JOURNAL OF RESOURCE MANAGEMENT AND TECHNOLOGY

ISSN NO: 0745-6999

Survey on features extraction methods for stress assessment from videos

FASIHA ANJUM ANSARI
Research Scholar
Department of Computer Science & Engg.,
Shri Jagdish Prasad Jhabarmal Tibrewala University Jhunjhunu, Rajasthan, INDIA
fasihaanjum123@gmail.com

Dr. SABERA BEGUM
Research Co-Guide
Department Of Computer Science & Engg.,
Shri Jagdish Prasad Jhabarmal Tibrewala University Jhunjhunu, Rajasthan , INDIA

Abstract—In modern work environment, stress is inevitable. Various work related factors like peer pressure, timelines, quality expectations etc drive the employees to higher stress levels. Continuous exposure to stress affects the mental and physical well being of the employees. It is very important for timely monitoring of stress and adopt suitable medications to ensure the general well being of the employee. Regular counseling sessions is most adopted stress relief program in many companies.. But these mechanisms are not effective as questionnaire based stress evaluation can be bypassed by employee due to various factors like lack of time, tired etc. Recently, non intrusive stress monitoring has been proposed to do automatic stress assessment based on video feeds collected from work environment. The effectiveness or accuracy of these methods is strongly dependent on the features collected from video feeds. This work does a survey on different features extraction methods for automatic stress assessment.

I. INTRODUCTION

Stress is human body's response to external physical and emotional events faced by a person. In modern work environment, stress has become a inevitable part. Increasing problem of stress and mental health in organization has become a major concern. According to World Health Organization's (WHO) report, workplace stress is causing low productivity among employees and costing the global economy nearly $1 trillion annually. Stress is a silent killer and it is the root cause for depression and workplace anxiety in India. Recent survey [1] by American institute of stress, 80% of employees feel stress on the job and nearly half of them needed help to manage the workplace stress. Early detection and medication is the solution to manage work place stress. In addition to mental health, stress also affects the physical health. For example, an observational data suggests an average 50% increased risk for coronary heart disease among employees with work stress. Continuous exposure to stress affects both physical and mental health causing problems like abnormal cardiac rhythm, depression etc. Studies carried on animals and humans suggest that stress can affect immune system leading to cancer.

Conventional questionnaire based stress assessment is ineffective as most employee's skip it due to lack of time, interest and fear of self esteem. Biochemical and physiological indicators can also be used for stress assessment. But these methods are intrusive and the methods themselves become a source of stress. Research has focused on finding objective, non-intrusive, continuous and quantitative ways to detect stress.

Automatic stress assessment by analyzing the visual cues from the video surveillance cameras at work place is a non intrusive technique for stress assessment. The effectiveness of this method depends on the type of visual features extracted from the videos frames. In this work, we survey the different features that can be extracted from the surveillance videos and used for stress assessment. The pros and cons of each of the method and the open issues for further research are analyzed.

ISSN NO: 0745-6999

II. SURVEY

Zhang et al (2020) extracted deep learning features from faces and used it for classifying the face to two level of expressive and non expressive. Emotions like angry, disgust, fear, happy, sad, and surprise were under the category of expressive. The subjects were shown different videos and their expression is collected through surveillance camera and use for classification. The entire face image is passed to Resnet deep learning model for classification. From the entire video two key frames for expression classification is selected manually.

The method was able to classify expressions with an accuracy of 85.42%. The experimental setup is limited. It groups stress and non stress factors into one category of expressive.

Giannakakis et al (2017) developed a framework for detection of stress from facial cues recorded in a video. Features of eye related events, mouth activity, head motion and heart rate collected through photoplethysmography are used to classify three states of neutral, stressed and relaxed. The facial cues used in this study involved eye related features (blinks, eye aperture), mouth activity features (VTI, ENR, median, variance, skewness, kurtosis and Shannon entropy), head movement amplitude, head velocity and heart rate estimation derived from variations in facial skin colour.Facial features were found to achieve about 90.5% classification accuracy. The accuracy was measured in a restricted way only for four different activities and working was not considered in the activities.

Pampouchidou et al (2016) proposed an algorithm to extract mouth activity features and use it for stress detection. Eigen features were extracted from the mouth opening and mouth deformations and used as features for stress assessment. The method was able to achieve a classification accuracy of 89.17%. The sample size was very limited to 23 volunteers. The approach is semi automatic and involves manual selection of baseline.

Gao et al (2014) proposed a system to monitor the attentive and emotional state of the drivers through facial expressions. Anger and disgust are considered as stress related symptom in this work. Holistic affine wrapping and local descriptors are used for feature extraction from the facial regions. Discrete cosine transform is used for extracting holistic affine features. Scalable invariant feature transform is used for extraction of local descriptors. The method was able to achieve a classification accuracy of 85%. The features are sensitive to pose variations, so the method is pose specific.

