Emotion Analysis using Thermal Images based on Kernel Eigen Spaces

Ajaya AR
Department of CSE
B M II College of Engineering
Sasthamcotta, Kollam

P. Petchimuthu
Department of CSE
SCAD College of Engineering
Cheranmahadevi, Tirunelveli

Kavitha VK
Department of CSE
BM II College of Engineering
Sasthamcotta, kollam

ABSTRACT
Emotion recognition using facial expression has become an active research topic in recent years. In this paper we present an efficient method for emotion recognition, which has better performance over previous art of works. This work proposes an efficient attempt to investigate the suitability and sensitivity of the thermal imaging technique to detect specific muscles heat patterns and there by predicting the emotions. In this work, feature extraction is carried out by Kernel PCM and emotion classification is performed using Multi Class SVM. Thermal imaging is used for the investigation of Action Unit (AU) productions. A facial AU represents the contraction of a specific muscle or a combination of muscles, and earlier research had demonstrated that such muscle contraction induces an increase in skin temperature. For this reason, thermal imaging analysis might be well suited to detect AU production and there by predicting the emotional state of a person. We used a multi class SVM approach to classify nine different AUs or combinations of AUs and to differentiate their speed and strength of contraction. The Multi class SVM classifier gives promising results for the emotion classification process

General Terms
Classification, Algorithms, Learning, Features, Emotions

Key words
Emotion Recognition, Action Unit, Thermal Imaging, Eigen Faces, Kernel PCA

1. INTRODUCTION
Everyday life is awash in frustration, whether it is derived from dealing with bureaucracies or driving on a busy road. Frustration is caused by the occurrence of an unexpected result from an event .Unfortunately, computers are also a significant source of frustration and the automated detection of a user's emotional state has been the focus of considerable study. Early detection of a user's emotional state can be used for better management of computers and software. A computer and its software can respond to an individual's changing psychophysiology by either eliciting feedback from the user as to what is causing the frustration, or by offering assistance and guidance. Emotion recognition has potential applications in E-learning, Robotics and Human Computer interactions.

The face is a rich source of information about human behavior. Facial expressions are complex muscular patterns that carry complex social signals. A facial expression results from one or more motions or positions of the muscles of the face Fig. 1 shows different muscles located on the head and neck. By analyzing these muscle contractions we can predict the emotional state of the particular person. This art of emotion recognition by trained system comes under the area, Affective computing. Affective computing includes the study and development of systems and devices that recognize interpret process and simulate human emotions. The ability of humans to recognize a wide variety of facial expressions is unparalleled. Researchers in the recent past have been trying to automate this task on a computer, employing a combination of image/ video processing techniques, along with machine learning techniques like ANNs

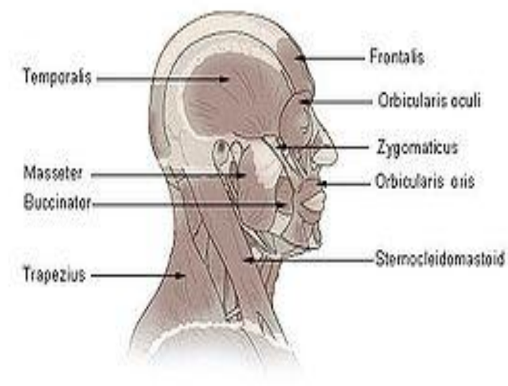

Fig.1 Muscles on the Face and Head

2. ON THE WAY TO THERMAL IMAGING
Ekman et al. [1] have developed the most popular standard system to classify the physical aspects of facial expressions: the FACS. This system is based on the anatomy of the facial muscles and is composed of action units (AUs) that describe all visible facial movements at different intensities (fig.2). Although very informative, this process of coding each action unit in a facial expression is very time-consuming. Several attempts have been made to automatize the coding though. For instance, Lien et al. [3] used video-taped images under visible spectrum lighting to automatically detect, track, and classify the AUs implied in the expressions. Unfortunately, the influence of lighting on image\ quality (contrast fluctuations or low light) limits this visible spectrum imagery technique. To circumvent this problem, researchers can directly record the electrical activity of muscles that subtend the AUs by means of facial EMG, which measures muscle contraction (even the visually imperceptible). This technique is particularly sensitive to measure the kinetics and intensity of that muscular contraction [4]. However, EMG recording is not without drawbacks: 1) It can be difficult to record the precise activity of a specific muscle involved in a given AU because of the diffusion of electrical activity from one muscle to

another (i.e., cross-talk phenomenon); 2) electrodes must be fixed on many areas of the face, a constraint that could hamper natural muscular contraction; and 3) theoretically, there should be as many electrodes as there are different muscles related to the AUs. This last point constitutes a severe limitation for the use of EMG as a non-invasive method. However, to date no technique has been developed that would allow the simultaneous recording of all facial muscle activity, being sensitive to the intensity and the temporal dynamics of the contractions, and without hindering natural AU production or facing light problems.

