Facial Expression Recognition Using Support Vector Machines

Philipp Michel and Rana El Kaliouby
Computer Laboratory
University of Cambridge
Cambridge CB3 0FD, U.K.
pnm22@cl.cam.ac.uk, rana.el-kaliouby@cl.cam.ac.uk

1 Introduction

Human beings naturally and intuitively use facial expression as an important and powerful modality to communicate their emotions and to interact socially (Ekman, 1982). There has been continued research interest in enabling computer systems to recognise expressions and to use the emotive information embedded in them in human-machine interfaces. This poster presents the application of the machine learning system of support vector machines (SVMs) to the recognition and classification of facial expressions in both still images and live video.

2 On Automated Facial Expression Recognition

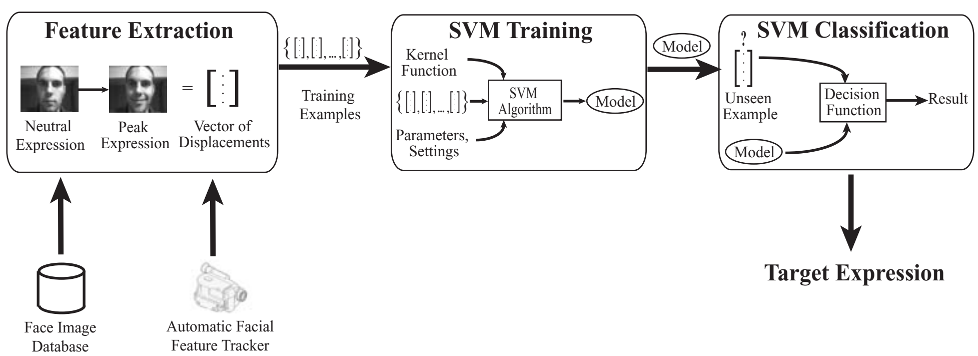

Figure 1: Stages of our automated expression recognition approach

Most approaches to automated expression recognition need to deal with the issues of face localisation, facial feature extraction and the training as well as the classification stages of the learning method used (Pantic & Rothkrantz, 2000). Our initial experiments used displacements of user-defined feature points across a sequence of images as the input to the SVM. Subsequently, we employed an automatic facial feature tracker to perform face localisation and to extract 22 feature displacements from a live video stream. This allows the user to perform training in an interactive manner. The trained SVM model is subsequently used to dynamically classify unseen feature displacements and the result is then returned to the user. Classification thus identifies the feature displacement pattern of an unseen example with the characteristic feature displacement pattern of the most similar expression (e.g. for the basic emotions of ‘anger’, ‘disgust’, ‘fear’, ‘joy’, ‘sorrow’ or ‘surprise’) supplied during training.

3 Support Vector Machines

Support Vector Machines (Cristianini & Shawe-Taylor, 2000) are a maximal margin hyperplane classification method that relies on results from statistical learning theory to guarantee high generalisation performance. Kernel functions are employed to efficiently map input data which may not be linearly separable to a high dimensional feature space where linear methods can then be applied. SVMs exhibit good classification accuracy even when only a modest amount of training data is available, making them particularly suitable to a dynamic, interactive approach to expression recognition. The often subtle differences distinguishing separate expressions such as 'anger' or 'disgust' in our displacement-based data as well as the wide range of possible variations in a particular expression when performed by different subjects led us to the adoption of SVMs as the classifier of choice. Selection of an appropriate kernel function allowed further adjustment and optimisation of the SVM classifier to our particular domain of facial expression recognition. Support vector machines have previously been successfully employed in a variety of classification applications including identity and text recognition as well as DNA microarray data analysis.

4 Evaluation and Conclusion

Table 1: Classification accuracy for each pairing of basic emotions. Rightmost column gives overall mean accuracy for each row's emotion.

Emotion	**Anger**	**Disgust**	**Fear**	**Joy**	**Sorrow**	**Surprise**	*Overall*
Anger	–	72.5%	61.8%	97.1%	93.8%	97.1%	84.1%
Disgust	72.5%	–	63.2%	94.9%	88.9%	100.0%	83.9%
Fear	61.8%	63.2%	–	90.9%	66.7%	100.0%	76.2%
Joy	97.1%	94.9%	90.9%	–	96.8%	97.1%	95.3%
Sorrow	93.8%	88.9%	66.7%	96.8%	–	100.0%	89.4%
Surprise	97.1%	100.0%	100.0%	97.1%	100.0%	–	98.8%

Total accuracy: 87.9%

The Cohn-Kanade AU-coded facial expression database (Kanade, Cohn & Tian, 2000) was used to evaluate expression recognition in still images, while live sessions were carried out to evaluate video-based recognition of spontaneous expressions. Our initial implementation correctly recognised expressions in 78% of trials, with subsequent improvements including selection of a kernel function customised to the training data boosting recognition accuracy up to 87.9%, as illustrated in Table 1. Incorporating further possible enhancements such as adjusting data to account for head motion or performing automatic SVM model selection is likely to yield even better performance and further increase the suitability of SVM-based expression recognition approaches in building affective and socially intelligent human-computer interfaces.

References

Cristianini, N. & Shawe-Taylor, J. (2000) An Introduction to Support Vector Machines. Cambridge: Cambridge University Press.

Ekman, P. (1982) Emotion in the Human Face. Cambridge: Cambridge University Press.

Kanade, T., Cohn, J.F., & Tian, Y. (2000) Comprehensive Database for Facial Expression Analysis. *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00).*

Pantic, M. & Rothkrantz, L.J.M. (2000) Automatic Analysis of Facial Expressions: The State of the Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12), 1424–1445.*