A SURVEY ON HUMAN EMOTION ANALYSIS USING THERMAL IMAGING AND PHYSIOLOGICAL VARIABLES

C.M.Naveen Kumar[1], G. Shivakumar[2]
[1,2] Dept. of EIE, Malnad College of Engineering, Hassan-573202, Karnataka, India

Abstract

Enhancement of naturalism in interaction of human beings and computers is possible if computers can identify, perceive, and react to human emotions. Inspite of numerous available techniques for recognition of human emotions based on visual facial expressions, thermal facial imaging and anatomical parameters; performance, robustness and accuracy of an emotion recognizer can be enhanced by combining one modality with other. This paper investigates the strong suit and deficiencies of the currently available emotion recognition systems based on facial expressions, facial thermal imaging and anatomical parameters.

Index Terms: Emotion recognition, physiological variables, facial expression, thermal ir imaging.

I. INTRODUCTION

The involvement and impact of computer and its applications in todays world are magnified. Like our human-human interaction, improving naturalism in the interaction with them is critical. Naturalism can be enhanced through the development of capability that infers emotional states depending on overt / covert signals of these states. Based on the emotional states, we can alter our responses and behavioural patterns that ensure optimization and convergence of the interaction process. Literature related to emotions spread over several disciplines. Irrefutably, detailed investigation of the human brain is possible with improvements in the technology; specifically the emotional circuits' association in emotion recognition. Therefore a detailed and comprehensive understanding of the structure and function of emotion recognition in the brain can be achieved. Technological advances enhanced emotions investigation with its physical correlates which allows to develop powerful emotion analysis interface in multimodal approach [12].

Emotions are classified into: primary emotions and secondary emotions [20]. Primary emotions are simple emotions like attractions, innate aversions and will arise spontaneously in the low-level limbic circuit. Secondary emotions incorporate high-level cortical processing and require conscious awareness and cognitive processing because they are more sophisticated and refined.

Psychological theory categorizes human emotions into: surprise, fear, disgust, anger, happiness, and sadness. In order to avoid impeachment sometimes people would intentionally falsify their emotional states. Misrepresentation of emotions can be determined through physiological measurements [21]. People learn to willingly hinder spontaneous emotional expression in order to obey with group membership, culture and gender.. For example, anger is replaced with a fake smile since anger is treated as unsociable and is discouraged. [19].

Direct measurement of Human's state of emotion is not feasible. Emotions are interrelated to external and/or internal factors that are perceptible which allows its investigation. The internal factors come from different parts of the body in several forms such as electroencephalography, heart rate, heart rate variability and others[18].

In a dark environment, thermal imaging improves the visibility of objects through the determination of the objects' IR radiation and developing an image by this information. IR energy is emitted by all objects related to its temperature. Thermal IR spectrum based facial tracking has received enhanced attention in recent years. Since thermal imaging is less sensitive to lighting conditions; earlier it was used in surveillance applications and face recognition [2]. Based on the radiated heat from the face thermal imaging technology is used to analyse stress levels. Facial thermal images are better in comparison with visual images. Facial temperature measured using thermal camera is independent of darkness, skin color and lighting condition.

II. RELATED WORK

Temporal data corresponding to facial temperature is used for analysis of emotion recognition in [17]. Firstly, extraction of information related to facial temperature data is done and facial region is segmented into sub regions. Then extraction of histogram and statistical features related to differential temperature is done from the difference matrix corresponding to facial temperature. Classification is done with Discrete Hidden Markov Models for every feature. Here recognition results influence feature selection strategy in training set was proposed.

Regions of Interest tracking technique in thermal image videos are discussed in [16]. Here vibrant signs are measured to recognize emotions. It's framework includes three modules. Updating the template of the adaptive particle filter tracker is done which is correlated to learning decision module to circumvent drifting. ROI is detected from randomized classifier is used. The Classifier's output is enhanced by removing false positives using a projected geometrical constraint.

Analysis of human emotion using a novel non-invasive method to classify using thermal images of face is discussed in [13]. Hu's moment invariants of different patches are merged with histogram statistical feature and it is utilized as robust features in multiclass support vector machine based classification.

