiration, contempt, guilt, shame, etc.) [10] result from combining primary emotions to produce a new emotion. Social emotions [11], such as infamy, unfavourable judgments of one's behaviour, strong reserve and withdrawal, fear or excessive anxiety, infamy, and social fears in general, are hurtful emotions that have the role of guiding us towards better protection or better scenarios, while other feelings like admiration, camaraderie, desire, sympathy, cooperation, living in society, attachment, calming sensations and relative to beauty, moral feelings, are enormously motivating because they develop essential sensations of complacency or relief.

1.2. Theoretical models of emotions

The channel of reaction of the person with his environment is actuated by his emotions spread according to three cognitive aspects. They are related to reading the meaning of the stimuli, the physiological one that leads the reaction to be carried out, and the expressive one considered by the translation by what message to transmit to the environment. The physiological theoretical framework is founded by William James (January 11, 1842 - August 26, 1910) and Carl Lange (December 4, 1834 - May 29, 1900) and succeeded by several types of research that found that the emotional reactions are in solid link with the psychological experience and the emotional state. Another theoretical framework presided over by Charles Darwin (February 12, 1809 - April 19, 1882) is that the adult emotional manifestation expresses the continuation of complex behavioural systems derived from other animal species and that the face reflects the emotional state. Another organ that is also very important is the brain that participates in the physiology of emotions.

2. Representation of emotions

The representation of emotions poses several problems since the modelling is in simple formalism and must be consistent with psychology results. Based on the work in psychology, the vision differs between a categorical approach [12] and the other a multidimensional construction.

2.1. Categorical approach

The main interest of the categorical approach is that emotions are distinct classes from each other, and researchers disagree on a single classification, which is why there are several lists of basic emotions (six for Ekman et al. and up to ten for Izard and Carroll E). This approach proposes linking to each emotion the associated facial expressions and does not consider the links between these states.

2.2. Dimensional approach

The dimensional approach is based on the continuity of emotions on several axes and dimensions and proposes that the passage from one state to another is not discontinuous [13]. Emotion comprises three components: the behavioural component, the physiological, and cognitive/subjective elements. The behavioural components are the facial expressions and the prosody. We cite here the example of William James concerning the confrontation of a bear, we have fear (cognition), our body manifests the desire to flee, this action is accompanied by open eyes and feet that begin to run (behaviour) this causes the increase of the heart rate and breathing that increases (physiology) [14].

3. The recognition of feelings

The works that exist in the literature interested in recognizing emotions use the following modalities: facial expressions, physiological tests, or speech to deduce the emotional state in human beings. We note that there are other modalities, and they are not treated in this work; we do not find enough works trying to explain the emotion combining two or more modalities, that is why we have tried to treat this aspect and to clarify the interests, the constraints and the perspectives.

3.1. Facial expressions

A facial expression results from a deformation of the traits of the eyes, mouth, nose and eyebrows that lead to a particular emotion. To measure the emotion, we find either facial expressions, physiological measurements, voice or gestures.

The measurement of emotion from facial expressions based on a video is essentially done in three stages: the extraction of facial features in the first frame, tracking its features. The final stage is the coding and classification of expressions for giving the emotion. The images to be processed in general must be well centred on the face of the person and in a well-lighted environment. The detection of the face is done by several techniques such as HAAR descriptors [15] and then using classifiers such as AdaBoost [16] to determine if the bounding box represents a face or not.

The extraction of the facial characteristics is equivalent to determining the mouth, the eyes and the nose; the transition max calculates this extraction in the upper part of the face, which determines the line that contains the eyes and the same for the mouth and finally the mediating segment containing the nose. The location of the characteristic facial points is done by an anthropometric model that allows the measurement of the dimensional particularities of a technical man developed by Adolphe Quetelet. (February 22 1796 - February 17 1874). After knowing the points of interest, the tracking of the points is based on the variation of the locations of the pixels of the image, this concept of optical flow was studied by James Gibson (1904-1979). In the field of computer vision, the concept was introduced by Bruce D. Lucas and Takeo Kanade. In the stage of facial expression recognition, the purpose varies between recognizing discrete states of emotion or measuring the amplitude of facial action units from the face and applying the appropriate emotion classification algorithm.

