Interpolation on data with multiple attributes by a neural network

Hiroshi AZUMI[1], Kazuyuki HIRAOKA[2], Taketoshi MISHIMA[2]
[1]Graduate School of Science and Engineering, Saitama University
[2]Department of Information and Computer Science, Saitama University
255 Shimo-Okubo, Saitama, Saitama, 338-8570, Japan
Tel./Fax.: +81-48-858-3723
[1]E-mail: azumi@me.ics.saitama-u.ac.jp

Abstract: High-dimensional data with two or more attributes are considered. A typical example of such data is face images of various individuals and expressions. In these cases, collecting a complete data set is often difficult since the number of combinations can be large. In the present study, we propose a method to interpolate data of missing combinations from other data. If this becomes possible, robust recognition of multiple attributes is expectable. The key of this subject is appropriate extraction of the similarity that the face images of same individual or same expression have. Bilinear model [1] has been proposed as a solution of this subject. However, experiments on application of bilinear model to classification of face images resulted in low performance [2]. In order to overcome the limit of bilinear model, in this research, a nonlinear model on a neural network is adopted and usefulness of this model is experimentally confirmed.

1. Introduction

Real-world data often have two or more attributes. For example, face images have attributes "individual", "expression", and so on. Of course, one can apply individual-recognition and expression-recognition independently to such data. However, effective use of information in the data cannot be expected in this method since relevance between attributes is not fully exploited

Then, a method of taking both attributes into consideration is desirable. We can expect more efficient recognition with such a method.

However, there are problems in this technique. It is obvious that the structure of this method is more complicated than the former naive method. Furthermore, data for all combinations of multiple attributes must be prepared, and therefore, collection of data is difficult in reality from a physical factor. Then, in this research, we study interpolation of data for missing combinations of attributes. Although there was no restriction in the kind of attribute, in this paper, it was set as two attributes.

2. Model of high-dimensional data with two attributes

In many cases, we can expect that high-dimensional data (e.g. face images) lie on a low-dimensional surface S (Fig. 1). Then, the data can be expressed by the shape of S and small number of parameters that specify a location on S.

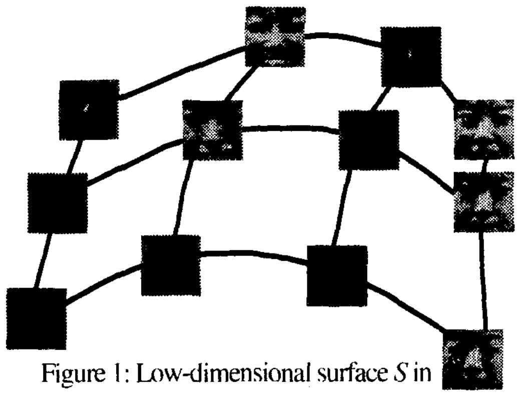

Figure 1: Low-dimensional surface S in high-dimensional data space

Based on this view, let us consider a model which consists of two parts:

1) Express each attribute of given high-dimensional data x by a low-dimensional vector. (Expression part)

2) Reproduce the original data x from the vector expressions obtained in 1). (Reproduction part)

An attribute is mapped into a low-dimensional analog-value vector at expression part. Thereby, the data for any combination of attributes can be expressed.

Parameters of the model are adjusted so that the existing data can be reproduced. Then one can guess missing data by simply inputting the attributes to the model.

This structure has been proposed as the bilinear model by [1]. Since the form of a low dimensional surface S and specification of a location on S are both restricted, the bilinear model has an advantage that only small amount of calculations are required for parameter fitting. However, when the bilinear model was applied to recognition of the individual and expression of face images, only low performance was obtained for classification [2]. This is because distribution of face images in the image space is not necessarily linear.

3. Proposed methods

Based on the consideration in the previous section, in this research, the bilinear model is extended so that nonlinear relations can be expressed. In order to realize general nonlinear functions, a method based on a neural network is proposed. The neural network that is used in this research has four-layers. (Fig. 2)

1) Input layers

Input layer is divided into two segments corresponding to two attributes. They receive the attributes of the data. Each attribute is expressed by a vector of 0-1 values: only the element which corresponds to the attribute value is 1, and other elements are 0. For example, the attribute value "3" is expressed by [0 0 1 0].

2) Expression layers

These are the layers that transform data from the input layers to low-dimensional vectors.

3) Hidden layer

Data is combined from the low dimension vector obtained in the expression layers.

4) Output layer

It is the layer connected to the output of this system. The output of this layer is reproduced data x' which is an estimation of the original data x.

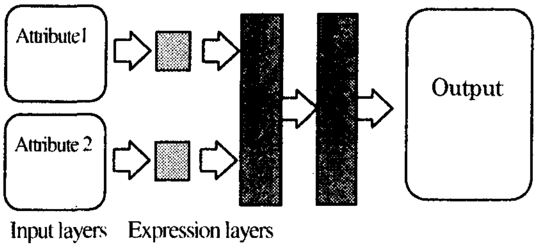

Figure 2: The four-layered neural network for the proposed method

In the above structure, 1) and 2) correspond to the expression part in the previous section, while 3) and 4) play the role of reproduction part. For training of this neural network, data x with labels s, c of two attributes are presented.

We will use standard backpropagation with momentum term [3] for training of the network.