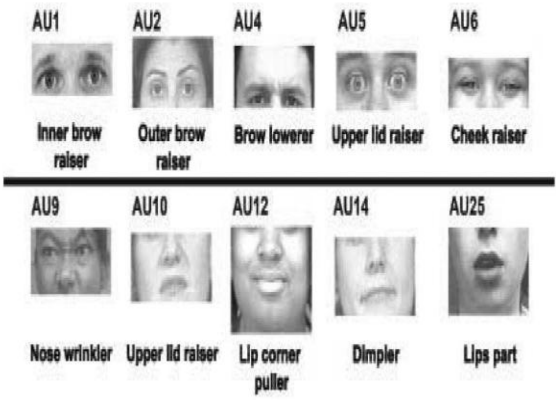

Fig.2 Action Units produced by FACS Coders

Thermal imaging [5] is a promising technique for the investigation of facial muscle contractions due to emotions. Thermal imaging can track dynamic patterns in facial temperature at any distance with high temporal and thermal resolutions. By rendering infrared radiation as visible light it avoids the illumination problem. Thermal images shows infrared output differentials, so two objects with the same temperature will appear to be the same color. A facial AU represents the contraction of a specific muscle or a combination of muscles, and research has demonstrated that such muscle contraction induces an increase in skin temperature. For this reason, thermal imaging analyses might be well suited to detect AU production. AU production analysis using thermal imaging [9] makes use of PCM for feature extraction but it has two major disadvantages as follows. PCA will not be able to capture the relationship among more than two variables. PCA cannot represent the variations caused by illuminations and expressions properly. So in this work our main focus is to eliminate these two disadvantages of PCM and thereby improve the efficiency of emotion prediction.

3. SYSTEM DECOMPOSITION

In this work, we have used Kernel PCA for feature extraction. Here the data is projected into nonlinear higher dimensional space using the kernel method. This enables to capture the nonlinear relationships among the pixels within the modules. For the emotion classification purpose Multiclass SVM is used. The basic SVM supports only binary classification, but Multi Class SVM supports multiple classifications. For classifying emotions Multi Class SVM is proposed in this work. This project work consists of mainly four modules.

3. 1 Image Preprocessing

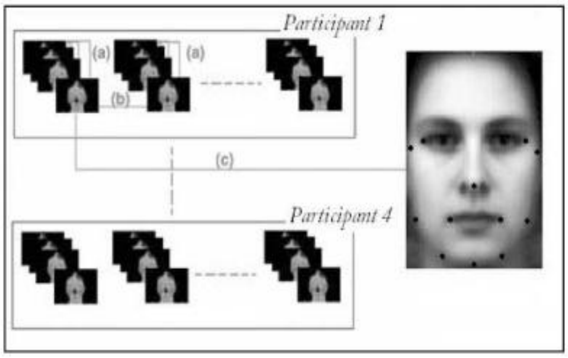

Fig.3 Face Normalization

All recorded images were reduced to the face area and rescaled to a particular size (210 pixels high _ 150 pixels wide) by means of a bilinear interpolation algorithm to optimize speed calculation and disk storage, using Matlab (Matlab, Release 14, The Mathworks, Inc.). First, a rigid 2D translation and rotation procedure was applied to each trial (baseline included) to align the images composing it. Images were realigned with each other using a Matlab's basic optimization routine (fminsearch) to find the transformation that restores the original image shape. The same procedure was then applied subsequently to align the trials within individual participants. For each participant, the algorithm determines an affine transformation that matched 12 control points placed on the individual's face (first image of participant's first trial) with those placed on the average face in the Karolinska Directed Emotional Faces database[6] (Fig. 3). All facial images of a given participant were then spatially normalized to the average face according to this affine transformation. Examples of the results obtained with this thermal image transformation technique are presented in Fig. 4.

