International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:6, No:3, 2012

Emotion Classification using Adaptive SVMs

P. Visutsak

***Abstract*—**The study of the interaction between humans and computers has been emerging during the last few years. This interaction will be more powerful if computers are able to perceive and respond to human nonverbal communication such as emotions. In this study, we present the image-based approach to emotion classification through lower facial expression. We employ a set of feature points in the lower face image according to the particular face model used and consider their motion across each emotive expression of images. The vector of displacements of all feature points input to the Adaptive Support Vector Machines (A-SVMs) classifier that classify it into seven basic emotions scheme, namely neutral, angry, disgust, fear, happy, sad and surprise. The system was tested on the Japanese Female Facial Expression (JAFFE) dataset of frontal view facial expressions [7]. Our experiments on emotion classification through lower facial expressions demonstrate the robustness of Adaptive SVM classifier and verify the high efficiency of our approach.

***Keywords*—**emotion classification, facial expression, adaptive support vector machines, facial expression classifier.

I. Introduction

A major role in human interaction is non verbal and among these nonverbal cues a large portion is in the form of facial expressions. Merhrabin, a psychologist has claimed that verbal languages convey only a 7% portion of the message in face-to-face communication between humans, whereas more than 55% portion is conveyed by the nonverbal method such as facial expression [9]. Facial expression is the natural means for human beings to show their emotions and motivations. The ability to recognize and understand facial expression has become a new challenge of a human-computer interaction research. Pantic and Rothkrantz clearly defined three basic problems, related to facial expression analysis approach need to accomplish: detection of an image segment as a face, extraction of the facial expression information, and classification of the expression according to emotion categories [11]. In the past decade, many of researches have been done on machine recognition of human facial expressions. Convention approaches extracted features of facial components, such as eyes, nose, and mouth, in gray or color images of frontal view faces. Properties and relations (e.g., areas, distances, angles) between the features were used as descriptors of faces for recognizing the expressions from their changing properties or their geometrical relationships by different particular facial expressions [9].

P. Visutsak is with the Computer Science Department, Faculty of Liberal Arts and Science, Kasetsart University, Nakhon Pathom, 73140 THAILAND, (e-mail: faaspwv@ku.ac.th).

Advantages included economy and efficiency in achieving data reduction and insensitivity to variations in illumination and viewpoint. However, the estimation of their precise properties in real images was difficult and facial features extraction techniques developed to date have not been reliable enough for the facial expression recognition.

More work has been made due to advances in the appearance-based and learning-based approaches. Viola and Jones used the Adaboost algorithm to scan an image by passing multiple scales sub-window for rapid face detection [13]. Littlewort used a bank of forty Gabor wavelet filters at different scales and orientations to perform a data extraction, then generated image sequences image-by-image to train two stages of support vector machines from Gabor filter jets [6]. Essa and Pentland performed spatial and temporal filtering together to extract motion blobs from image sequences, and evaluated them using the eigenfaces method [5]. They extracted positions of prominent facial features using eigen features and Principal Component Analysis (PCA), then calculated ideal motion energy templates for each expression category. Ekman and Friesen developed the Facial Action Coding System (FACs) to code expressions as a combination of forty four Action Units, and defined six basic emotions: angry, disgust, fear, happy, sad, and surprise [4]. Cohn used displacement vectors of facial feature points to represent the extracted expression information, applied separate discriminant functions together with variance-covariance matrices to different facial regions, and used feature displacements as predictors for classification [1]. Dailey used a six unit, a single layer neural network to classify into six basic emotion categories given Gabor jets extracted from static images [2]. Philipp and Rana proposed a method for automatically inferring emotions by recognizing facial expressions in live video, by using a real time facial feature tracker; they could deal with problems of face localization and a feature extraction in spontaneous expressions [10]. A set of displacements from feature motion in the video stream, gathered by a face feature tracker, were used to train the Support Vector Machines (SVMs) classifier to recognize previously unseen expressions. The extended version of the emotion recognition using SVMs has been proposed by Porawat [14], this method used the lower feature of facial points to train the dataset based on six basic emotions scheme.

In this paper, we propose the static approach to emotion classification through facial expression. We concentrate our method on the Adaptive Support Vector Machines (A-SVMs), We modify the A-SVMs from the original SVMs, which are the learning technique developed by V. Vapnik and his team at AT&T Bell Labs. [12]. The classification system consists mainly of two consecutive components: a training phase and a testing phase. In the training phase, vectors of displacements according to the motion of feature points in the lower face are used to train the A-SVMs classifier. Unlike the SVMs, The A-SVMs model directly adapts an existing model to the new data, and avoids the overhead of training over the existing data. In the testing phase, the decision function classifies an unseen sample into six basic emotions scheme defined by Ekman (angry, disgust, fear, happy, sad and surprise) [3]. The rest of the paper is organized as follows. The system structure is given in the next section, which describes an outline of an overall implementation of our approach to emotion classification. Experimental results are then presented and discussed. Finally, some conclusions are drawn.

