DECODING EMOTIONAL EXPERIENCES THROUGH PHYSIOLOGICAL SIGNAL PROCESSING

Maria S. Perez-Rosero[1], *Behnaz Rezaei*[1], *Murat Akcakaya*[2], *and Sarah Ostadabbas*[1]

[1]Electrical and Computer Engineering Department, Northeastern University, MA, USA.
[2]Electrical and Computer Engineering Department, University of Pittsburgh, PA, USA.

ABSTRACT

All modern emotion theoretical views assume a role for peripheral physiological changes during emotional experiences. In this paper, we explored the correlation between autonomically-mediated changes in multimodal bodily signals and discrete emotional states. In order to fully exploit the information in each modality, week learners based on individual signal modalities are built and then fused to formed a robust inference model. To validate our model, three specific physiological signals including Electromyogram (EMG), Blood Volume Pressure (BVP) and Galvanic Skin Response (GSR) recorded during eight emotional states were analyzed. Our approach showed 88.1% emotion recognition accuracy, which outperformed the conventional Support Vector Machine (SVM) classifier with 17% accuracy improvement. Furthermore, in order to avoid information redundancy and the resultant over-fitting, a feature reduction method is proposed based on a correlation analysis to optimize the number of features required for training and validating each weak learner. Despite the feature space dimensionality reduction from 27 to 18 features, our methodology preserved the recognition accuracy of about 85.0%.

Index Terms— Correlation analysis, emotional experience, fusion algorithm, physiological signals, weak learners.

1. INTRODUCTION

1.1. Study Motivation

Providing computers with emotional understanding of their users along with their current mathematical-logical capabilities is considered a breakthrough in creating more intelligent and less exacerbating behaviors for machines in Human-Computer Interaction (HCI) applications [1]. An example of an intelligent HCI is exploiting "feeling computers" in enhancing distance-education experience. In [2], a facial recognition software was introduced to detect specific feelings of students such as frustration and boredom during training sessions. Among the difficult challenges in these platforms is the ambiguity in recognizing emotions only from studying the taxonomy of facial behaviors. This ambiguity is partly due to the unrecognizable facial deformations such as wrinkling of the forehead which can express in different emotions. In addition, facial expressions can easily be manipulated, which produce faked affect signs.

Unlike the facial expressions, physiological signals can not be manipulated. This characteristic makes them a robust alternative for emotion recognition. Knowledge of the natural processes that occur at different scales inside our body can be obtained by exploring different physiological signals and by drawing conclusions about how these biological processes are triggered, executed and connected between each other. In particular, the aim of this paper is to explore the correlation between physiological changes and emotional experiences in order to develop a robust emotion recognition inference model. As authors concluded in [3], much work remains before emotion interpretation by machine intelligence can occur at the level of human abilities. When it comes to the implementation of the emotional understanding in machines with high-constrained computational resources, a simple but efficient knowledge of the key features that trigger and characterize human emotions could be game changing.

1.2. Related Works

Automatic emotion recognition is a field that has gained a lot of attention in the past few decades [4, 5, 6]. Much of the work explores diverse patterns drawn from physiological signals to train and test several supervised and unsupervised methods [7, 8, 9]. Recent studies show that autonomic affective regulation in two direction of arousal and valance is indexed by bodily signals such as skin conductance, respiration rate, and cardiac variables [10]. Also, there are strong evidences that physiological activity associated with psychology or mental states can be distinguished and systematically organized [11]. For example, electrocardiovascular, blood volume pressure (BVP), Galvanic skin response (GSR) and electromyogram (EMG) activities have been used to examine the dimension of pleasure, or valence (i.e, positive and negative affect) of human subjects [12, 13, 14]. However, a deeper understanding is needed to completely describe the relation between human emotional experience and each source

