Impact Factor (SJIF): 5.301

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 5, Issue 3, March-2018

Automatic stress recognition using gestures and speech modulation

Mrs.M.Theresa [*1], A.Carolin [*2], M.Divya [*3], P.Pavithra [*4]

Department of computer science, Panimalar Institute of technology

Abstract* - *This paper investigates how speech and gestures convey stress, and how they can be used for automatic stress recognition. Stress can be conveyed by the intended semantic message (e.g. spoken words for speech, symbolic meaning for gestures), and stress conveyed by the modulation of either speech and gestures (e.g. intonation for speech, speed and rhythm for gestures). The goal is having a surveillance system that would notify when the stress level is high and extra assistance is needed. The solution is that speech modulation is the best performing intermediate level variable for automatic stress prediction. Using gestures increases the performance and is mostly beneficial when speech is lacking.

Keywords- Evaluation, non-contact sensing, sleep monitoring, cost-effective , polysomnography(PSG) ballistocardiogram, photoplethysmography.

1.INTRODUCTION

While automatic detection of unwanted behaviour is desirable and many researchers have delved into it, there are still many unsolved problems that prevent intelligent surveillance systems to be installed to assist human operators. One of the main challenges is the complexity of human behaviour and the large variability of manifestations which should be taken into consideration. Emotions and stress play an important role in the development of unwanted behaviour. In it is distinguished between instrumental and affective aggression. Instrumental aggression is goal directed and planned, e.g. Pick-pocketing. Affective aggression results from strong emotional feelings, like anger, fear and frustration. Furthermore, there is a link between stress and aggression. Detecting negative emotions and stress at an early stage can help preventing aggression.

However, an early stage is characterized by more subtle behaviour compared to violence, which increases the difficulty of automatic detection. Stress is a phenomenon that causes many changes in the human body and in the way in which people interact. It is a psychological state formed as response to a perceived threat, task demand or other stressors, and is accompanied by specific emotions like frustration, fear, anger and anxiety. For a thorough overview of stress we refer the reader to. Since the final application of this work is in the surveillance domain, we are interested in the stress observed in the overall scene, and not per person. It is a psychological state formed as response to a perceived threat, task demand or other stressors, and is accompanied by specific emotions like frustration, fear, anger and anxiety. For a thorough overview of stress we refer the reader to. Since the final application of this work is in the surveillance domain, we are interested in the stress observed in the overall scene, and not per person.

II.LITERATURE SURVEY

A. Title: 3D Model-Based Continuous Emotion Recognition
Author Hui Chen

Abstract The design and characteristics of a real-time 3D model-based method that continuously recognizes dimensional emotions from facial expressions in natural communications. In our method, 3D facial models are restored from 2D images, which provide crucial clues for the enhancement of robustness to overcome large changes including out-of- plane head rotations, fast head motions and partial facial occlusions. To accurately recognize the emotion, a novel random forest-based algorithm which simultaneously integrates two regressions for 3D facial tracking and continuous emotion estimation is constructed. Moreover, via the reconstructed 3D facial model, temporal information and user-independent emotion presentations are also taken into account through our image fusion process. The experimental results show that our algorithm can achieve state-of- the-art result with higher Pearson's correlation coefficient of continuous emotion recognition in real time

B.Title: Emotion Representation, Analysis and Synthesis in Continuous Space
Author Hatice Gunes[1], Bj¨orn Schuller[2], Maja Pantic[3] and Roddy Cowie[4]

Abstract The design and characteristics of the affective computing research field, modelling, analysing, interpreting and responding to naturalistic human affective behaviour still remains as a challenge for automated systems as emotions are complex constructs with fuzzy boundaries and with substantial individual variations in expression and experience. Thus, a small number of discrete categories (e.g., happiness and sadness) may not reflect the subtlety and complexity of the affective states conveyed by such rich sources of information. Therefore, affective and behavioural computing researchers have recently invested increased effort in exploring how to best model, analyze, interpret and respond to the subtlety, complexity and continuity (represented along a continuum e.g., from -1 to +1) of affective behaviour in terms of latent dimensions (e.g., arousal, power and valence) and appraisals. Accordingly, this paper aims to present the current state of the art and the new challenges in automatic, dimensional and continuous analysis and synthesis of human emotional behaviour in an interdisciplinary perspective

