Proceedings of 1993 International Joint Conference on Neural Networks

QUANTIZATION LEVEL INCREASE IN HUMAN FACE IMAGES USING MULTILAYER NEURAL NETWORK

Kenji NAKAYAMA Yoshinori KIMURA Hiroshi KATAYAMA

Dept. of Electrical and Computer Eng., Faculty of Tech., Kanazawa Univ.
2-40-20, Kodatsuno, Kanazawa 920 JAPAN
E-mail: nakayama@haspnn1.ec.t.kanazawa-u.ac.jp

ABSTRACT

In this paper, quantization level increase in human face images using a multilayer neural network (NN) is investigated. Basically speaking, it is impossible to increase quality without any other information. However, when images are limited to some category, image restoration could be possible, based on the common properties in this category. The multilayer NN is trained using human face images of 32x32 pixels with 8-levels as the input data, and 256-level images as the targets. The standard back-propagation (BP) algorithm is employed. 20, 40 and 100 training data are examined. By increasing the training data, a general function of regenerating missing information can be achieved. The internal structure of the trained NN is analyzed using some special input images. As a result, it has been confirmed that the NN regards the input image as the human face, and extracts features of the face. The input image is transformed using these features and the common properties of the training data, extracted and held on the connection weights, to the human face image.

I INTRODUCTION

Multilayer neural networks (NN) are very attractive for pattern recognition, pattern classification, and so on [1]-[3]. Especially, when extracting a general rule is difficult, and desired responses are given for the input patters, the multilayer NN can be effectively applied. After training converge, the NN can extract a general rule, which can be valid for the other untraining data.

In this paper, capability of regenerating missing information by the multilayer NN is investigated. Quantization level increase in human face images is employed for this purpose. Basically speaking, it is impossible to increase quantization levels without any other information. However, by limiting the images to some category, it can be expected to extract common features of this category, and to use them for regeneration.

II TRAINING OF MULTILAYER NEURAL NETWORK

2.1 Human Face Images

The training and untraining images are represented using 32x32 pixels. This size was determined in order to save computation time, at the same time, to guarantee necessary image quality for evaluating performance of the NN.

2.2 Multilayer Neural Network

A two-layer NN is employed. Low-quality images with 8-levels are applied to the NN. The targets are the corresponding high-quality images with 256-levels. The input layer includes 32x32=1024 units. The number of hidden units vary according to the number of the training data. The output layer represents 32x32 pixel images with 256 levels. So, it also includes 1024 units. Furthermore, an offset unit is employed to provide bias to the hidden and output layers. An activation function is the following sigmoid function.

$$f(x) = \frac{1}{1 + \exp(-x)} \tag{1}$$

The standard back-propagation (BP) algorithm is employed [1]. Initial guess of connection weights are random numbers, uniformly distributed within ± 0.2. Magnitude of each pixel is distributed within [0,1]. Thus, magnitude 0 and 1 correspond to white and black, respectively. The output error is evaluated by the mean squared error given by

$$\text{Error} = \frac{1}{N_p} \sum_{p=1}^{N_p} \{ \frac{1}{N_k} \sum_{k=1}^{N_k} (t_{pk} - o_{pk})^2 \} \leqq \varepsilon \tag{2}$$

N_p and N_k are the numbers of training images and pixels in one image, respectively. o_{pk} is the output of the kth pixel of the pth training image. t_{pk} is the target for o_{pk}. ε is chosen to be 3.8×10^{-4}, which corresponds to 5-level error in the 256-level images. After training converges satisfying Eq.(2) with the above ε, difference between the output image and the target image cannot be visually recognized. At the same time, computation time can be saved.

20, 40 and 100 training data are examined in order to investigate effects of the number of the training data. The number of hidden units is 32, 64 and 100 for the above training data, respectively. The training converge in all cases.

III QUANTIZATION LEVEL INCREASE FOR UNTRAINING IMAGES

Since the training can converge, it is possible to increase quantization level for the training images. However, validity for arbitrary data is important feature of NN. In this section, performance of the trained NN for untrained data is evaluated. The multilayer NN, obtained by using 20, 40 and 100 training data, are denoted NN_{20}, NN_{40} and NN_{100}, respectively.

3.1 Visual Evaluation

Applying the untrained face images with 8-levels, the output image are calculated, and are evaluated visually. Examples of simulation results using NN_{20}, NN_{40} and NN_{100} are shown in Fig.1. It includes (a) the input image with 8-levels, (b) its original image with 256-levels, (c1)-(c3) the NN output images with 256-levels in NN_{20}, NN_{40} and NN_{100}, respectively, and (d1)-(d3) the targets, which are most close to the NN output image in the MSE sense.