Corresponding author's e-mail: ostadabbas@ece.neu.edu

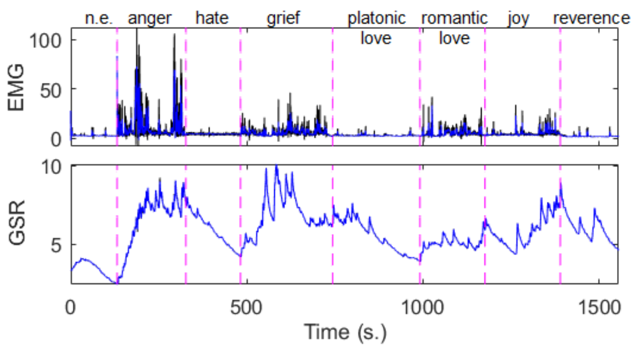

Fig. 1. Samples of electromyogram (EMG) and galvanic skin response (GSR) for 8 emotional states obtained from [16].

of biosignals [15]. Another interesting direction to explore within automatic emotion recognition is the optimal feature selection, reduction, or transformation methods to cost efficiently (in terms of power, speed, storage, etc.) exploit the related information content of human PHY signals [1].

1.3. Our Contribution

The present work extracts proper attributes from three physiological modalities (EMG, BVP, GSR), and feeds them to an innovative fusion-based classifier which decodes the correlation between the bodily signal expressions and different affective categories. Fig. 1 shows a sample of EMG and GSR obtained from [16], which were recorded over 20 minutes for eight emotional experiences. Our automatic emotional experience decoding approach is based on the fusion of three weak learners that are built upon features extracted from three specific physiological modalities using a linear discriminant analysis classifier (LDA). Then, a fusion algorithm consolidates the prediction weights from each weak learner. Finally, a feature selection approach based on correlation analysis reduces the complexity of the algorithm while keeping the classification performance within an acceptable range. The contribution of this paper capitalizes the effect of a fusion algorithm, which extracts highly relevant attributes from each modality, and fuses the prediction outputs rather than mixing all of the data into a single classifier. Moreover, our emotion decoding system provides lower computational complexity by reducing the dimensionality of each linear weak learner through discarding highly correlated features.

2. METHODOLOGY

Fig. 2 presents the general algorithm for the proposed automatic emotional experience recognition method, where given 19-day observations (training set), the objective is to recognize the corresponding emotions from the remaining 1-day observation (test set). A 20-fold with leave one fold out cross validation was performed to reduce the observation-dependent predication error. The algorithm implementation was executed in MATLAB R2015b.

2.1. Database Description

Our algorithm was evaluated using PHY data obtained from the database provided in [16], which contains recordings of three PHY signals: Electromyogram (EMG), Blood Volume Pressure (BVP), and Galvanic Skin Response (GSR), during eight emotional states: (1) Baseline-No emotion (N.E.), (2) Anger, (3) Hate, (4) Grief, (5) Platonic Love, (6) Romantic Love, (7) Joy, and (8) Reverence. We will use these numbers as the emotion's IDs throughout this paper. In the experiment, bodily signals were recorded at a sampling rate of 20Hz for a 25-minute time period over 20 days (one observation recorded per day) from a healthy female subject [3].

2.2. Signal Pre-processing

Each physiological signal was filtered using a first order low-pass Butterworth filter with cutoff frequency of 10Hz for EMG, and 19Hz for BVP and GSR accordingly to the criteria explained in [3]. Then, the signals were smoothed by computing the average of the upper and lower envelope of their corresponding filtered versions.

In order to compensate the nonlinear phase distortion introduced by the Butterworth filter, specially around the cut-off frequencies, the original Butterworth coefficients were applied to the signal using a zero-phase digital filter known as *filtfilt* in MATLAB. *filtfilt* process the input data in both the forward and reverse directions: after filtering the data in the forward direction, it reverses the filtered sequence and runs it back through the filter [17]. By this mechanism, the output signal achieves the desired zero-phase behavior.

Finally, a min-max normalization was performed across the 20-day measurements to avoid the effects of human initial statics in the resulting feature space. This step is highly recommended for a robust performance of the classifier.

