5D－9

Recognition of Facial Expressions Based on the Hopfield Memory Model

Ying Dai Yoshitaka Shibata Kozi Hashimoto
(Iwate Pref. University)
Yasuaki Nakano (Shinshu University)

1.Introduction

The study of the face recognition techniques has importance in person authentication since it has many applications such as security systems, criminal identifications, teleconferences and so on.

In this paper, a method for the facial expression recognition using the Hopfield memory model combined with the pattern matching is proposed. The experimental results for facial expression recognition are also represented and analyzed.

2.Hopfield Net for Face Images

The Hopfield net [1] has only one layer of units. These units play a triple role as input, output, and processing units. The units are globally interconnected and every unit is thus connected to every other units.

It is known that the Hopfield model can store patterns, which can be recalled. In other words, it can work as an associative memory. In order to give the network the capability of the associative memory, the weights are determined as below:

$$W_{(i,j)(i',j')} = \frac{1}{N}\sum_{m=1}^{M} s^m_{(i,j)} s^m_{(i',j')} \qquad (1)$$

where, (i,j) and (i',j') are the units which are linked to each other, N the total number of units, and $s^m_{(i,j)}$ the value of the (i,j)-th pixel of the m-th pattern. M indicates the number of patterns to be memorized.

M faces can be stored in the Hopfield memory by giving the weight $W_{(i,j)(i',j')}$ according to (1). There, one pixel (i,j) corresponds to one unit (i,j). This process corresponds to learning of M faces, each of which will be recalled by the process explained below.

When a facial image is given as an input to the Hopfield memory, the units of the system converge to a stable state finally, if they are updated asynchronously at random times. The value of all the units at the stable states give an associated recall pattern from the face given as input.

By using the stable values of states $\mathbf{s}^r = (s^r_1, ..., s^r_N)$, the weight modification is calculated according to the formula

$$\Delta W_{(i,j)(i',j')} = -k s^r_{(i,j)} s^r_{(i',j')} \qquad (2)$$

where, k is called as an unlearning parameter. It is expected that the pseudo memory is obliterated, if $\Delta W_{(i,j)(i',j')}$ is added to the weight $W_{(i,j)(i',j')}$.

As motioned in [2], when M is equal to 1, the pattern which is memorized in the Hopfield memory is only the standard face of a person. In this case, the system recalls the original image completely. In other words, the similarity between the input pattern and the recalled one is equal to the unity.

Here, the similarity is defined as

$$u = \frac{\sum_x \sum_y (f(x,y) - \bar{f})(t(x,y) - \bar{t})}{\sqrt{\sum_x \sum_y (f(x,y) - \bar{f})^2 \sum_x \sum_y (t(x,y) - \bar{t})^2}}$$

where $f(x,y)$ and $t(x,y)$ indicate two patterns and $\bar{f}$ and $\bar{t}$ do the average values of the patterns respectively.

For $M > 1$, the smaller the value of M is, the more similar the recalled pattern of a test pattern becomes to its standard pattern. In order to inspect the conjecture above, we define

$$d = \frac{u_s - u}{u_s} \qquad (3)$$

as the distance measure of how the recall of the test pattern is close to its standard pattern and separated from the other patterns. Here, u_s indicates the similarity between the recalls of the test pattern and the standard pattern of the same class (person), and u does the one between the recalls of the test pattern and the other classes' standard patterns. If the value of d is large, we will have higher reliability on the decision that the test pattern belongs to the same class, not to the other classes.

表 1: Decision strategy

M	combination of 1st and 2nd candidates	
6	C_1, C_2	C_1, C_2
2	C_1, C_2	C_2, C_1
decision	$C_2(\|d_2\| < \|d_6\|)$ $C_1(\|d_2\| > \|d_6\|)$ $rej(\|d_2\| = \|d_6\|)$	C_2

3.Face Expression Recognition by the Hopfield Model

The six principle emotions are: happiness, sadness, surprise, fear, anger, and disgust. Based on the formulas (1) and (2), the standard templates of six emotions are stored in the Hopfield memory to construct the Hopfield memory model. As the unknown facial expression, the test facial image is inputted into the Hopfield memory to yield its recalled pattern. Calculating similarities between the recalled pattern associated from the input and each template stored in the Hopfield memory, the first and second candidate(C_1 and C_2) are selected from those six classes. If we suppose that the judging of C_1 is correct to the test facial image, the value of d between C_1 and C_2 in eq. (3) are calculated, and is denoted as d_6. As a result, a new Hopfield memory model is organized to store templates for C_1 and C_2. Then the similarities between the recall of the input pattern and each of the two patterns stored in the new Hopfield memory are calculated to determine the new first and second candidates. The value of d between the new first and second candidates are also calculated, and is denoted as d_2.

Table 1 shows the algorithm for deciding the expressions of the unknown face. In the case that C_2 is still the second candidate for $M = 2$ and $M = 6$, C_2 is outputted as the recognition result, if $|d_2| < |d_6|$; while the C_1 is outputted, if $|d_2| > |d_6|$. If $|d_2| = |d_6|$, the output result is rejected.

In the case that C_2 becomes the first candidate for $M = 2$, the C_2 is outputted as the recognition result.

4.Experimental Results and Analysis

The standard templates of six emotions are the facial images which are selected from the face database supported by ORL, and are re-sampled into 30 × 30 pixels. As a result, the Hopfield memory model is encoded with 900 units.

表 2: Experimental results with the facial images

expression	correct	missed	rate
happiness	28	2	93%
anger	4	0	100%
fear	3	1	75%
disgust	4	0	100%
suprise	4	0	100%
sadness	3	1	75%

The 50 facial images, which are selected randomly from the face database of ORL, and re-sampled into the resolution of 30 × 30 pixels, are used as the test patterns in the experiment.

Table 2 shows the details of our results.

According to the table 2, it can be seen that for the facial images of the happiness, the recognition rate is 93%. For the facial images of the anger, disgust and surprise, the recognition rate are 100%. For the facial images of the fear and sadness, the recognition rate is 75%. It is fact that for the expressions of the anger, fear, disgust, suprise and sadness, used samples are too few. But, the experimental results for the facial images of the happiness is satisfactory.

On the other hand, from the results described above, we can get an conclusion that the resolution of 30 × 30 pixels is enough to recognize the expressions of the six emotions.

参考文献

[1] J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Abilities", Proc. Natl. Acad. Sci. U.S.A, 79, pp. 2554-2558 (1982)

[2] Y. Dai, et al. "Recognition of Facial Images with Low Resolution Using A Hopfield Memory Model", Pattern Recognition, No. 2, pp. 159-167, (1998).