C.Title: Affective Body Expression Perception and Recognition
Author: Andrea Klein smith[1] and Nadia Bianchi-Birdhouse[2]

Abstract The design and characteristics of decreasing cost of whole-body sensing technology and its increasing reliability, there is an increasing interest in, and understanding of, the role played by body expressions as a powerful affective communication channel. The aim of this survey is to review the literature on affective body expression perception and recognition. One issue is whether there are universal aspects to affect expression perception and recognition models or if they are affected by human factors such as culture. Next we discuss the difference between form and movement information as studies have shown that they are governed by separate pathways in the brain. We also review psychological studies that have investigated bodily configurations to evaluate if specific features can be identified that contribute to the recognition of specific affective states. The survey then turns to automatic affect recognition systems using body expressions as at least one input modality. The survey ends by raising open questions on data collecting, labeling, modeling, and setting benchmarks for comparing automatic recognition systems.

D.Title: A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions
Author: Zhihong Zeng, Member

Abstract The design and characteristics of Automated analysis of human affective behavior has attracted increasing attention from researchers in psychology, computer science, linguistics, neuroscience, and related disciplines. However, the existing methods typically handle only deliberately displayed and exaggerated expressions of prototypical emotions, despite the fact that deliberate behavior differs in visual appearance, audio profile, and timing from spontaneously occurring behavior. To address this problem, efforts to develop algorithms that can process naturally occurring human affective behavior have recently emerged. Moreover, an increasing number of efforts are reported toward multimodal fusion for human affect analysis, including audiovisual fusion, linguistic and paralinguistic fusion, and multicue visual fusion based on facial expressions, head movements, and body gestures.

This paper introduces and surveys these recent advances. We first discuss human emotion perception from a psychological perspective. Next, we examine available approaches for solving the problem of machine understanding of human affective behavior and discuss important issues like the collection and availability of training and test data. We finally outline some of the scientific and engineering challenges to advancing human affect sensing technology.

E.Title: Modeling User Affect and Sentiment in Intelligent User Interfaces
Author: Bj¨ornW. Schuller

Abstract The design and characteristics of the computer-based automatic analysis of human sentiment, and affect are broadly expected to play a major role that will likely make 'that difference' in future Intelligent User Interfaces, as they bear the promise to lend interactive systems emotional intelligence. Such comprise intelligent digital games, e. g., for empowerment and inclusion, tutoring systems, information systems or virtual companions, e. g., in the car to name but a few. This tutorial aims to give a good introduction into the related fields of user Sentiment Analysis and user Affect Modeling. Its intention is to show the general technology, and its current reliability, the ways for technical integration and efficient embedding of solutions in a user interface context, and the latest trends in this young and ever emerging field. Emphasis is laid on highlighting the range of toolkits available at this moment with the aim of empowering one to immediately craft own solutions. This description contains the general motivation, goals, objectives, and topics

F.Title: Multi model aggression detection in train
Author: Z.Yang
Abstract The design and characteristics of work addressing the challenges of video analysis for automatic detection of aggression in a train. Using data from surveillance cameras, the system assists human operators in their work. It is unobtrusive and respects the privacy of passengers. We used existing algorithms to recognize and classify human behavior. While evaluating the algorithms we paid special attention to their ability to cope with environment specific issues, such as varying lighting conditions and (self)occlusions. A passenger behavior model was developed based on many hours of observing and studying professional operators as they analyze and respond to surveillance data. Experiments were conducted in a real train to evaluate the detection system

G. Title: Intergration of gestures and speech in human -robot interaction
Author: Raveesh Meena[1], Kristiina Jokinen[2], Graham Wilcock[3]

Abstract The design and characteristics to enhance the interaction abilities of the Nao humanoid robot by extending its communicative behavior with non-verbal gestures (hand and head movements, and gaze following). A set of non-verbal gestures were identified that Nao could use for enhancing its presentation and turn-management capabilities in conversational interactions. We discuss our approach for modeling and synthesizing gestures on the Nao robot. A scheme for system evaluation that compares the values of users' expectations and actual experiences has been presented. We found that open arm gestures, head movements and gaze following could significantly enhance Nao's ability to be expressive and appear lively, and to engage human users in conversational interactions .

