AN ADVANCED TECHNIQUE FOR FACIAL EMOTION DETECTION USING ARTIFICIAL INTELLIGENCE

Mr. V. Vigneswara Reddy[1]
3rd Year Student,
Department of Computer Science,
SV U CM & CS, Tirupati.

Prof. G.Anjan Babu[2],
Professor,
Department of Computer Science,
SV U CM & CS,, Tirupati.

Abstract: Human psychological stress and human emotion are very much interconnected, mainly in computational psychology study the relationship between stress and emotions is the key to understanding the underlying effects of surrounding and inner challenges of a human being. In this research we accurately predicted emotions individually and as a mixture from coded facial expressions. We also explored the relationship between visually perceptible psychological stress and the seven basic human emotions (Anger, Disgust, Contempt, Fear, Sad, Happy and Surprise) as depicted by facial expressions, both individually and in combination and found that visually perceptible psychological stress varies as a logarithmic function of emotion percentages as observed on the face of subjects.Substantial research has been done in the field of stress detection and analysis but the number of researches in the field of computerized or automated stress detection from the face is few. Our future goal is to develop a system to detect and monitor stress levels of employees in workplaces in real-time. All researches done so far on stress detection are not very practical for use as a real world application for stress detection. This is due to the fact that it is either intrusive and if not intrusive are cumbersome due to attachment to gadgets to the hands or other parts of the body which might restrict free movement of the subject at workplaces. Another method exists which measures stress levels of subjects, which is the self-reporting questionnaire method. This method is very much susceptible to reporting bias. Also, research in this domain suggests that both stress and emotionsaffect facial muscle movements. Accordingly a relationship seems to exist between stress and emotions as depicted on the face. So we proposed an emotion based method to detect and evaluate stress levels from facial expressions of a subject. This method eradicates the cumbersomeness of attaching electrodes, collecting biological samples or the bias in questionnaire methods and is very well applicable to real scenarios. The stress we evaluate is not a medical evaluation of psychological stress but just an indicator that can be used to recommend medical attention or consultation.Also, there has been plenty of research that attempts at unraveling emotion information of human subjects. There are a range of approaches to deal with emotion detection. Facial expressions, speech, bio-medical and psychophysiological methods are the leading approaches used for this field of research. Again, in this case as well, the facial expression method is the most suitable for our future goal of a stress monitoring system at workplaces as it does not require isolation as in speech method, neither does it require intrusive medical procedures as in bio-medical methods, nor is cumbersome with electrodes attached to the body as in non-intrusive psychophysiological approach. In this research facial expression is used as the input and using our proposed pre-trained Hidden Markov Model (HMM) [1] network emotional mixture of any particular facial expression is evaluated. Our findings include the following:

1. Accuracy of Emotion prediction is enhanced when we consider all facial muscle movements together rather than considering only the prominent muscle movements.
2. There is no gender difference in terms of emotion response among different genders but accuracy of overall emotion prediction is improved by gender segmentation during training and testing.
3. Visually Perceptible Psychological Stress can be quantitatively expressed as a function of seven basic emotions (Anger, Contempt, Disgust, Fear, Happy, Sad and Surprise).
4. When stress is considered as the response of occurring emotions on the face of a subject the relationship between stress and emotion is found to be logarithmic, this is in accordance to the famous Weber-Fechner law [2, 3] of stimuli.
5. Accuracy of Stress evaluation is enhanced when we consider all the insignificant emotions as well along with the lead emotion rather than only considering the lead emotion.

INTRODUCTION

The understanding of psychological stress and emotions are well related but has not been explicitly taken up for research yet. They are always dealt in researches as two different fields of study. In this research we quantified an indicator for visually perceptible human psychological stress from the face (hereafter mostly referred to as stress) in terms of emotions represented on the human face, which we term as facial expressions. Facial musculature is greatly varied among humans and to establish a universal or generalized method to evaluate facially observed stress would be impossible if we were to use only the muscle movement information. So to overcome this problem we need to express the muscle movement on the face to an intermediate form that would be more generalized before we could build a method to evaluate facially observed stress. According to the Facial Action Coding System developed by Ekman and Friesen [6, 7] seven basic emotions anger, disgust, contempt, fear, happy, sad and surprise are innate and universal to humans. These universal basic emotions can be used as an intermediate form which can then be used to develop a more general quantitative method of evaluating our intended stress indicator. Also, if we want to evaluate stress without the knowledge of the interplay of underlying emotions then any discussion on stress maybe incomplete. So in order to have an insight on the stress response of subjects we first need to develop a system which would identify the basic emotions depicted by a facial expression. Breaking it down further, to understand the emotion exhibited we needed to identify the muscle movements on the face.We did not build our own dataset of facial expressions and neither had we identified facial muscle movements from images. We used a dataset known as the Cohn Kanade Dataset (CK+) [4, 5] which readily gave us the facial muscle movement data as well as their depicted emotion for facial images according to the Facial Action Coding System developed by Ekman and Friesen in 1978 [6, 7]. We use these muscle movement data as inputs to our proposed model and arrive at emotion data as outputs. The manually obtained emotion labels provided in the CK+ dataset provided for the ground truth data which we used for training as well as testing of our model. Once the emotion detection phase was completed we conducted surveys that related stress and emotions. From these surveys and our emotion detection results we analyzed the relationship between emotions and visually perceptible psychological stress as reported by psychologists in the survey responses using regression analysis.