2.3. Feature Extraction

According to [3, 18, 19, 20, 21, 22], nine features were extracted for each physiological signal to capture time, frequency, statistical and spectral relevant characteristics of the signals, including: (1) max value, (2) min value, (3) number of peaks, (4) mean: first statistical moment, (5) variance: second statistical moment, (6) kurtosis: forth statistical moment, (7) entropy, (8) signal power, and (9) signal spectral power.

2.4. Classifier Design

Our proposed emotional experience decoding approach is based on fusion of specialized weak learners for each of the three physiological signals under study (EMG, BVP, GSR)

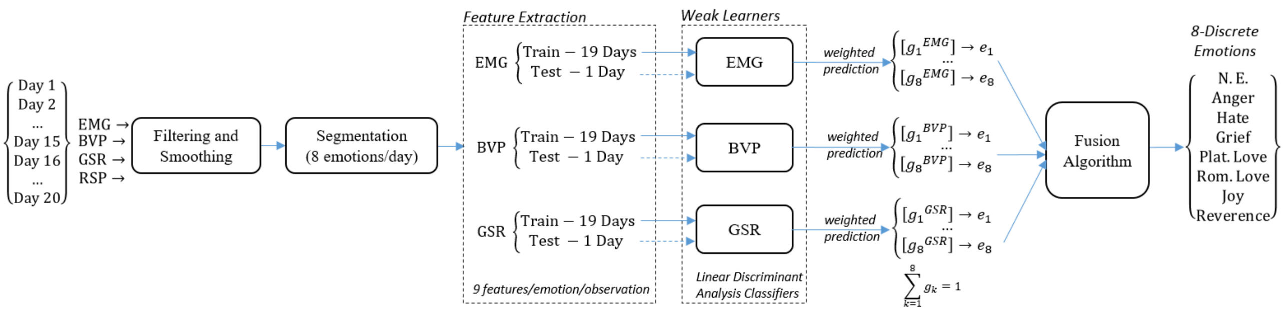

Fig. 2. Automatic emotional experience decoding algorithm based on fusion of specialized weak learners.

as shown in Fig. 2. A linear discriminant analysis (LDA) classifier is implemented as the weak learner due to its computational simplicity and reasonable robustness, even when the classes do not behave as normal distributions [23, 24]. The aim of a linear discriminant classifier is to find decision rules $g_i(x)$ in terms of the minimum total error of classification and a monotonic transformation of the posterior probabilities $P(e_i|x)$:

$$g_i(x) = \ln P(e_i|x) \quad \text{for} \quad i = 1, \cdots, 8. \tag{1}$$

where each of the eight emotion categories is considered as a target emotional experience class e_i, and x is the set of given features from the specific physiological signal. Assuming that each class has multivariate normal distribution and all classes have different mean values μ_i, but equal covariance matrix Σ, by Bayes' theorem, the joint posterior probability $g_i(x)$ for the eight emotions can be written as a linear system:

$$g_i(x) = W_{io} + W_i^T x \tag{2}$$

where $W_i = \Sigma^{-1}\mu_i$ and $W_{io} = -\frac{1}{2}\mu_i{}^T\Sigma^{-1}\mu_i + \ln P(e_i)$. Note that for a single LDA classifier, x belongs to emotion class e_i if $g_i(x) > g_j(x), \forall i \neq j$. Also, $P(e_i)$ implies the prior information obtained for each emotion class through the feature extraction of the corresponding training set.

However, in our classification design, rather than concluding the emotion class e_i from each LDA-based weak learner, the output from each of the three weak learners is expressed as a weighted prediction vector of the eight emotions $[g_1^j, \cdots, g_8^j]$ that the jth ($j = 1, 2, 3$) physiological input signal is likely to be constituted of. In turn, the three weighted predictions from the three weak learners are consolidated in a robust inference model by means of a fusion algorithm.