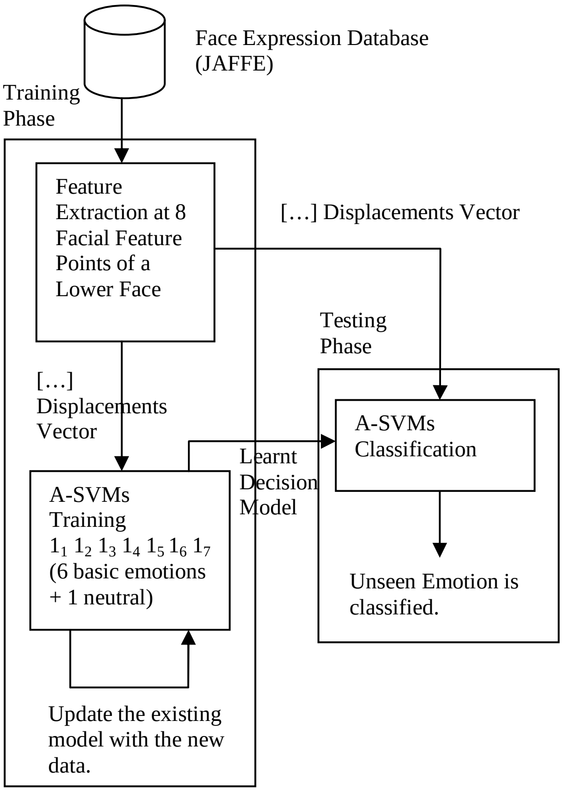


Fig. 1 Block diagram of the system

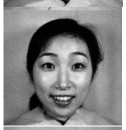
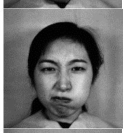
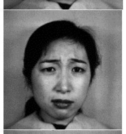
II. System Structure

The system structure is shown in figure 1. The feature extraction on the images is performed in order to gather the numerical A-SVMs input data for the training phase, then training samples are labeled corresponding to emotion classes: neutral, angry, disgust, fear, happy, sad and surprise. By setting up the A-SVMs parameters (e.g., kernel function type), the decision model is then created from the input data. In the testing phase, the extracted information of unseen example image is classified according to the decision model.

A. Facial Expression Database

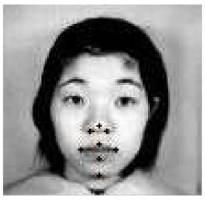
The database used in our experiment is the Japanese Female Facial Expression (JAFFE) database [7]. The database contains 213 images of ten Japanese female models. Each model posed 3 or 4 examples for each of 6 basic facial expressions and 1 neutral. All images are roughly frontal view, under well controlled lighting condition. Original image are grayscale, with resolution of 256x256 pixels in tiff format. The useful information attached with this database is semantic rating from psychological experiments using the images, averaged over sixty Japanese female students. A five level scale (5-high, 1-low) is rated for each of six basic emotional adjectives. This semantic rating is used in our experiment for evaluating the state of each expression. Some sample images and corresponding semantic ratings are shown in table 1.

TABLE I
Example Images And Corresponding Semantic Ratings

Class Label	Expressions Posed	HAP	SAD	SUR	ANG	DIS	FEA
Neutral		3.03	2.45	1.74	2.00	1.90	1.77
Happy		4.84	1.42	1.35	1.23	1.23	1.23
Sad		1.45	4.61	1.52	1.55	2.13	2.39
Surprise		2.35	1.81	4.87	1.81	1.71	2.29
Angry		1.19	2.16	1.55	4.74	2.84	1.39
Disgust		1.35	2.90	2.45	2.65	4.81	3.32
Fear		1.23	4.19	3.74	2.16	3.42	4.26

B. Feature Extraction Method

The markup scheme of a lower facial model is shown in figure 2; eight feature points (fiducial points) of lower face are manually placed in the facial images. By taking the displacements (Euclidean distances) of the feature points between a neutral facial expression and the six particular emotive expressions (angry, disgust, fear, happy, sad, and surprise), it can establish a characteristic motion pattern for each emotion expressed, as shown in figure 3.