H.Title: Gestures :Discovering Gerstures Descriptors Associated with Spoken Utterances
Author: Shogo Okada[1], Kahuri Otsuka[2]

Abstract The design and characteristic of modeling the relationship between gestures and spoken words during continuous hand motions, speech and language data. The problem setting and modeling is defined as "gesture association" (associating gestures with words). We present a framework to associate spoken words with hand motion data observed from a optical motion-capture system. The framework identifies pairs of hand motions (gesture) and uttered words as training samples by autonomously aligning the speech and the co-occurring gestures. Using the samples, a supervised learning approach is undertaken to learn a model to discriminate between gestures that co-occur with a spoken word (w) and gestures when the word w is not being spoken. To detect gestures and associate them with spoken words, we extract (1) a trajectory signal feature set, (2) features of various gesture phases from hand motion data and (3) primitive gesture patterns learned by Sift Invariant Sparse Coding (SISC). Then Hidden Markov Model (HMM) and Support Vector Machine (SVM) classifiers were trained with these three-feature sets. In experiments, the classification accuracy achieved for 80 words was above 60 % (the maximum accuracy achieved was 71 %) using the proposed framework. In particular, SVMs trained with features (2) and (3)

outperform an HMM trained with feature (1) (prepared as a baseline model). The results show that the gesture phase and primitive patterns trained by SISC are effective gesture features for recognizing the words accompanying gestures

I Title: 3D Model Classification and Recognition Method Based on Tensor Principal Component Analysis
Author: Xinying Wang[1],Yuanyang Yue[2]

Abstract This paper propose a 3D model classification and recognition method based on tensor principal component analysis. Firstly, the 3D model is transformed into 2D view and expressed in tensor, and then feature vector from multiple angles of tensor is extracted. On the basis of multi-linear principal component analysis and linear discriminate analysis, we present a method for merged weighted multi-linear principal component analysis with weighted linear discriminate analysis (WMPCA + WLDA). This method is not only can save spatial correlation of the projected view, but also can improve feature recognition of the 3D model through using class label information and features weight. Experiments on Princeton Shape Benchmark has shown that the method is superior to the conventional MPCA, MPCA + LDA, WMPCA + LDA and other methods on classification results

J. Title: A gesture recognition system for detecting behavioural patterns of ADHD
Author: Lyn Chao-ling Chen[1], Kuan-Wen Chen[2], Yi-Ping Hung[3]

Abstract A paper to develop a non-invasive sleep monitoring system to distinguish sleep disturbances based on multiple sensors. A device with an infrared depth sensor, a RGB camera, and a four-microphone array is used to detect three types of events: motion events, lighting events, and sound events. The events are classified by an epoch approach algorithm and provide a graphical sleep diagram. Experimental results in sleep condition show the efficiency and reliability of our system, and it is convenient and cost effective to be used in home context.

III.PROPOSED WORK

In this paper, proposed to computer vision and acoustic model based find out the human stress using ARM11 processor. Computer vision means contour points based stress identification. How many contour points are detected in the camera region, hence low contours are detected in the camera region this is low level stress and high contour points detected means high level stress. Acoustic model means some of voice modulation data are stored in database, that voice data are stored in various frequencies, so compared to that voice data, so known the which state of stress in human. Finally known the human stress are compared to the acoustic model and computer vision data. And send the stress people information to the service desk.

DATA FLOW DIAGRAM:

