Recognizing Facial Expressions with a Neural Network

Christine L. Lisetti (LISETTI@PSYCH.STANFORD.EDU)
David E. Rumelhart (DER@PSYCH.STANFORD.EDU)
Parallel Distributed Processing Laboratory, Department of Psychology
Stanford University
Stanford, CA 94305-2130

Introduction

Quite a lot of research has been done already in the field of face and gesture recognition with some outstanding results. We take the research effort a step further to ask if a computer can recognize not only which user is facing the computer screen, but also what facial expression [s/]he is showing. The expressions are then categorized in terms of the emotions they signal. The ability to recognize facial expressions can enable: (1) computers to become "aware" of the emotional states the user is experiencing and that information can be included in the user model; (2) human-computer communication to become more natural and is starting to be included in multimodal human-computer interfaces (Sakamoto et al., 1997); (3) software agents with such "awareness" to evolve into more complex and self-motivated artificial systems (Lisetti, 1997); (4) expert systems in automatic real-time recognition of facial expressions to facilitate the development of psychological emotion theories by making easier the experimental data collection of facial expressions.

Some Experimental Results

We started working with two expressions, namely neutral and happy. We used the FERET data base of faces which included smiling faces and neutral faces[1]. We preprocessed the images to crop the relevant portions of the face and to perform histogram equalization for normalizing intensity across the different images. We isolated the third area of the face which is capable of independent movement. It includes the mouth, most of the nose and the chin (Ekman, 1975).

We trained the network on 116 input images of size (74X73). We had images of 52 different persons, two images per person. We included 69 images in our training set and 22 images in our testing set for generalization.

The images contained in the FERET data base were sometimes graded in terms of expressiveness, an issue that we addressed by using graded targets as well. We used targets with range $0.1 < target < 0.6$.

Table 1 summarizes the testing session for generalization. The first column identifies the person. The second and fourth columns list the target values evaluating the degree of expression of each image. The third and fifth column give the generalization results of the network after the training. The network learned and generalized accurately with no errors on this data set.

Table 1: Results of lower face generalizing

PERSON ID	TARGET NEUTRAL	NETWORK OUTPUT	TARGET SMILE	NETWORK OUTPUT
s192	.2	.212	.6	.468
s193	.1	.175	.3	.369
s194	.1	.211	.4	.441
s212	.1	.184	.6	.410
s213	.2	.226	.6	.457
s214	.2	.247	.5	.403
s215	.2	.249	.4	.414
s095	.2	.295	.3	.339
s233	.2	.187	.4	.355
s234	.1	.194	.1	.191
s235	.1	.205	.5	.409

Conclusion

The development of our *Automatic Integrated Facial Expression Interpreter System* is designed with the goal of rendering computers more sensitive to their user's cognitive states. The sub-system proposed in this part of the research is concerned with making the computer more finely aware of the user's body language – in particular, the face. Such a system can be integrated within a larger computer perceptual front-end sending collected information to various sub-systems.

References

Ekman, P., & Friesen, W. V. (1975) *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*, Englewood Cliffs, New Jersey: Prentice Hall, Inc.

Lisetti, C. L. (1997). Motives for Intelligent Agents: Computational Scripts for Emotion Concepts. *Proceedings of the Sixth Scandinavian Conference on Artificial Intelligence*, IOS Press Frontiers in Artificial Intelligence.

Sakamoto, K, & Hinode, H., & Watanuki, K., & Seki, S., & Kiyama, J. and Togawa, F. (1997). A Response Model for a CG Character Based on Timing of Interactions in a Multimodal Human Interface. *Proceedings of the International Conference on Intelligent User Interfaces*, 257-260.

[1]Portions of this research in this paper use the FERET database of facial images collected under the ARPA/ARL FERET program.