1. Nose tip
2. Right nostril
3. Left nostril
4. Upper lip center
5. Right mouth corner
6. Left mouth corner
7. Lower lip center
8. Tip of chin

Fig. 2 The markup scheme of a lower facial model, consists of eight feature points

The data extracted from an example thus consists of the vector of displacements of all feature points. After training a machine learning algorithm on such data, classification essentially consists of matching a new, unseen motion pattern with the class of a training pattern, it most closely corresponds to.

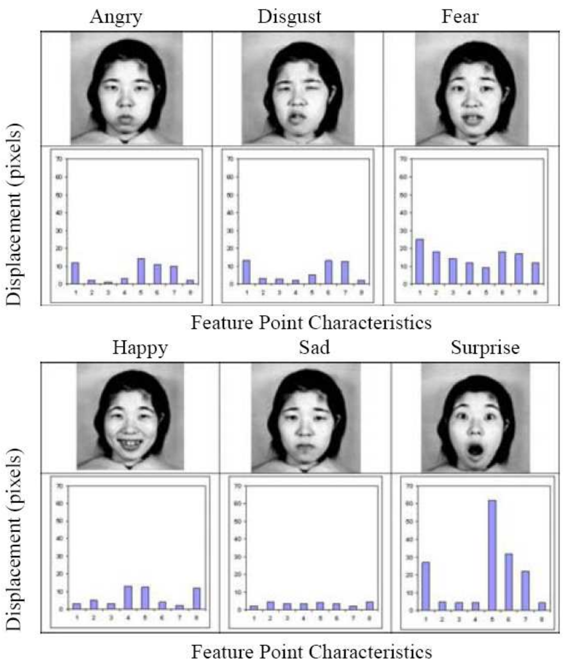


Fig. 3 Eight feature points motion patterns for six basic emotive expressions

C. Facial Expression Classification: the A-SVMs Classifier

In this section, we briefly sketch the A-SVMs algorithm and its motivation. The A-SVMs are the pattern classification algorithms, which receive input data during a training phase, then build a decision model according to the input and generate a function that can be used to predict a future data.

By given a set

$$D = \{(x_i, y_i)\}_{i=1}^{l}$$

of labeled training examples, where

$$y_i \in \{-1, 1\}$$

Learning systems typically try to find a decision function of the form

$$f(x) = sgn([w.x] + b)$$

where *w* is a vector of weights and *b* is called bias, that yields a label ∈ {-1, 1} (for the basic case of binary classification) for an unseen example *x*, which have the smallest generation error.

The A-SVMs technique perform projection the original set of variables *x* in higher dimensional *feature space*:

$$\mathrm{x} \in R^d \rightarrow z(x) \equiv (\emptyset_1(x), \dots, \emptyset_n(x)) \in R^{2n}$$

where a linear algebra and a geometry may be used to separate data that is only separable with nonlinear rules in an input space. By formulating the linear classification problem in the feature space, the solution will have the form

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i\right) + b)$$

Associates with using the kernel functions, allowing an efficient computation of inner products directly in a feature space, by given a nonlinear mapping Φ that embeds input vectors into a feature space, kernels have the form

$$K(x, z) = [\Phi\,(\mathrm{x}).\,\Phi\,(\mathrm{z})]$$

where the α_i are Lagrange multipliers of a dual optimization problem. It is possible to show that only small number coefficients α_i are different from zero, and since every coefficient corresponds to a particular data point, this means that the solution is determined by the data points associated to the non-zero coefficients. These data points, which are called, *support vectors*. These induce sparseness in the solution and give rise to efficient approaches to optimization.

In table II we list some choices of the kernel function proposed by [12], notice how they lead to well known classifiers, whose decision surfaces are known to have good approximation properties.

TABLE II
SOME POSSIBLE KERNEL FUNCTIONS AND THE TYPE OF DECISION SURFACE THEY DEFINE

Kernel Function	Type of Classifier
$K(x,x_i) = \exp(-\|x-x_i\|^2)$	Gaussian RBF
$K(x,x_i) = (x^T x_i+1)^d$	Polynomial of degree *d*
$K(x,x_i) = \tanh(x^T x_i-\Theta)$	Multi Layer Perceptron

III. EXPERIMENTATION RESULTS

In this section, we present and discuss results achieved on our experiment. Five subjects for each basic emotion are used in classifying emotion expressions, two are included in the training set and the other three are included in the test set. In the training phase, the highest two consecutive semantic rating images are used as the training set. Facial features are manually defined for each image and displacements are subsequently extracted from pairs of images consisting of a neural and a representative frame for each basic expression. The other three subjects, used in the testing phase are selected randomly from the rest of nine subjects in JAFFE dataset. The experimental result is shown in table III.