In single trial, weights cannot be adjusted to a desirable value in many cases because of local comergences. To overcome this problem, in this research, training is repeated with different initial values of parameters, which are generated randomly [4]. Then the average of the estimated x' is calculated and it is output as the final guess of the whole system. Thereby, the accuracy of the result is greatly improved compared with single-shot trial.

4. Experiments on face image data

An experiment on face images has been performed. Twelve images of four persons and three expressions were prepared (Fig. 3). In the training phase, eleven images of them have been presented as samples. Then the estimation of the rest one image, which was not shown in the training phase, has been performed. The size of each image is 72x72 pixels.

In this experiment, these face image data have been compressed into ten dimensions by Principal Component Analysis (PCA) [5][6]. A neural network was trained using eleven images $x(1), x(2), \ldots, x(11)$ in Fig. 3. After this, we expect that the neural network can correctly reproduce the rest one image y in Fig. 3. Table 1 is the setup of the neural network that is used in this experiment.

In Fig. 4, left image y is data that was not presented in training phase. It is the target image to be estimated. Right side image y' is output of our neural network, which is the estimation of the target image.

In order to evaluate this result objectively, (A) similarity between target data y and reproduced data y', and (B) similarity between other samples $x(n)$ (n=1,2,...,11) and y' are measured. Then (A) was the minimum value among them. In this meaning, interpolation of this experiment is successful. Similar successful results have been obtained for ten cases of different target images (Fig 5). Failure is observed only in the rest one case (Fig. 6). These results suggest usefulness of the proposed method towards robust recognition of multiple attributes.

However, for human eyes, the results do not look so good. It is because only eleven images $x(1), x(2), \ldots, x(11)$ were used for PCA in the stage of preprocessing. Since the same images for training were used for PCA, reproduced images are restricted to linear combination of these $x(1), x(2), \ldots, x(11)$. If more number of eigen faces based on a larger database is used as bases of compressed data, visual impression of the result will be improved.

Figure 3:Face images for experiment (four persons and three expressions)

y y'

Figure 4: Target image which was not shown in training phase (left) and output of our neural network (right)

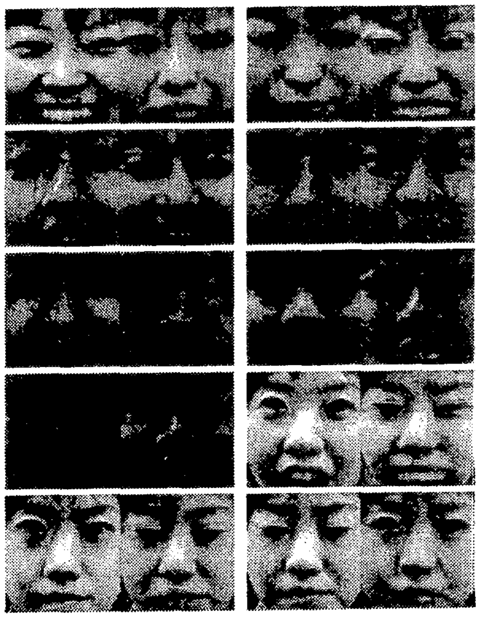

Figure 5: Target images (left) and output of our neural networks (right)
(Successful cases)

Figure 6: Target image (left) and output of our neural networks (right)
(Failure case)

Table 1: Setup of the neural network used in the experiment

Number of attributes	2
Kind of attribute 1	4
Kind of attribute 2	3
Number of neurons in the expression layer 1	2
Number of neurons in the expression layer 2	2
Number of neurons in the hidden layer	20
Number of neurons in the output layer	10
Learning rate	0.1
Moment coefficient	0.9
Number of epochs	50000
Number of trials	20

5. Conclusions

In this paper, about data with multiple attributes, we proposed a method to interpolate data of missing combinations from other data. In the proposed method, attributes are expressed by low-dimensional vectors, and data are reproduced from these vectors. As extension of the conventional bilinear model, the neural network was used for the proposal method. The average value of the data of several trials was used as a final guess, because accuracy was low in single trial. We experimented using the twelve face images: four persons and three expressions. Eleven images were used for training, and rest one was estimated by trained network. Similarity between output data and target data, and similarity between output data and data for training were compared. Estimation of this experiment was successful. These results suggested usefulness of the proposed method towards robust recognition of multiple attributes.

This work has been partly supported by JSPS (Japan Society for the Promotion of Science), 14580405.

References

[1] Joshua B. Tenenbaum and William T. Freeman: Separating Style and Content with Bilinear Models, *Neural Computation*, 12, 2000, pp.1247-1283.

[2] Tetsuya KUROO: Realization of the individual discernment and the expression recognition using bilinear model. *Graduation thesis of Information and Computer Science*, Saitama University, 2001.

[3] D. E. Rumelhart, G. E. Hinton, R. J. Williams: Learning internal representations by error propagation, D. Rumelhart and J. McClelland, editors. *Parallel Data Processing*, Vol.1, Chapter 8, the M.I.T. Press, Cambridge, MA 1986, pp.318-368.

[4] Nguyen, D. et al: Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights, *IJCNN International Joint Conference on Neural Networks*, San Diego, CA,USA 17-21 June 1990, p.21-6 vol.3.

[5] J. Edward Jackson: *A User's Guide To Principal Components*, A Wiley-Interscience Publication, 1991.

[6] Richard O. Duda, Peter E Hart, David G.Stork: *Pattern Classification*, A Wiley-Interscience Publication, pp.568-569, 2001.