Thermal infrared data is utilized to derive time, frequency, and time-frequency features that can be used to differentiate self-reported individual emotional states in retort to visual stimuli obtained from the International Affective Pictures System. To differentiate baseline and affect states, a total of six binary classification tasks were investigated. Emotional states were distinguished from subject-reported levels of arousal and valence [3].

The thermal IR imaging is proven validated to be a reliable tool for the non-invasive contactless evaluation of vital signs, psychophysiological responses and emotional states [15]. The current trends in thermal infrared imaging in the recognition of emotional states are highlighted to know the shortcomings and capabilities of data analysis in image processing.

Current approaches to know the human emotional states based on facial images, gestures and physiological signals demands considerable and exclusive equipment. Topographic distribution of facial temperature demonstrated exclusive features which are dependent on emotional arousal including related measures of anatomical variables. Facial expression for specific emotion is predicted by Facial Thermal Feature points (FTFP) and is mapped with underlying facial muscles [14].

Tracking the facial tissue is influenced and affected with the physiological including positional changes; to eliminate this, a tracker dependent on particle filter which is based on probabilistic template function for spatial including temporal smoothing components is projected in [2]. In predict and update loop, the non-linear motion can be handled by this filter. Update operation is affected by probabilistic template algorithm. In facial thermal imaging, particle filter can generate well established tracking behaviour and may track region of interests' position in the current frame based on template matching.

Levels of stress in humans are identified from the heat radiated from the face [1]. Deviations in initial thermal levels can be compensated by baseline thermal corrections meanwhile skin surface temperature is correlated with the facial expressions body metabolism, deviations in musculo-thermal activities, surrounding environments' thermal emissions and finally illness that may be accomplished by subtracting the average of the four baseline questions by every response of interview of the subject. Extracted signatures are concatenated and are then condensed to lesser dimensions using PC analysis. For data representation, classifier dependent on k-nearest neighbour algorithm can be utilized to categorize the thermal responses with different strategies.

Ability to predict lie/truth responses based on with-in person methodology is achieved.

The physiological signal processing dependent emotion recognition system is divided into four sections: signal pre-processing extracted signals, extraction of biological features, matching and classification of features [9]. In each section; details, performance and characteristics of current methods are analysed. Here relationships between factors affecting human emotions and emotional state is significant to simplify model and reduce dimension is suggested as a future work.

Summarization of techniques for physiological measurements for the autonomic and central nervous system is done in [10]. Empirical studies of psychophysiology proved a strong relationship between physiological reactions and emotional/affective states of humans. Mapping of the physiological patterns with its exact types of emotions is a complex task. Physiological patterns differ from one user to another and with different situations. Combining physiological signals with other modalities is validated as a next positive step in recognizing emotions.

Facial expression and anatomical variable based multimodal drivers' emotion recognition technique is proposed in [11]. Here experiments to determine associations of heart rate, skin conductance and temperature are determined for two different emotions: fear and amusement.

Carlos et. al., discussed associations of the facial expressions, acoustic information and its uses and shortcomings in [19]. Using three different systems for audio, facial expression and bimodal information four emotions are recognized. Facial expression has good result than acoustic information. Combinations of both the information are increases efficiency and accuracy. With audio signals, sadness and fear can be recognized easily and with video analysis, anger and happiness can be better recognized.

Eun H J et. al., discussed the difference of pain, surprise and boredom with physiological signals [5]. Possibility of six different classification techniques for their classification is discussed. The recognition accuracy of machine learning algorithms dependent on non-linear methods including SOM and CART is lower in comparison with the one based on linear methods. This method can be validated if gender and age effects are considered which will contribute to anthropology in an effective way.

M Malkawi et. al., designed a trainable adaptive neuro-fuzzy system that deals with 14 human factors [7]. Human factors are linked with some specific rules that have correlation with human emotions and can link specific emotion to some of the factors. Diverse variety of input/output membership functions are used to build six models. Diverse kinds of input arrays are used in training.

N Krupa et. al., designed physiological signals acquisition system which is a wearable wristband for [4]. Emotions such as neutral, involvement and happy are influenced Autism Spectrum Disorder (ASD) which can be classified using SVM algorithm. Changes in Galvanic SR and HR Variability are used in categorizing emotions.

The deep Boltzmann machine (DBM) is designed [6] that will cram thermal features in order to recognize emotions using TIR facial images [6]. The face is traced at the starting stage and then normalized using TIR images. Then, a DBM model which consists of two layers is trained. DBM parameters are fine-tuned for recognizing emotions then pre-training of feature learning is done.