3.2. Physiological measurements

The emotions of humans are related to physiological changes that manifest themselves in several organs such as the heart, skin, lungs and other limbs. The sensation of anger is usually related to an increase in body temperature, so sensors that determine this constant lead to deduce the emotional state. In practice, by using packs that can provide the capture and measurement of different information such as respiratory rate (VR) and heart rate (BVP), electromyography (EMG), electroencephalography (EEG), electrocardiography (ECG), skin conductance (SKC)... The sensors are connected to the person either by wired or wireless links transmitting the data to an encoder which takes care of the processing. We note that the use of these sensors is more or less flexible. In the case of ECG measurement, we need to undress our subject. In general, these measurements are costly in resources, and sometimes the data contains noise related to the measurement tool, which imposes the use of ways to correct these constraints. The data processing allows to extract characteristic information from each physiological signal and then use a classifier to match the measurement with the

state already characterized by some parameters and finally validate the model on a reliable database containing the studied emotions. In summary, the combination of several physiological signals determines the emotional state with great accuracy.

3.3. Fusion of physiological signals and facial expressions

The use of facial expression or physiological signals alone does not give a satisfactory result, so combining these two or more methods is necessary. We find works on the fusion between two or more modes: speech, facial expressions, physiological signals, and text input. For the use of facial expressions, we cite here the work of Michael J. Black & Yaser Yacoob on the recognition of facial expressions in image sequences using local parameterized models of image motion [17], the work of I.A. Essa and A.P. Pentland on coding, analysis, interpretation, recognition of facial expressions [18], the work of Kenji MASE entitled recognition of facial expression from optical flow [19] and the work of Ying-li Tian et al. entitled recognition of lower face action units for facial expression analysis [20]. For the use of physiological signals, we cite here the work of K. H. Kim et al. entitled Emotion Recognition System Using Short-Term Monitoring of Physiological Signals [21], the work of Pramila Rani et al. entitled Online Stress Detection Using Psychophysiological Signals for Implicit Human-Robot Cooperation [22] and the work of R.W. Picard et al. entitled Towards the emotional intelligence of machines: analysis of the physiological affective state [23]. For the use of speech, we cite the paper written by Tin Lay Nwe et al. entitled classification of emotions based on speech [24] and the work of F. Dellaert et al. entitled recognizing emotions in speech [25]. For the fusion of several modalities, we cite the work of L.S. Chen et al. entitled multimodal recognition of human emotions/expressions [26] and the work of L.C. De Silva et al. entitled bimodal recognition of emotions [27].

4. Results and discussion

The variation of the distances of the characteristic features of the face is a way that allows the determination of emotions from facial expressions; the information is extracted from the variation of the shape of the muscles of the face based on the fixed points and dynamic points, the detection of the face is the first step of this method, Many algorithms are introduced to improve the accuracy of the centring of the face such as the technique of Shi-Tomasi for example, the next phase is to use classification methods to separate the states from each other, the choice of the algorithm is directly related to the data and of course to the expertise of the person who conducts the research. The validation of the model is desirable on approved databases; this approach is an external vision of the person who participates in this experiment, so another approach that is required is the inner vision, applying to the participant psychological tests such as signals (VR) and (BVP), (EMG), (EEG), (ECG), (SKC) ... physiological tests are very interesting and cannot be manipulated because they are related to the activities of the nervous system. However, the sensors are generally parasitic such as a sensor on the head. Some questions remain open such as the inclusion of images, the excellent resolution, the algorithms to use and the calculation time for real-time applications; the test and validation databases are sometimes very typical which do not take into consideration the diversity on our planet in terms of the human race, age, occupation, emotional diversity... For physiological signals, the use of sensors with wires can be annoying to the user; smaller sensors without wires or contacts are strongly recommended.

5. Conclusion

Human emotion is a physiological and behavioural component that complicates this study because of this diversity and the validation of the results. The privilege of physiological signals is due to the excellent robustness despite the simplicity of facial expressions. However, for better detection of emotions, the fusion of the two methods can lead to better results. An important thing to mention is that emotion is the human interface of its communication with the environment, proving its importance. Other modalities can be used, such as speech and gestures, to enrich the recognition of emotions.