Specifically, the fusion algorithm combines the individual weighted predictions from the three weak learners in a likelihood weight prediction matrix G, whose dimension is 3×8 due to the 3 physiological signals as inputs and the 8 discrete emotional experience categories as the classification outputs. Then, in order to consolidate all three independent decisions, the mean vector of G across physiological signals is computed to obtain $[\tilde{g}_1, \cdots, \tilde{g}_8]$. From here, a discrete emotion classification is obtained by selecting the emotion class e_k for which the maximum weight $\tilde{g}_k$ is obtained. Within this concept, the idea of prediction weights can be associated with the probability of decoding 8 discrete emotions from a given multimodal physiological input.

3. RESULTS

To demonstrate the performance of our emotion recognition approach, we compared its behaviour against a classical support vector machine (SVM) classifier implemented with a linear, radial basis, and polynomial kernel functions. Table 1 presents the results obtained from a 20-fold leave one fold out cross validation, where the best accuracy of 72.5% for the *SVM classifiers* was obtained through a radial basis kernel function, and for our proposed *weak learner fusion* algorithm, the accuracy summed up to 88.1%.

Moreover, a variant of the proposed weak learners algorithm was implemented to reduce the feature space by means of a *feature correlation analysis*. This approach does not intend to increase the computational complexity as when transforming the feature space to other dimensions (e.g., principal component analysis (PCA)). In our implementation, feature correlation analysis computes the correlation between each pair of features in the whole set prior passing relevant features to each weak learner, and drops those that are correlated with a factor greater than a predefined threshold, which we set to 0.8. As a result, from a feature space of 27 features per observation per emotion, the reduced feature space resulted

Table 1. Recognition performance of different classifiers

Classifier	**Accuracy (%)**
SVM - linear kernel	70.0
SVM - radial basis kernel	72.5
SVM - 1^{st} order polynomial kernel	68.1
SVM - 2^{nd} order polynomial kernel	70.0
SVM - 3^{rd} order polynomial kernel	71.9
SVM - 4^{th} order polynomial kernel	69.4
SVM - 5^{th} order polynomial kernel	67.5
Weak learners	**88.1**
Weak learners & correlation analysis	**85.0**

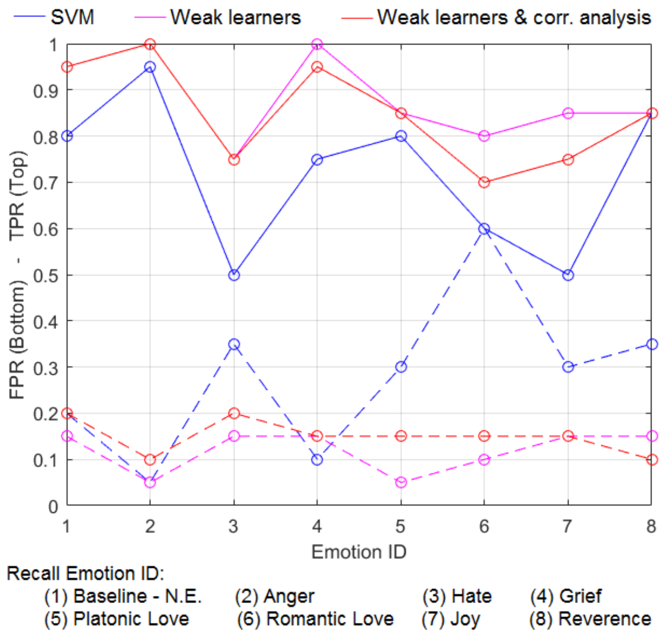

Fig. 3. True Positive Rate – TPR (solid lines) and False Positive Rate – FPR (dashed lines) curves for the 8 target emotion experience categories.

in 18 uncorrelated features. Note that 27 features stand for the 9 features per each of the 3 physiological signals under study, and all those 27 features need to be in turn computed for each of the 20 day observations during the 8 emotion segments. Under this scenario, feature reduction from 27 to 18 features per observation per emotion represents a significant computational complexity reduction.