Viegas et al (2018) used 17 facial action units for stress detection. Authors extracted 17 different Action Units (AUs) from upper-level to lower-level face frame-wise. They were able to detect stress with an accuracy of 74%. The stress was evaluated for activity of typing with rest between typing sessions. The approach is not generic and subject dependent.

Gavrilescu et al (2019) proposed an algorithm to evaluate Depression Anxiety Stress Scale (DASS) levels by analyzing facial expressions using Facial Action Coding System (FACS). It is a three layer approach. In the first layer, Active Appearance Models (AAM) and a set of multiclass Support Vector Machines (SVM) are used for Action Unit (AU) classification; in the second layer, a matrix is built containing the AUs' intensity levels; and in the third layer, an optimal feedforward neural network (FFNN) analyzes the matrix from the second layer in a pattern recognition task, predicting the DASS levels. The method was achieve a classification accuracy of 90.2% for stress. The approach is very specific to health screening and it applicability to job environment is not explored.

Prasetio et al (2018) proposed a histogram based feature extraction method for stress detection from face. From the three areas of pair of eye, nose and mouth , histogram features are extracted using Gabor and histogram of gradient algorithms. This method is able to provide a classification accuracy of 86.7%. The method was tested only against 10 subjects.

Tamura et al (2018) proposed a method to detect stress based on facial frontal features. The facial image is split to pair of eyes, nose and mouth parts. Facial features are extracted from these three parts using DoG,HoG and DWT. The method was able to provide an accuracy of 86% for stress classification.

Pediaditis et al (2015) experimented with different facial features for stress detection. Head motion, eye, mouth related features are extracted from face and used for stress classification. Eye related features like eye blink was found to be a better indicator of stress. The experimented was conducted across 22 subjects and the accuracy of the method was about 75%.

Bevilacqua et al (2018) proposed an algorithm for stress detection from facial features during online games. Seven different facial features were extracted from the facial region. These features correlated to the activity of set of facial muscles. The study found that features of mouth outer, mouth corner, eye area, eyebrow activity, and face area were found to have higher correlation to stress during online gaming. The study was conducted only with limited sample size of 20.

Zacharatos et al (2014) studied the correlation between body postures and emotions. The study found a strong correlation between the body postures and the emotions identified in the Valence-arousal space model (Posner et al)

The summary of the survey is given in Table 1.

III. ISSUES

The open issues in the existing solutions for stress assessment from videos is listed below

- Almost all the methods are tested with limited samples and variances.
- Body postures is not given importance for increasing the accuracy of classification
- Variance in features clues across the frames is not considered in many of the works.
- Personalization is stress assessment is lacking
- Almost all the methods were tested for limited activities and don’t consider work environment.
- Manual feature selection

IV. DISCUSSION ON OPEN ISSUES

Issue 1: The current methods for stress assessment from videos are tested against very limited dataset less than 50 undermining its reliability. The solutions must be tested against large dataset with different demographics. Only then the reliability of the solution can be ensured.

Issue 2: Most of the approaches are based only on facial features. Integrating body posture features along with facial features can increase the accuracy of the stress classification. The correlation between body postures and the emotional state is proved in many works. But the selection of relevant body postures and its relevance to emotion considering the demographics is not yet explored in any of the works. This can be explored and the relevant body posture features can be augmented with facial features for more accurate stress assessment.

Issue 3: Most of stress assessment methods are snapshot based. They do stress assessment for an image. But with the availability of video stream, the variance in feature cues can be used to establish personalized baselines. Stress assessment methods based on this variance in features and baselines can provide higher accuracy.

Issue 4: Personalization is very important in stress assessment as some of the facial features for stress assessment is subject dependent. In most of existing solution, establishing base line for subject and personalizing the stress assessment based on the base line of the subject is not considered.

Issue 5: Most of the stress assessment methods are tested for specific activities and none of them consider work environment which involves multitudes of activities.

Issue 6: In most of the approaches, the facial regions and the features to be extracted are selected manually. With availability of Deep learning feature extraction methods, the best set of features with higher correlation to stress can be learnt in a better way.

V. CONCLUSION

The paper summarizes the current works on stress assessment from videos. The existing solutions have been detailed and the problems in each solution are documented. The open areas for further search on the problem studies are listed with discussion on the issue and prospective solution for those issues. Further work will be on design on efficient solutions to address the identified open issues.

