Facial expressions analysis by similarity-based neural network

Kenji Suzuki[†] Hiroshi Yamada[‡] Shuji Hashimoto[†]
[†]Dept. of Applied Physics, Waseda University [‡]Dept. of Psychology, Nihon University
Contact email: kenji@ieee.org

Judging and categorizing emotional facial expressions have long been a research subject. It is well-known that Schlosberg proposed a three-dimensional model of facial expressions, which states that facial expression is located along three scales: pleasant - unpleasant, attention - rejection, and sleep - tension. In the model, the facial expressions are arranged onto a low-dimensional semantic space based on conventional statistical approaches. The analyzed result, however, does not contain the relationship between the physical features and the semantic parameters, although the semantic scaling is subsequently given. We consider that more significant perspective of multidimensional perceptual scaling is that high-dimensional physical features perceived by humans are projected to another low-dimensional perceptual space in a nonlinear fashion.

In this paper, we introduce a novel model for the measurement of human subjective evaluation by using *distance mapping learning* network that is a similarity-based multilayer perceptron [1]. The purpose of this work is to realize a multidimensional perceptual scaling that associates physical features of a face with its semantic vector in low-dimensional space. Unlike the conventional multilayer perceptron that learns from a set of an input feature vector and the desired output, the proposed network can obtain a nonlinear mapping between input feature vectors and the outputs by providing a pair of objects and the desired distance of them. The desired distance represents the dissimilarity between two objects obtained by a method of Scheffe's paired comparison. We have conducted an analysis of facial expressions both with a psychological model of line-drawing image of facial expression [2] and real image set.

As illustrated in Fig. 1, when the desired distance (dissimilarity) between two input face images ($\mathbf{R}^m$) is given as a teacher signal, the network can obtain a nonlinear mapping under the condition so that the given distance is preserved in the output space ($\mathbf{R}^n$). The faces are then arranged onto another low-dimensional semantic space. In addition, new data, which are not used in the training in the network, can also be evaluated with the aid of generalization ability of the network. We evaluate the obtained arrangement of faces in comparison with the result of principle component analysis (PCA) and multi-dimensional scaling method (MDS). We also discuss the theoretical and mathematical background underlying the network learning with some related works [3]. Furthermore, by learning from an individual data set on dissimilarity, visualization of the nonlinear mapping is given to represent the differences among individuality and the characteristics common to all.

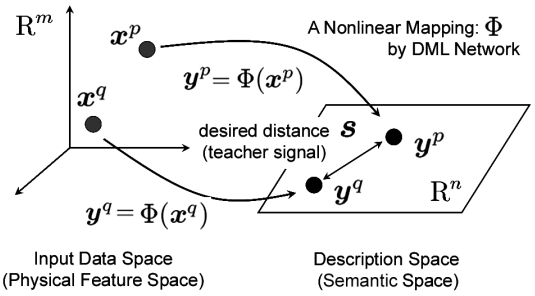

Fig. 1 *Distance mapping learning* network

[1] K. Suzuki, H. Yamada and S. Hashimoto, "Interrelating physical feature of facial expression and its impression," *Proc. of IEEE/INNS Intl. Conf. on Neural Networks*, pp. 1864–1869.
[2] H. Yamada, "Visual information for categorizing facial expression of emotion," *Applied Cognitive Psychology*, **7**, 1993, pp. 257–273.
[3] J. W. Sammon, "Nonlinear mapping algorithm for data structure analysis," *IEEE Trans. Computer*, C-**18**, 1969, pp. 401–09.