Anna E et. al., provided a detailed information on experimental arrangement, the acquiring scenario, the stimuli generation and the data to get naturally prompted emotional facial expressions using the high emotions content videos and it reports experimental data to determine the effects of emotions on tasks of memory word recognition. The database that results include contains thermal as well as visible emotional facial expressions [8].

III. RESULT ANALYSIS

Fear, happy and disgust great impact in facial temperature when compared with other basic emotions. The features corresponding to forehead's region is the superior and that of mouth region's features are the minimum [17].

The tracker proposed in [16] has the finest performance because it magnifies the efficiency of the Tracking-Learning-Detection by using geometrical constraint so that false positives and circumvent drifting are eliminated [16].

KTFE database is used in experiments. Here training set contains 60% of images and testing set contains 40% of images. The classification performance of third order polynomial kernel for is superior in comparison with linear kernel, RBF kernel, Quadratic kernel [13].

Classification among an affective state and baseline produced sophisticated mean adjusted accuracies than classification between high versus low arousal or valence. The features used in classification task are picked highly within the periorbital and nasal/maxilliary regions[3].

Manual segmentation of ROI in every frame produces the data of ground-truth. For larger database, manual ground truthing may not be possible so a strategy to give tracking results using three trackers are utilized. If a tracker is failed then manual reposition is done from that particular onwards and error is minimized. Ground-truth trackers are formed as the means of the individual corrected trackers [2].

Consequences of Between-Person approach validated that the deceptive behaviour is not utilized for generalisation across the whole population. Using Within-Person Approach based on PCA from thermal signatures, an average classification accuracy with for each subject measured over 50 runs and tested using leave-one-question-out , five-fold cross-validation and two-fold cross-validation is robust and accurate[1].

Classes of features and combination of features that may be more appropriate for categorization of particular emotion is to yet be recognised. An emotional model that conforms the laws of human emotions is to be developed [9].

Several well-known interactions prevent simple elucidation of physiological signals by means of emotion channel. A deep breath may influence supplementary measures because respiration is diligently associated to cardiac function [10].

Higher heart rate and lesser skin conductance will be yielded for emotion of delight in comparison with the emotion of fear. The change between average of skin temperature in those emotions are trivial[11].

The emotions pair that are jumbled in one modality are effortlessly classified in the other. Eventhough audio based system exhibits eviler performance, its extracted features carries valuable data corresponding emotions that may not be obtained using facial expressions. [19].

The Discriminant Function Analysis produced finest classification results for pain, surprise and boredom emotions. These emotions can be precisely classified using linear methods in comparison with non-linear methods like SOM, CART [5].

The section of dataset for training has an influence on emotions outcome. Diverse training data sets can produce different emotions output eventhough diverse groups of people may react inversely to emotions stimuli [7].

RBF kernel function is used for classification. A 5 fold cross validation is done to compute the accuracy and counting matrix of the algorithm with acquired data samples[4].

The performance of the DBM is superior in comparison with previous techniques for statistical parameters as wel as the area under ROC curve of the DBM. The rotation of the head or the occlusion by the hair is restored up to some extent. The information corresponding to forehead region, mouth and eye is more trustworthy to categorize low and high valence when compared to remaining facial areas. [6].

Concatenation of various features together is considered to enhance the overall performance due to availability of numerous diverse features. Pre-processing of features vectors dataset using PCA or Linear Discriminant Analysis (LDA) before the classification phase enhances the accuracy for categorization [14].

IV. CONCLUDING REMARKS

Thermal images are unaffected for variation in illumination and shadow whereas visible image based facial expression recognition is influenced by those parameters [8]. Topographic distribution of facial cutaneous temperature demonstrated explicit features which directly correlated to emotional arousal and associated measures of standard anatomical variables [14]. The TIR is validated as a stable technique for the non-invasive contactless assessment of vibrant signs, psychophysiological reactions and state of emotion [15]. The automatic recognition of face, feature extraction, optimization of parameter in classification system and decision strategy for ultimate results boosts its performance. [17]. Extensive study is required to know likelihood of integration of physiological facts with automatic recognition of facial expressions for obtaining an atlas of the emotions' thermal signatures [15]. The amalgamation of various modalities for recognizing emotions improves rate of accuracy for recognition in comparison with single modal analysis [10]. Hybrid models may be identified at various phases for boosting the performance recognition [4]. Patterns of anatomical variables are subjected to vary from one subject to another and from one circumstance to another [10].