In NN_{20}, the NN output image is different from the original. Since the number of training data is small, effect of individual training data is strong. Therefore, the output image is pulled by one of the targets. In this case, increasing quantization levels cannot be realized for the untraining data.

By increasing the number of the training data to 40, the output image approaches to the original. However, it is still different from the original. In NN_{100}, the output image is more close to the original than the target. Thus, by increasing the training data up to about 100, generalization can be obtained.

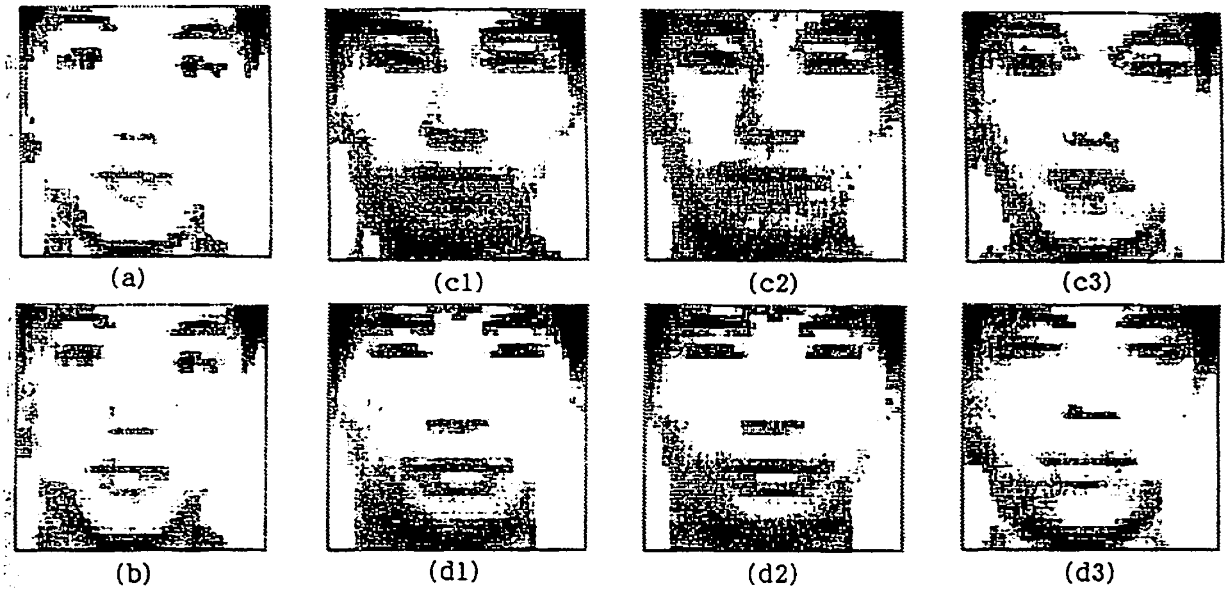

Fig.1 Examples of image transformation. (a)Input image with 8-levels, (b) original image with 256-levels, (c1)~(c3) NN output images in NN_{20}, NN_{40} and NN_{100}, respectively. (d1)~(d3) targets, closest to the output images.

3.2 Evaluation by Mean Square Error

The mean square error is defined by

$$MSE = \frac{1}{N_K} \sum_{K=1}^{N_K} (p_{oK} - p_{rK})^2 \qquad (3)$$

p_{oK} is the kth pixel in the NN output image. p_{rK} is used to express the kth pixel in the original or the targets. Table 1 shows examples. Data 1~5 mean 5 untrained images. A and B indicate MSE between the NN output and the original, and the target, which is closest to the NN output image, respectively. These results are consistent with the previous visual evaluation.

Table 1 Mean square errors between NN output and original(A), and target(B), closest to NN output.

Data	NN_{20}		NN_{40}		NN_{100}	
	A	B	A	B	A	B
1	0.0184	0.0152	0.0098	0.0106	0.0056	0.0086
2	0.0380	0.0110	0.0240	0.0212	0.0180	0.0123
3	0.0098	0.0088	0.0056	0.0104	0.0032	0.0098
4	0.0184	0.0120	0.0128	0.0148	0.0074	0.0134
5	0.0194	0.0078	0.0156	0.0066	0.0038	0.0078

3.3 Transformation of Very Low-Quality Images

Using NN_{100}, transforming 2-level images to 256-level images was tried. Figure 2 shows some results. Since information is drastically reduced from its 8-level version, the NN outputs are a little degraded. However, it can still produce similar images to the originals. Furthermore, missing parts, such as nose and mouth are regenerated.