References

1. F. Khalfi, 'Reconnaissance automatique des émotions par données multimodales: expressions faciales et des signaux physiologiques', p. 161.
2. 'Comment reconnaître les expressions faciales des émotions', *Bernard Guevorts*, May 09, 2018. <https://www.bernard-guevorts.com/article/comment-reconnaitre-les-expressions-faciales-des-emotions/> (accessed May 21, 2021).
3. 'Les émotions en détail: la joie', *Bernard Guevorts*, May 16, 2018. <https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-joie/> (accessed May 21, 2021).
4. 'Les émotions en détail: la tristesse'. https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-tristesse/#.YKb9Vd0o_IU (accessed May 21, 2021).
5. 'Les émotions en détail: la surprise'. https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-surprise/#.YKb9d90o_IU (accessed May 21, 2021).
6. 'Les émotions en détail: la peur'. https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-peur/#.YKb9l90o_IU (accessed May 21, 2021).
7. 'Les émotions en détail: la colère'. https://www.bernard-guevorts.com/article/les-emotions-en-detail-la-colere/#.YKb9tt0o_IU (accessed May 21, 2021).
8. 'Les émotions en détail: le dégoût'. <https://www.bernard-guevorts.com/article/les-emotions-en-detail-le-degout/> (accessed May 21, 2021).
9. 'Les émotions en détail: le mépris'. <https://www.bernard-guevorts.com/article/les-emotions-en-detail-le-mepris/> (accessed May 21, 2021).
10. E. Duplâa, 'Les émotions: outils et corps social pour l'apprentissage en ligne', *stice*, vol. 14, no. 1, pp. 289–308, 2007, doi: 10.3406/stice.2007.961.
11. R. Coenen, 'Les émotions sociales, une clé pour la délinquance?', *Journal du droit des jeunes*, vol. N° 271, no. 1, pp. 25–27, 2008.
12. J. Redet and M. Vian, 'Comparaison d'outils adaptés au terrain pour l'évaluation du ressenti à l'usage des IHM', p. 4.
13. M. Courgeon, C. Clavel, and J.-C. Martin, 'Approche Comparative des Modèles Informatiques des Émotions pour l'Animation Faciale en Situation d'Interaction', p. 8.
14. S. Gil, 'Comment étudier les émotions en laboratoire?', no. 4, p. 10, 2009.
15. A. A. Mekonnen, F. Lerasle, A. Herbulot, and C. Briand, 'Détection de personnes par apprentissage de descripteurs hétérogènes sous des considérations CPU', p. 8.
16. Y. Freund and R. E. Schapire, 'Experiments with a New Boosting Algorithm', p. 16.
17. M. J. Black, 'Recognizing Facial Expressions in Image Sequences Using Local Parameterized Models of Image Motion', p. 26, 1997.

18. I. A. Essa and A. P. Pentland, 'Coding, analysis, interpretation, and recognition of facial expressions', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 757–763, Jul. 1997, DOI: 10.1109/34.598232.
19. K. Mase, 'Recognition of Facial Expression from Optical Flow', *IEICE TRANSACTIONS on Information and Systems*, vol. E74-D, no. 10, pp. 3474–3483, Oct. 1991.
20. Y. Tian, T. Kanade, and J. F. Cohn, 'Recognizing lower face action units for facial expression analysis', in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, Mar. 2000, pp. 484–490. DOI: 10.1109/AFGR.2000.840678.
21. 'Emotion recognition system using short-term monitoring of physiological signals | SpringerLink'. <https://link.springer.com/article/10.1007/BF02344719> (accessed June 08, 2021).
22. 'Online stress detection using psychophysiological signals for implicit human-robot cooperation - ProQuest'. <https://www.proquest.com/openview/cdfd1510ad24d462dcbda33a6a98cd9e/1?pq-origsite=gscholar&cbl=37292> (accessed June 08, 2021).
23. R. W. Picard, E. Vyzas, and J. Healey, 'Toward machine emotional intelligence: analysis of affective physiological state', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, Oct. 2001, DOI: 10.1109/34.954607.
24. T. L. Nwe, F. S. Wei, and L. C. De Silva, 'Speech based emotion classification', in *Proceedings of IEEE Region 10 International Conference on Electrical and Electronic Technology. TENCON 2001 (Cat. No.01CH37239)*, Aug. 2001, vol. 1, pp. 297–301 vol.1. DOI: 10.1109/TENCON.2001.949600.
25. F. Dellaert, T. Polzin, and A. Waibel, 'Recognizing emotion in speech', in *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, Oct. 1996, vol. 3, pp. 1970–1973 vol.3. DOI: 10.1109/ICSLP.1996.608022.
26. L. S. Chen, T. S. Huang, T. Miyasato, and R. Nakatsu, 'Multimodal human emotion/expression recognition', in *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, Apr. 1998, pp. 366–371. DOI: 10.1109/AFGR.1998.670976.
27. L. C. De Silva and P. C. Ng, 'Bimodal emotion recognition', in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, Mar. 2000, pp. 332–335. DOI: 10.1109/AFGR.2000.840655.