Within this context, the overall accuracy was preserved to 85.0%. Note that neither PCA nor LDA provide information regarding what exact features can be removed from the feature space to reduce complexity. The new feature space provided by these methods still considers a linear combination of all the original features.

3.1. Emotion Decoding

Fig. 3 shows a variant of the receiver operating characteristic (ROC) curve, which is specialized to demonstrate true positive rate (TPR) and false positive rate (FPR) of a multi-class classifier. TPR represents the successful classification rate of a given class (emotion), and FPR represents the corresponding misclassification rate of that class among others. From Fig. 3, it is clear that the weak learner-based classifier with and without correlation analysis classifier outperforms (higher TPR and lower FPR) the SVM approach.

Also, analyzing the trends exposed in Fig. 3, we explored which emotional sources provoked the TPR to fall below 0.9 which were hate, platonic love, romantic love, joy, and reverence. We found out that reverence gets confused with hate, platonic love with grief, hate with both baseline and reverence, joy gets misclassified with romantic love, and romantic love with anger or joy. The negative effect of the correlation-based feature selection method on certain emotions for which the TPR decreases surfaces one of the potential weaknesses of this method, since it completely discards features that are highly correlated without considering the case in which these features may contain useful information for differentiation in other domains as well.

3.2. Discussion

One interesting question that arises is why the LDA classifier outperforms the SVM approach? Even when considering the case of SVM with a linear kernel function (which did not achieved the best behaviour among the SVM group). Since the data showed to be linearly separable, the SVM approach of mapping data into a higher dimensional space in which it would be linearly separable turns out to be redundant. If the data was already observed to be linearly separable, LDA represents a good approach towards classification without overfitting the classification model [25]. The strength of our model then is mainly due to the fusion of multiple decisions provided by different weak learners each tuned for a specific input modality.

4. CONCLUSION AND FUTURE WORK

Our proposed classification paradigm based on specialized weak learners for each physiological signal modality reported a higher recognition accuracy than the classical SVM approach. Our algorithm also highlights modularity when adding new modalities without compromising the overall complexity. In fact, the proposed fusion algorithm is an alternative boosting approach for consolidating multiple decisions provided by different weak learners in a strong inference model. By working with prediction scores as outputs of each weak learner, we are considering more than one discrete emotion that a single test input is likely to be constituted of; and this flexibility obtained from the prediction scores can be extremely relevant when a final classification decision is made by combining the criteria provided by different learners.

Furthermore, simple methods such as feature correlation analysis can turn out to be very practical tools to analyze redundancy across the feature set. By performing this prior correlation study, a fine-grained feature selection can be passed as input to each weak learner to improve individual classification accuracy. Also, since for each training set the suitable features might be different, the number of selected features for each set changes accordingly. Finally, one of the remaining questions regarding feature selection is how should a penalization term be included in the algorithm to rather than only considering the criteria of redundant information, also contemplate the case of the confusion that overlapped information might cause in the feature space.

5. REFERENCES

[1] Rosalind W Picard, Elias Vyzas, and Jennifer Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 10, pp. 1175–1191, 2001.

[2] Alexandria Katarina Vail, Joseph F Grafsgaard, Joseph B Wiggins, James C Lester, and Kristy Elizabeth Boyer, "Predicting learning and engagement in tutorial dialogue: A personality-based model," *Proceedings of the 16th International Conference on Multimodal Interaction*, pp. 255–262, 2014.

[3] 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.

[4] D. Giakoumis, D. Tzovaras, K. Moustakas, and G. Hassapis, "Automatic recognition of boredom in video games using novel biosignal moment-based features," *IEEE Transactions on Affective Computing*, vol. 2, no. 3, pp. 119–133, July 2011.

[5] 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, pp. 35–45, Nov 2014.

[6] E. Mower, M. J. Mataric, and S. Narayanan, "A framework for automatic human emotion classification using emotion profiles," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 5, pp. 1057–1070, July 2011.