METHODOLOGY

Emotion assessment from facial expression is filled with complications and complexities. There are 64 identifiable facial muscle movements connected to emotion representation on the face according to Ekman and Friesen [6, 7]. Also, each muscle referred to as an Action Unit (AU) may exhibit varying degrees of intensity or deviation from the neutral position. A Neutral facial expression refers to a facial expression which does not exhibit any emotion at all and the position and orientation of any AU in this expression is referred to as the neutral position of that particular AU. These varying intensities are graded from level 1 to level 7 in terms of their increasing intensity. Level 1 represents the absence of any movement for an AU and Level 2 indicates that an AU is showing presence but the deflection is negligible and so on to Level 7 which indicates the maximum deflection of an AU from its neutral position. So to say the least for 64 AUs interplaying together to make up a facial expression there could be 647 = 4.398 x 1012 combinations that may represent an emotion or an emotion mixture.

This huge number of combination are not all valid or possible expressions on a human face due to interrelated muscle movements and the near impossibility of facial muscle control to move every AU individually. Nevertheless roughly 7000 combinations are known to be valid [44] and there are no rules that can directly relate these combinations to respective emotions. So a robust and learning capable model needs to be proposed to solve the problem.

We use Hidden Markov Models in a complex orientation to find the probabilities that facial expression represent a particular emotion for all emotions. These probabilities are then normalized to understand the proportions of the basic emotions that constitute the whole facial expression. We assume here that the normalized probabilities are representative of the degrees of mixed emotion as often done in the field of engineering.

After we have deciphered emotion information from the facial expression we need to evaluate stress levels and to that effect we conducted two surveys among psychology researchers and psychiatrists, one relates facial expression images to stress and the other relates emotion degrees to stress levels. After these surveys were conducted we use the results of the survey and the deciphered emotion information data to predict a regressional model that best describes the stress and emotion data. We propose five different linear and non-linear regressional models for prediction of stress levels from emotion degrees or percentages. We do this both for individual basic emotions and mixed emotions and choose the best model based on parameters like goodness of fit and root mean square errors. After the best model is chosen we present the model equations by determining the coefficients of unknown variables using regression analysis.

The above mentioned approach can finally be used to evaluate visually perceptible stress levels as observable from the face of a subject in terms of seven basic emotions and the emotions can be deciphered using facial muscle movement information. The stress evaluation done here would be representative of the way how psychology researchers and psychiatrists recognize stress from the face.

CONCLUSION

This dissertation discussed about two different models of emotion classification and a model for evaluation of facially observed stress. We found that stress visible on the face varies logarithmically with emotion percentages. From the perspective of stimulus and response this follows Weber Fechner's law of stimuli. This research also proposed emotion detection model with improved performance both for individual or lead emotion and emotion mixtures. The model for lead emotion detection is very fast without much loss of generality and accuracy. In the emotion mixture model gender segmentation was used to detect gender differences about emotion display on the face and our findings indicate no gender difference between males and females in emotion representation. Although we found no gender bias in emotion display but creating separate models for male and female groups have boosted accuracy.

The stress detection model is intended not only to establish the logarithmic relationship between stress and emotions but it is a yardstick to quantitatively assess visually perceptible psychological stress levels using the equations generated by regression analysis. Once we train our model with more and more data and do a survey with much wider demographics, we would be able to very accurately predict stress in real scenarios. Emotions as well as stress response vary immensely along socio-cultural lines, and lines of ethnicity. In this respect there is need to study the cross-cultural effects of emotions and stress. Right now the CK+ [4, 5] dataset is not demarcated into ethnicitiesor into socio-cultural groups so these studies were not possible. This is one of the areas of our future study interest.