REFERENCES

[1] Bashar A. R et.al., "Thermal Facial Analysis for Deception Detection", IEEE TIFS, Vol. 9, No. 6, Page(s) : 1015-1023, JUNE 2014.

[2] Yan Zhou et.al.,"Spatio-temporal Smoothing as a Basis for Facial Tissue Tracking in TI", IEEE Transactions on Biomedical Engg, Vol. 60, No. 5, pp. 1280-1289, MAY 2013.

[3] Brian R N et.al., "Classifying Affective States Using TIR Imaging of the Human Face", IEEE Transactions on Biomedical Engg , Vol.57 , Issue: 4, Page(s): 979 – 987, April 2010.

[4] N Krupa et.al., "Recognition of emotions in autistic children using physiological signals", DOI 10.1007/s12553-016-0129-3, IUPESM and Springer-Verlag Berlin Heidelberg, 2016.

[5] Eun-H et.al.,"Analysis of physiological signals for recognition of boredom, pain, and surprise emotions", Journal of Physiological Anthropology, Page(s): 25-34, 2015.

[6] Shangfei W et.al.,"Emotion recognition from thermal infrared images using deep Boltzmann machine", pp: 609–618, DOI 10.1007/s11704-014-3295-3, HEP and Springer-Verlag Berlin Heidelberg 2014.

[7] M Malkawi et.al., "Artificial neuro fuzzy logic system for detecting human emotions", HCIS, Springer Open Journal, pp. 1-13, 2013.

[8] Anna E et.al., "A Naturalistic Database of Thermal Emotional Facial Expressions and Effects of Induced Emotions on Memory", LNCS 7403, pp. 158–173, Springer-Verlag Berlin-Heidelberg 2012.

[9] Naiyu Wu et.al., "Emotion Recognition Based on Physiological Signals", BICS, LNAI 7366, pp. 311–320, Springer-Verlag Berlin-Heidelberg 2012.

[10] R. B Knapp et.al., "Physiological Signals and Their Use in Augmenting Emotion Recognition for HMI", Emotion-Oriented Systems, Cognitive Technologies, Springer-Verlag Berlin Heidelberg, pp. 133-159, 2011.

[11] H. Leng et.al., "An Experimental Study on Physiological Parameters Toward Driver Emotion Recognition", HCII , LNCS 4566, , Springer-Verlag Berlin Heidelberg, pp. 237–246, 2007.

[12] N. Fragopanagos et.al.,"Emotion recognition in human–computer interaction", Elsevier, Neural Networks 18, pp. 389–405, 2005.

[13] A Basu et.al., "Human emotion recognition from facial thermal image based on fused statistical feature and multi-class SVM", 2015 INDICON, pp. 1-5, 17-20 Dec. 2015.

[14] M. H A Latif et.al., "Thermal imaging based affective state recognition", 2015 IEEE-IRIS ,pp. 214-219, 18-20 Oct. 2015.

[15] D Cardone et.al., "The thermal dimension of psychophysiological and emotional responses revealed by thermal infrared imaging", 2014 IEEE-ICIP ,pp. 1942-1946, 27-30 Oct. 2014.

[16] E Mostafa et.al.,"Long term facial parts tracking in thermal imaging for uncooperative emotion recognition", IEEE- BTAS, pp. 1-6, Sept. 29 2013-Oct. 2 2013.

[17] Z Liu et.al., "Emotion Recognition Using Hidden Markov Models from Facial Temperature Sequence", International Conference on Affective Computing and Intelligent Interaction (Springer LNCS 6975), pp. 240-247 October 2011.

[18]Kreibig S: "Autonomic nervous system activity in emotion: a review", biological psychology, Elsevier, Volume 84, Issue 3, pp. 394–421, July 2010.

[19] C Busso et.al., "Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information", International Conference on Mutimodal Interfaces, State College, PA, ACM 1-58113-890-3/04/0010, 2004.

[20] Antonio R ,"Descartes' error: Emotion, reason, and the human brain". New York: PPG, 1994, ISBN: 0-380-72647-5.