Fig.4 Thermal Image Normalization

3.2 Facial Feature Extraction based on Kernel Eigen Spaces

Facial feature extraction is carried out by means of an algorithm called Kernel PCA [8]. In this algorithm the data is projected into nonlinear higher dimensional space using the kernel method. This enables to capture the nonlinear relationships among the pixels within the modules. Using the kernel PCA one can compute the higher order statistics using only dot products of the input patterns. Kernel PCA has been applied to face recognition applications and is observed to be able to extract nonlinear features.

PCA encodes the pattern information based on second order dependencies, i.e., pixel wise covariance among the pixels, and are insensitive to the dependencies of multiple (more than two) pixels in the patterns. Since the eigenvectors in PCA are the ortho-normal bases, the principal components are uncorrelated. In other words, the coefficients for one of the axes cannot be linearly represented from the coefficients of the other axes. Higher order dependencies in an image include nonlinear relations among the pixel intensity values, such as the relationships among three or more pixels in an edge or a curve, which can capture important information for recognition. Explicitly mapping the vectors in input space into higher dimensional space is computationally intensive. Using the kernel trick one can compute the higher order statistics using only dot products of the input patterns.Fig.5 shows the average face extracted from the training set. Kernel PCA has been applied to face recognition applications and is observed to be able to extract nonlinear features.

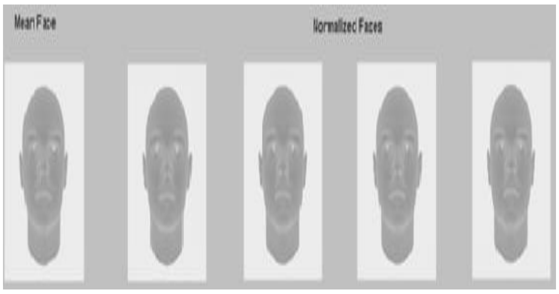

Fig.5 Mean face Calculation

The process of obtaining the weights for the input patterns in the kernel principal component analysis transformed space is described below. Let xi be are the vectors belonging to the training sample set **X={x1, x2, x3…..xm}** where **xi €R^{n}**' and **c** be the number of classes in the training set. The data in this case is assumed to be centered.

That is the mean of the data **m0** is given by

$$m_0 = \frac{1}{m}\sum_{k=1}^{m} x_k = 0.$$

The Covariance Matrix **C** is given by

$$C = \frac{1}{m}\sum_{j=1}^{m} x_j x_j^T .$$

The eigenvectors corresponding to the nonzero eigenvalues of the covariance matrix are calculated as

$$CV = \lambda V.$$

The matrix V represents the eigenvectors and represents the corresponding eigenvalues of the matrix. The eigenvectors arranged in the descending order of the eigenvalues correspond to the variations in the data in that order in those directions. Let be the mapping between the input space and the feature space. It is possible that the feature space can be infinite dimensional. Hence, can be assumed to be a Hilbert space. The covariance matrix in the feature space is given by

$$C_\phi = \frac{1}{m}\sum_{j=1}^{m} \phi(x_j)\phi(x_j)^T$$
$$\phi : X \rightarrow H.$$

It is assumed that the data in the feature space is also centralized, that is, Similar to principal component analysis, and are the eigenvectors and eigen values of the covariance matrix

$$C_\phi V_\phi = \lambda_\phi V_\phi.$$

The eigenvectors of a covariance matrix which is calculated from a given set of data vectors lie within the span of those vectors. Hence, the eigenvectors can be expressed as the linear combination of the vectors in the data set. The following equation gives the relationship between the eigenvector and the sample training vectors in the feature space is given by

$$C_\phi = \frac{1}{m}\sum_{j=1}^{m} \phi(x_j)\phi(x_j)^T$$
$$\phi : X \rightarrow H.$$

The eigenvectors of a covariance matrix which is calculated from a given set of data vectors lie within the span of those vectors. Hence, the eigenvectors can be expressed as the linear combination of the vectors in the data set.

The following equation gives the relationship between the eigenvector and the sample training vectors in the feature space:

$$V_\phi = \sum_{i=1}^{m} \alpha_i \phi(x_i).$$

3.2.1 Apex determination.

The temperature peak (apex) is the moment when the facial temperature is maximal. For each trial, a distribution is calculated averaging the two first component loadings convoluted with the trial thermal values (time points _ pixels) in time. We calculated the apex when the maximum amplitude of this temporal distribution is reached after the baseline period. For each trial, the apex was calculated by performing an automatic baseline-to-peak analysis to find the maximum temperature variations. The latency of the apex gave us information about the speed of temperature changes related to AU production, and the amplitude of the signal at the apex informed us about the intensity of temperature change related to the AU(s) (Fig. 6).