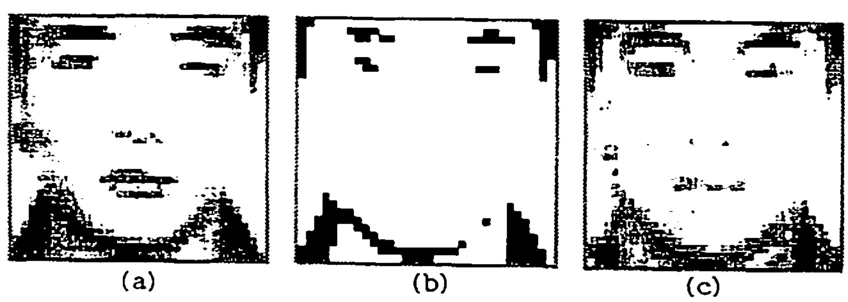

Fig.2 Transformation of untraining 2-level image. (a) Original image with 256-levels, (b) 2-level image applied to NN, (c) NN output image.

IV ANALYSIS OF INTERNAL STRUCTURE

As shown in the previous section, the face images can be regenerated from very limited information. Therefore, it can be expected that the NN gets some capability of transforming arbitrary images to human face images. The NN internal structure is analyzed by using some special input images.

All pixels of the input image take the same value, denoted δ. Therefore, the input image has no information as a pattern. Figure 3 shows the NN output images using δ =0, 0.04, 0.5 and 1. These results show information, held in the NN structure, that is connection weights. When δ =0, that is no input, the NN still outputs the human face image. This information is held on connection weights from the offset unit to both the hidden and output layers, and from the hidden layer to the output layer. Because the sigmoid function defined by Eq.(1) can produce 0.5 for a zero input.

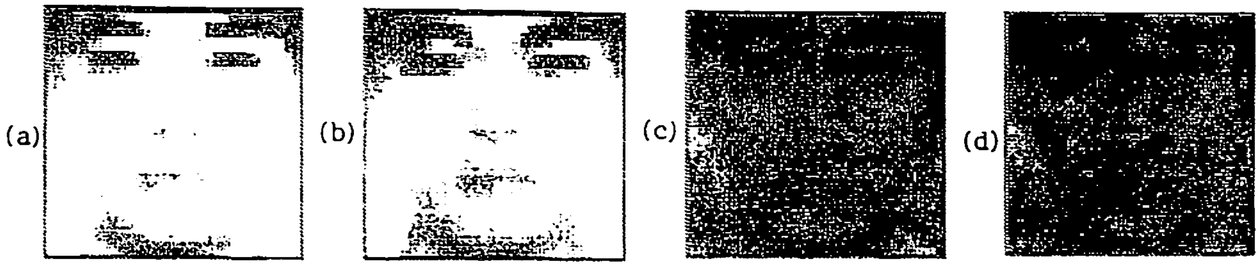

Fig.3 Output images for uniform images. (a) δ =0, (b) δ =0.04, (c) δ =0.5, (d) δ =1

Furthermore, by increasing δ, the output image gradually changes, and magnitude of pixels also increase. Information of face images held in connection weights from the input layer to the hidden layer is a little different from that held in the other connection weights.

Finally, a random image is applied. Figure 6 shows the input and output images. The NN can still generate a human face like image. It is a little different from the output image using δ =0.5 in Fig.5. In this case, the random image is treated as a human face, and features of face are extracted. These features are also used to generate the output image.

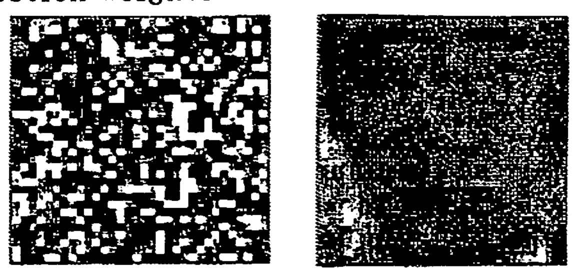

Fig.4 Output image for random input image.

VI CONCLUSIONS

Image restoration by the multilayer NN has been investigated. By limiting images to human faces, and using 100 training data, regeneration of missing information can be achieved. The NN can extract common feature of human faces.

REFERENCES

[1]D.E.Rumelhart and J.L.McClelland, Parallel Distributed Processing, MIT Press 1986.

[2]S.Geman and D.Geman,"Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", IEEE Trans vol.PAMI-6, pp.721-741 1984.

[3]T.Miyajima et al.,"A facial expression recognition using a neural network", Trans IEICE Japan, vol.J75-D-Ⅱ, no.3, pp.671-673, March 1992.