[7] B. Meuleman and K. R. Scherer, "Nonlinear appraisal modeling: An application of machine learning to the study of emotion production," *IEEE Transactions on Affective Computing*, vol. 4, no. 4, pp. 398–411, Oct 2013.

[8] E. H. Jang, B. J. Park, S. H. Kim, M. A. Chung, M. S. Park, and J. H. Sohn, "Emotion classification based on bio-signals emotion recognition using machine learning algorithms," *Information Science, Electronics and Electrical Engineering (ISEEE), 2014 International Conference on*, vol. 3, pp. 1373–1376, April 2014.

[9] W. Wen, G. Liu, N. Cheng, J. Wei, P. Shangguan, and W. Huang, "Emotion recognition based on multi-variant correlation of physiological signals," *IEEE Transactions on Affective Computing*, vol. 5, no. 2, pp. 126–140, April 2014.

[10] J. T. Cacioppo and L. G. Tassinary, "Inferring psychological significance from physiological signals," *American Psychologist*, vol. 45, no. 1, pp. 16–28, 1990.

[11] Margaret M Bradley and Peter J Lang, "Emotion and motivation," *Handbook of psychophysiology*, vol. 2, pp. 602–642, 2000.

[12] J Papillo and D Shapiro, "Principles of psychophysiology: Physical, social and inferential elements, chapter the cardiovascular system," *Cambridge University Press, Cambridge, England*, 1990.

[13] John T Cacioppo and Louise C Hawkley, "Social isolation and health, with an emphasis on underlying mechanisms," *Perspectives in biology and medicine*, vol. 46, no. 3, pp. S39–S52, 2003.

[14] Anna Pecchinenda, "The affective significance of skin conductance activity during a difficult problem-solving task," *Cognition & Emotion*, vol. 10, no. 5, pp. 481–504, 1996.

[15] Karen S Quigley and Lisa Feldman Barrett, "Is there consistency and specificity of autonomic changes during emotional episodes? guidance from the conceptual act theory and psychophysiology," *Biological psychology*, vol. 98, pp. 82–94, 2014.

[16] Jennifer Healey and Rosalind W. Picard, "Eight-emotion sentics data," `http://affect.media.mit.edu/share-data.php`.

[17] MathWorks Inc., "Matlab documentation: filtfilt," `http://www.mathworks.com/help/signal/ref/filtfilt.html`.

[18] Egon L Van Den Broek, Viliam Lisỳ, Joris H Janssen, Joyce HDM Westerink, Marleen H Schut, and Kees Tuinenbreijer, "Affective man-machine interface: unveiling human emotions through biosignals," *Biomedical Engineering Systems and Technologies*, pp. 21–47, 2009.

[19] Gyanendra K Verma and Uma Shanker Tiwary, "Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals," *NeuroImage*, vol. 102, pp. 162–172, 2014.

[20] Cornelia Setz, Johannes Schumm, Claudia Lorenz, Bert Arnrich, and Gerhard Tröster, "Using ensemble classifier systems for handling missing data in emotion recognition from physiology: one step towards a practical system," *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pp. 1–8, 2009.

[21] Guillaume Chanel, Julien Kronegg, Didier Grandjean, and Thierry Pun, "Emotion assessment: Arousal evaluation using eegs and peripheral physiological signals," *Multimedia content representation, classification and security*, pp. 530–537, 2006.

[22] M. Soleymani, M. Pantic, and T. Pun, "Multimodal emotion recognition in response to videos," *IEEE Transactions on Affective Computing*, vol. 3, no. 2, pp. 211–223, April 2012.

[23] Ludmila Kuncheva, "Pattern classifiers methods and algorithms," pp. 45–46, 2004.

[24] Richard Duda, Peter Hart, and David Stork, "Pattern classification, 2nd edition," pp. 215–218, 2001.

[25] K. P. Lin and M. S. Chen, "On the design and analysis of the privacy-preserving svm classifier," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 11, pp. 1704–1717, Nov 2011.