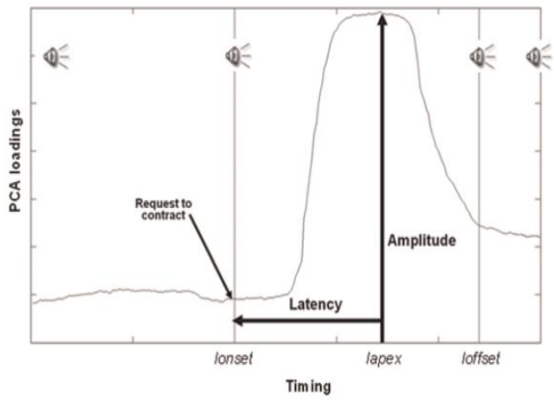

Fig.6 Apex Calculation

3.3 Expression Classification based on Multi class SVM

In this work we use Multi Class SVM [7] for the classification of expressions. The basic SVM supports only binary classification, but Multi Class SVM supports multiple classifications. Support Vector Machines are among the most robust and successful classification algorithms. They are based upon the idea of maximizing the margin i.e. maximizing the minimum distance from the separating hyperplane to the nearest example.

SVMs as binary classifiers have drawn much attention because of their high classification performance and thorough mathematical foundations rooted in statistical learning theory. They are here considered in a normalized feature space [6] by using modified kernel

Functions $\frac{K(\vec{x},\vec{y})}{\sqrt{K(\vec{x},\vec{x})K(\vec{y},\vec{y})}}$, the latter still satisfying Mercer's theorem. The data in such a space lies on a unit hypersphere. A modification in the SV algorithm taking advantage of this geometrical property is proposed in [6] where the offset of the optimal separating hyperplane (OSH) is shifted. In the standard SVM algorithm, the OSH is placed in the middle of the margins. In a normalized feature space, however, the OSH is chosen so as to separate equidistantly the unit hypersphere in between the margins as shown in figure 1. In other words, the distance metric is projected on the unit hypersphere and the correction of the offset can be written as

$$\frac{\hat{b}}{||\vec{w}||} = \cos\left(\frac{\arccos(\frac{b-1}{||\vec{w}||}) + \arccos(\frac{b+1}{||\vec{w}||})}{2}\right)$$

Allwein et al. generalized the idea of ECOC to apply general coding techniques, where the coding matrix M is allowed to take values $\{-1, 0, +1\}$. The value of +1 in the entry M (k, n) means that examples belonging to class k are considered as positive examples to classifier n. A value of -1 denotes that these examples are considered negative examples. The OVA approach has a matrix M such that each column contains exactly one +1 value with the rest filled with -1. The AVA scheme has columns with exactly one +1 value and one -1 value with the rest set to zero's. The ECOC has the matrix M filled with +1 and -1 values. When testing an unknown example, the codeword "closest" to the output corresponding to that example is chosen as the class label. There is the usual hamming distance between the two code words. In Multi Class SVM [4], additional parameters and constraints are added to handle the separation of the different classes. Fig.7 shows the emotion classification task/

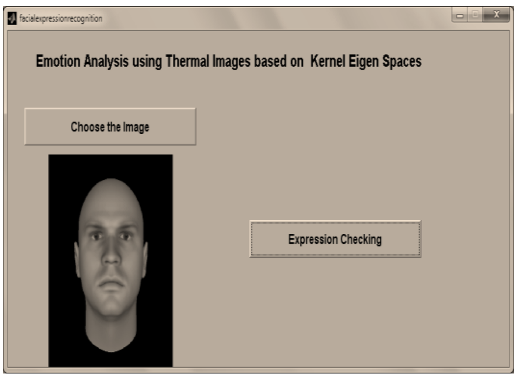

Fig.7 Emotion Classification

4. CONCLUSION AND FUTURE ENHANCEMENT

The use of HMM and dense-flow extraction has improved the performance in visible light emotion analysis over previous state-of-the-art methods. Thus, the work presented here focuses mainly on emotion recognition by avoiding the illumination problem, which is a major drawback of the previous art methods. For this purpose we have presented an efficient method for predicting the emotions of a particular person by analyzing the facial muscle contractions .All the images used in this work are in the thermal mode, which will avoid the illumination problem.