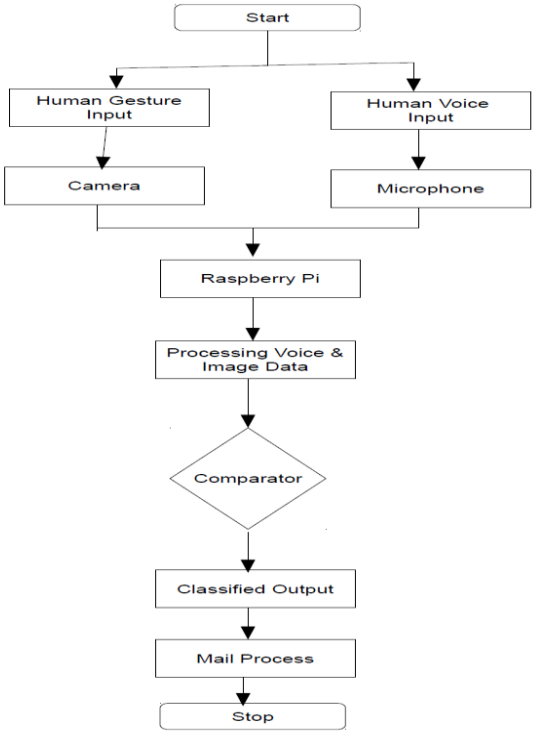

Figure 1.Data flow diagram

IV.CONCLUSION

To summarize, in the framework of automatic surveillance , we investigated modalities of how speech and gestures communicate stress and how they can be used for automatic assessment of stress. For this purpose we proposed a human model of stress communication, which distinguishes between the semantics expressed by speech and gestures, and the way in which the messages are delivered (modulation). We assessed how these components convey stress based on human annotated labels. As a next step, we proposed a new method for automatic stress prediction based on a decomposition of stress into a set of intermediate level variables. The intermediate level variables were obtained by operationalizing the communication components of the human model. We validated our model for automatic stress prediction and obtained significant improvements over a baseline predictor based on decision level fusion on the audio, text and video features.

V.REFERENCES

[1] Z. Yang, "Multi-modal aggression detection in trains," Ph.D. dis-sertation, Delft Univ. Technol., Delft, Netherlands, 2009.

[2] L. Berkowitz and R. G. Geen, "Affective aggression: The role of stress, pain, and negative affect," in Human Aggression: Theories, Research, and Implications for Social Policy, E. Donnerstein, Ed. Amsterdam, The Netherlands: Elsevier, 1998, pp. 49–72.

[3] J. Sprague, E. Verona, W. Kalkhoff, and A. Kilmer, "Moderators and mediators of the stress-aggression relationship: Executive function and state anger," Emotion, vol. 11, no. 1, pp. 61–73, 2011.

[4] R. S. Lazarus and S. Folkman, Stress, Appraisal, and Coping. New York, NY, USA: Springer, 1984.

[5] D. McNeill, "So you think gestures are nonverbal?" Psychological Rev., vol. 92, no. 3, pp. 350–371, 1985.

[6] A. Kendon, Gesture: Visible Action as Utterance. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[7] R. M. Krauss, R. Dushay, Y. Chen, and F. Rauscher, "The commu-nicative value of conversational hand gestures," J. Exp. Social Psy-chol., vol. 31, pp. 533–552, 1995.

[8] P. Juslin and K. Scherer, "Vocal expression of affect," in The New Handbook of Methods in Nonverbal Behavior Research, J. Harrigan, R. Rosenthal, and K. Scherer, Eds. London, U.K.: Oxford Univ. Press 2005, pp. 65–135.

[9] K. R. Scherer, T. Johnstone, and G. Klasmeyer, Vocal Expression of Emotion, series R. J. Davidson, H. Goldsmith, K. R. Scherer (Eds.) - Handbook of the Affective Sciences. London, U.K.: Oxford Univ. Press, 2003.

[10] K. Scherer, "Voice, stress, and emotion," in Dynamics of Stress. M. H. Appley, and R. Trumbull, Eds. New York, NY, USA: Plenum., 1986, pp. 159–181.

[11] A. Batliner, S. Steidl, B. Schuller, D. Seppi, T. Vogt, J. Wagner, L. Devillers, L. Vidrascu, V. Aharonson, L. Kessous, and N. Amir. (2011). Whodunnit - Searching for the most important feature types signalling emotion-related user states in speech. Comput. Speech Language [Online]. 25(1), pp. 4–28 Available: http://www5. informatik.uni-erlangen.de/Forschung/Publikationen/2010/Bat liner10-WSF.pdf

[12] P. Ekman, "Emotional and conversational nonverbal signals," in

Language, Knowledge, and Representation, M. Larrazabal, L. Miranda. Eds. Norwell, MA, USA: Kluwer, 2004.

[13] P. Ekman and W. V. Friesen, "The repertoire of nonverbal behav-ior: Categories, origins, usage, and coding," Semiotica, vol. 1, pp. 49–98, 1969.

[14] J. Hansen and S. Patil, "Speech under stress: Analysis, modeling and recognition," in Speaker Classification I, C. Muller,€ Ed. New York, NY, USA: Springer, 2007, pp. 108–137.

[15] J. Hansen, S. Bou-Ghazale, R. Sarikaya, and B. Pellom, "Getting started with SUSAS: A speech under simulated and actual stress database," in Proc. 5th Eur. Conf. Speech Commun. Technol., 1997, vol. 97, no. 4, pp. 1743–46.