We have mentioned in this dissertation that combination emotional ground truth data is not available at the moment, but studies regarding emotion mixtures and combinations are needed to better understand human perception, and humanized communication. Also, thinking in term of mixture of emotion model, we can visualize a system which recognizes all 7 emotions from the face recursively. To elucidate this lets think of a recursive function that call itself. In the first run it finds the lead emotion, in the next call it finds the next lead emotion or the emotion which is next to the lead emotion in prominence. As in a recursive function parameters do not change, only the context does, the function leaves out the most prominent emotion in each step and attempts to find another prominent emotion until all 7 emotions are detected. Thinking this way the generated emotion percentage values should be fairly accurate. Our model can be thought of in a similar way. As the HMMs were trained without any bias or incomplete AU information all of the emotion HMMs would be equally responsive in detecting their type of emotion. In other words the probabilities for an emotion that is calculated can be thought of as the probability of the emotion of being the lead emotion. Considering all the emotion blocks prying on the AU code to find the possibility of corresponding emotion to be prominent, it is like a multithreaded parallel recursion. So, to conclude a database built with this model in combination with manual or automated FACS coding of facial image will successfully be authentic.Another area that we want to extend our research to is the real-time stress monitoring of personnel in organizations. This would be achieved in coordination of an automated FACS encoding program, historical and real-time stress analysis and the consideration of cardio and body temperature sensing. A cardio mouse is now readily available in the market and a remote body temperature sensor, which is pretty expensive at the moment. With these systems in place a stress monitoring system with alerting facility can be developed which would alert senior management or other responsible members of the organization about the stress conditions of an employee once it crosses a preset threshold. This would to some extent help employers to avoid stress related mishaps of employees. Also, patients suffering from mental illness if monitored through this system will give psychiatrists and medical professionals a great deal more information and insight about patients' stress traits outside the doctor's chamber, which is currently inaccessible.

REFERENCES

[1] Rabiner L.R., 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition, Proceedings of the IEEE, vol. 77 (2), pp. 257–286.

[2] Glezer V. D., 2007. The Meaning of the Weber-Fechner Law and Description of Scenes in Terms of Neural Networks, Human Physiology, Volume 33(3), pp. 257-266.

[3] Glezer V. D., 2011. The Meaning of the Weber-Fechner Law: IV. The Psychometric Curve and Interhemispheric and Intrahemispheric Interactions, Human Physiology, Volume 37(1), pp. 57-65.

[4] Kanade T., Cohn J., &Tian Y.L., 2000. Comprehensive database for facial expression analysis, Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00), Grenoble, France, pp. 46-53

[5] Lucey P., Cohn J.F., Kanade T., Saragih J., Ambadar Z. & Matthews I., 2010. The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression, Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, pp. 94-101.

[6] Ekman P. & Friesen W.V., 1978. Facial Action Coding System: Investigator's Guide, Consulting Psychologists Press, Palo Alto, CA.

[7] Ekman P., Friesen W.V, & Hager J.C., 2002. The New Facial Action Coding System (FACS), Research Nexus division of Network Information Research Corporation.

[8] Dellaert F., Polzin T., &Waibel A., 1996 Recognizing emotion in speech. In Spoken Language, 1996.ICSLP 96. Proceedings, vol. 3, pp. 1970-1973.

[9] Nogueiras A., Moreno A., Bonafonte A., &Mariño J. B., 2001. Speech emotion recognition using hidden markov models.In Proc. Eurospeech, vol. 1, pp. 2679-2682.

156

[10] Bhatti M. W., Wang Y., & Guan L., 2004. A neural network approach for human emotion recognition in speech, In Circuits and Systems, ISCAS'04, vol. 2, pp. II-181.

[11] Wilhelm F.H., Pfaltz M.C. & Grossman P., 2006, Continuous electronic data capture of physiology, behavior and experience in real life: towards ecological momentary assessment of emotion. Interacting With Computers, vol. 18 issue. 2, pp. 171-186.
[12] Murugappan M., Rizon M., Nagarajan R., Yaacob, S., Hazry, D., &Zunaidi, I., 2008. Time-frequency analysis of EEG signals for human emotion detection. In 4th Kuala Lumpur International Conference on Biomedical Engineering, pp. 262-265.
[13] Khalili Z., &Moradi M. H., 2008. Emotion detection using brain and peripheral signals.In Biomedical Engineering Conference, 2008, CIBEC 2008, Cairo International, pp. 1-4.
[14] Darwin C., 1898. The Expression of the Emotion of man and animals, D. Appleton & Co., New York.
[15] Mase K., 1991, Recognition of facial expression from optical flow, IEICE Transactions, E74 (10), pp. 3474–3483.
[16] Otsuka T. and Ohya J., 1996, Recognition of Facial Expressions Using HMM with Continuous Output Probabilities, Proceedings International Workshop Robot and Human Comm., pp. 323-328.

Authors Profile

VENNAPUSA VIGNESWARAREDDY, received Bachelor of Computer Science degree from Yogi Vemana University, Kadapa in the year of 2013-2016. Pursuing Master of Computer Applications from Sri Venkateswara University, Tirupati in the year of 2016-2019. Research interest in the field of Computer Science in the area of Artificial intelligence,Machine learning,Big Data,Network Security ,Networking and Software Engineering.