ISSN NO: 0745-6999

Table 1 Survey summary

Solution	Brief	Problems
Zhang et al (2020)	The entire face image is passed to Resnet deep learning model for classification. From the entire video two key frames for expression classification is selected manually	The experimental setup is limited. It groups stress and non stress factors into one category of expressive
Giannakakis et al (2017)	Features of eye related events, mouth activity, head motion and heart rate collected through photoplethysmography are used to classify three states of neutral, stressed and relaxed	The accuracy was measured in a restricted way only for four different activities and working was not considered in the activities.
Pampouchidou et al (2016)	Eigen features were extracted from the mouth opening and mouth deformations and used as features for stress assessment	The sample size was very limited to 23 volunteers. The approach is semi automatic and involves manual selection of baseline.
Gao et al (2014)	Holistic affine wrapping and local descriptors are used for feature extraction from the facial regions	The features are sensitive to pose variations, so the method is pose specific.
Viegas et al (2018)	Authors extracted 17 different Action Units (AUs) from upper-level to lower-level face frame-wise	The approach is not generic and subject dependent
Gavrilescu et al (2019)	Three layer approach based on facial features	The approach is very specific to health screening and it applicability to job environment is not explored.
Prasetio et al (2018)	From the three areas of pair of eye, nose and mouth , histogram features are extracted using Gabor and histogram of gradient algorithms	Limited dataset of 10 subjects
Tamura et al (2018)	The facial image is split to pair of eyes, nose and mouth parts. Facial features are extracted from these three parts using DoG,HoG and DWT.	Limited dataset of 10 subjects
Pediaditis et al (2015)	Head motion, eye, mouth related features are extracted from face and used for stress classification	Limited dataset of 20 subjects
Bevilacqua et al (2018)	Seven different facial features were extracted from the facial region. These features correlated to the activity of set of facial muscles	Limited dataset of 20 subjects

REFERENCES

[1] Zhang, H., Feng, L., Li, N., Jin, Z., & Cao, L. (2020). Video-Based Stress Detection through Deep Learning. Sensors (Basel, Switzerland), 20(19), 5552. https://doi.org/10.3390/s20195552

[2] Giannakakis, G.; Pediaditis, M.; Manousos, D.; Kazantzaki, E.; Chiarugi, F.; Simos, P.G.; Marias, K.; Tsiknakis, M. Stress and anxiety detection using facial cues from videos. Biomed. Signal Process. Control. 2017, 31, 89–101

[3] Pampouchidou, A.; Pediaditis, M.; Chiarugi, F.; Marias, K.; Simos, P.; Yang, F.; Meriaudeau, F.; Tsiknakis, M. Automated characterization of mouth activity for stress and anxiety assessment. In Proceedings of the IEEE International Conference on Imaging Systems and Techniques (IST), Chania, Greece, 4–6 October 2016; pp. 356–361.

[4] Gao, H.; Yüce, A.; Thiran, J. Detecting emotional stress from facial expressions for driving safety. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 5961–5965

[5] Viegas, C.; Lau, S.; Maxion, R.; Hauptmann, A. Towards Independent Stress Detection: A Dependent Model Using Facial Action Units. In Proceedings of the International Conference on Content-Based Multimedia Indexing (CBMI), La Rochelle, France, 4–6 September 2018; pp. 1–6

[6] Gavrilescu, M.; Vizireanu, N. Predicting Depression, Anxiety, and Stress Levels from Videos Using the Facial Action Coding System. Sensors 2019, 19, 3693

[7] Tamura, H.; Tanno, K. The Facial Stress Recognition Based on Multi-histogram Features and Convolutional Neural Network. In Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 881–887

[8] Prasetio, B.H.; Tamura, H.; Tanno, K. Support Vector Slant Binary Tree Architecture for Facial Stress Recognition Based on Gabor and HOG Feature. In Proceedings of the 2018 International Workshop on Big Data and Information Security (IWBIS), Jakarta, Indonesia, 12–13 May 2018; pp. 63–68.

[9] Pediaditis, M.; Giannakakis, G.; Chiarugi, F.; Manousos, D.; Pampouchidou, A.; Christinaki, E.; Iatraki, G.; Kazantzaki, E.; Simos, P.G.; Marias, K.; et al. Extraction of facial features as indicators of stress and anxiety. In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 3711–3714

[10] Fernando Bevilacqua, Henrik Engström, Per Backlund, "Automated Analysis of Facial Cues from Videos as a Potential Method for Differentiating Stress and Boredom of Players in Games", International Journal of Computer Games Technology, vol. 2018, Article ID 8734540, 14 pages, 2018. https://doi.org/10.1155/2018/8734540

[11] H. Zacharatos, C. Gatzoulis, and Y. L. Chrysanthou, “Automatic emotion recognition based on body movement analysis: A survey,” IEEE Computer Graphics and Applications, vol. 34, no. 6, article no. 106, pp. 35–45, 2014

[12] Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology, 17(3), 715–734. https://doi.org/10.1017/S0954579405050340