TABLE III
RECOGNITION RATE OF THE A-SVMS CLASSIFICATION ON DISPLACEMENTS EXTRACTED FROM STILL IMAGE (EIGHT FACIAL FEATURE POINTS SCHEME)

Emotion	Percent Accuracy
Angry	67.7%
Disgust	62.3%
Fear	66.7%
Happy	91.5%
Sad	61.0%
Surprise	97.5%
Average	74.5%

By comparing our method to the previous paper, Lyons et al. [8] reported a classification rate of 75% on the same database. Considering that we use only eight facial feature points on the lower face together with the displacements vector corresponding to a particular emotion, it could be said that the performance of our method is comparable to theirs.

IV. CONCLUSION AND FUTURE WORK

Emotion classification is a challenging application for many future human-computer interaction scenarios. This paper describes the emotion classification through lower facial expressions using the A-SVMs**.** The system extracts the motion of eight feature points (fiducial points) of a lower face, as displacements vector, and feed it into the A-SVMs classifier to categorize result expression.

There are two main contributions of this paper. First, we demonstrate that the facial feature extraction scheme enables robust emotion classification in still image. Second, we show that the A-SVMs are the suitable engine for reliable classification. Although the average classification rate of all six basic emotions is 75% approximately, the high percentage of accuracy can be achieved at classifying two distinct emotional expressions (happy and surprise).

Though the combination of feature point displacements and the A-SVMs classification approach is satisfied the requirement of the emotional classification task, more work is need to extend the implementation on another approaches of the feature extraction from face images (e.g. Gabor wavelets, and eigenfaces) and classification methods (e.g. AdaBoost, and Bayes).

ACKNOWLEDGMENT

We would like to thank Dr. Michael Lyons for providing us with the JAFFE database [7].

REFERENCES

[1] J. F. Cohn, A. J. Zlochower, J. J. Lien, and T. Kanade. 1999. Automated face analysis by feature-point tracking has high concurrent validity with manual faces coding. Psychophysiology. 36:35-43.
[2] M. N. Dailey, G. W. Cottrell, and R. Adolphs. 2000. A six-unit network is all you need to discover happiness. In Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society. Mahwah. NJ. USA.
[3] P. Ekman. 1982. Emotion in the Human Face. Cambridge University Press. Cambridge. UK.
[4] P.Ekman and W. Frisen. 1978. Facial Action Coding System (FACS): Manual. Consulting Psychologists Press. Palo Alto. CA. USA.
[5] I. Essa and A. Pentland. 1997. Coding, Analysis, Interpretation and Recognition of Facial Expressions. IEEE Transaction on Pattern Analysis and Machine Intelligence. 19(7): 757-763.
[6] G. Littlewort, I. Fasel, M. Stewart Bartlett, and J. R. Movellan. 2002. Fully Automatic Coding of Basic Expressions from Video. Technical Report 2002.03, UCSD INC MPLab.
[7] M. J. Lyons, Shigeru Akamatsu, Miyuki Kamachi & Jiro Gyoba. 1998. Coding Facial Expressions with Gabor Wavelets. Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition. April 14-16. Nara Japan. IEEE Computer Society. pp. 200-205.
[8] M. J. Lyons, J. Budynek, and S. Akamatsu. 1999. Automatic Classification of Single Facial Images. IEEE Trans. On Patt. And Anal. And Mach. Intel. vol. 21. no. 12. pp. 1357-1362.
[9] K. Matsuno, C-W Lee, S. Kimura, S. Tsuji. 1995. Automatic Recoginition of Human Facial Expressions. In Proceeding of the Fifth International Conference on Computer Vision (ICCV' 95).
[10] P. Michel and R. E. Kaliouby. 2003. Real Time Facial Expression Recognition in Video using Support Vector Machines. ICMI'03. November 5-7. Vancouver. British Columbia. Canada.
[11] M. Pantic and L. Rothkrantz. 2000. Automatic Analysis of Facial Expressions: The State of the Art. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22. no. 12. December.
[12] V. Vapnik. 1995. The Natural of Statistical Learning Theory. Springer. New York.
[13] P. Viola and M. Jones. 2001. Robust Real-Time Object Detection. Technical Report 2001/01. Compaq Cambridge Research Lab..
[14] Porawat Visutsak. Emotion Recognition through Lower Facial Expressions Using Support Vector Machines, In Proceedings of The Fifth National Symposium on Graduate Research, Kasetsart University, BKK, Thailand. 10-11 Oct. 2005.