In this work we used a kernel PCA based approach called Kernel Eigen Spaces for performing the feature extraction. The paper adopts Multi Class SVM for the classification of emotions, which is a robust and successful algorithm and provides better accuracy in prediction. Since the work is based on the thermal imaging technique it avoids the visible light illumination problem which exists in the previous techniques. Moreover, in contrast to EMG recordings, this technique is Noninvasive (no electrodes on the face).

Thermal imaging, which has recently been used in domains such as security, firefighters, and military and medical diagnosis, may be a promising alternative for the investigation of AU[9] production. A facial AU represents the contraction of a specific muscle or a combination of muscles, and research has demonstrated that such muscle contraction induces an increase in skin temperature [5]. For this reason, thermal imaging analyses might be well suited to detect AU production and there by predicting the emotional state of a person.

Finally, thermography may prove to be a useful tool to unobtrusively analyze fine-grained elements of facial expressions. The ideas of hand gesture and prosodic information [10] are also useful and can be incorporated to further improve the performance of the emotion prediction.

5. ACKNOWLEDGEMENTS

We would like to thank Satyanadh Gundimada and Vijayan K Asari for their discussion about Kernel PCA also Sophie Jarlier for her valuable discussion about thermal Imaging

6. REFERENCES

[1] P. Ekman, W.V. Friesen, and J.C. Hager, Facial Action Coding System. Consulting Psychologist Press, 1978.

[2] J.J.J. Lien, T. Kanade, C.C. Li, and J.F. Cohn, "Detection, Tracking and Classification of Action Units in Facial Expression,"IEEE J. Robotics and Autonomous Systems, special issue: face expression in human-robot interaction systems, vol. 31, pp. 131-146, 2000.

[3] S. Delplanque, D. Grandjean, C. Chrea, L. Aymard, I. Cayeux, C. Margot, M.I. Velazco, D. Sander, and K.R. Scherer, "Sequential Unfolding of Novelty and Pleasantness Appraisals of Odors: Automatic Reactions, Emotion, vol. 9, no. 3, pp. 316-328, 2009.

[4] Z. Zhu, J. Fei, and I. Pavlidis, "Tracking Human Breath in Infrared Imaging," Proc. Fifth IEEE Symp. Bioinformatics and Bioeng., pp. 227-231, 2005

[5] J. Gonza´lez-Alonso, B. Quistorff, P. Krustrup, J. Bangsbo, and B. Saltin, "Heat Production in Human

Skeletal Muscle at the Onset of Intense Dynamic Exercise," J. Physiology, vol. 524, pp. 603-615, 2000.

[6] D.Lundqvist and J.E. Litton, The Averaged Karolinska Directed Emotional Faces (AKDEF) Dept. of Clinical Neuroscience, Psychology Section, Karolinska Inst., 1998.

[7] Irene Kotsiay, Stefanos Zafeiriouy, Nikolaos Nikolaidisy and Ioannis Pitasy, Multiclass Support Vector Machines and Metric Multidimensional Scaling for Facial Expression Recognition, Aristotle University of Thessaloniki, Department of Informatics Thessaloniki, Greece, 2009

[8] Satyanadh Gundimada and Vijayan K. Asari, Facial Recognition Using Multisensor Image Based on Localized Kernel Eigen Spaces, IEEE Transactions On Image Processing, Vol. 18, No. 6, June 2009

[9] Sophie Jarlier, Didier Grandjean, Sylvain Delplanque, Karim N'Diaye, Isabelle Cayeux, Maria Ine´sVelazco, David Sander, Patrik Vuilleumier, and Klaus R. Scherer, Thermal Analysis of Facial Muscles Contractions, IEEE transactions on Affective Computing, vol. 2, no. 1, January-March 2011

[10] Chung-Hsien Wu, Senior Member, IEEE, and Wei-Bin Liang, Emotion Recognition of Affective Speech Based on Multiple Classifiers Using Acoustic-Prosodic Information and SemanticLabels,IEEE Transactions On Affective Computing,Vol. 2, No. 1, January- March 2011

[11] Xin Chen, Patrick J.Flynn, Kevin W.Bowyer, "IR and Visible light face Recognition", University of NotreDame, USA, http://www.identix.com/products/.

[12] Bai-Ling Zhang, Haihong Zhang, and Shuzhi Ge, "Face Recognition by Applying Wavelet Subband Representation and Kernel Associative Memory", IEEE Transaction on Neural Networks.