Mental Rotation by Neural Network

Conference Paper · September 2009

DOI: 10.1109/IIH-MSP.2009.282 · Source: DBLP

4 authors, including:

Toshihiko Sasama
Tottori University

Hiroshi Mitsumoto
Osaka Electro-Communication University

Shinichi Tamura
Osaka University

IIH-MSP 2009
The Fifth International Conference on
Intelligent Information Hiding and Multimedia Signal Processing
September 12 - 14, 2009 Kyoto, Japan

Mental Rotation by Neural Network

Toshihiko Sasama
Course of information and electronics,
Graduate school of engineering,
Tottori University
Tottori, Japan
sasama@ike.tottori-u.ac.jp

Hiroshi Mitsumoto
Department of environmental,
Osaka Electro-Communication
University
Osaka, Japan
mitsumot@kamilab.osakac.ac.jp

Kazuyo Yoneda and Shinichi Tamura
International technovators institude,
NBL Co., Ltd
Osaka, Japan
{yoneda, tamuras}@nbl.ac

Abstract

In this paper, we propose a neural network with a mental rotation function composed of two, rotation and comparison functions, and discuss its feasibility. The proposed network will output its result of judgment concerning their coincidence or non-coincidence, when two figures are input. If they coincide, the network will output the angle of difference between the two figures. The training for learning the mental rotation is made by teaching the network all rotated figures at every 22.5 degrees. Consequently, realization of the target network was confirmed by simulation experiments.

Key words; neural network, pattern recognition, mental rotation

I. Introduction

The human brain constructed by many neurons has very excellent image recognition ability. Among them, ability of mental rotation has been studied from psychological view point by many researchers [1, 2]. Our goal is to create rotation recognition function using artificial neural network where simple neurons are connected mutually. There are many kinds of studies for rotation recognition using neural network [3]. Fukumi et al. proposed a coin recognition neural network system utilizing rotaion invariant features etc.[4, 5]. Carcenac [6] proposed a modular neural network to rotate mental images. These researches use neural networks to recognize or rotate images. Sankoda [7] studied the normalization function of rotation, that detects inclination angle of input image, and rotate input image to the correct direction. However because of its system input is only one image, its ability is limited to the trained images and its variations. On the other hand, there is a research theme of mental rotation that can rotate any images. The system of mental rotation's input is two images, and it compares and detects inclination angle of two inputs if identical.

In this paper, we simulate the mental rotation function using back propagation neural network, and confirm realization of the target network.

II. Mental Rotation Function

The mental rotation is defined that the one image of the two presented images is rotated one after another, and check whether one is corresponding to another. Fig.1 shows the example of the mental rotation. The left image matched to the right image by 90 degree rotation. The mental rotation includes some functions: rotate, compare, and repeat these actions. Then, one method to create the mental rotation is to integrate some networks that has each ability of functions, and another is to create one network that has all ability of the functions. In this paper, we use one 3-layer back propagation neural network that is the type of the latter. In some studies, the rotation systems calculate and obtain invariant feature values from images by preprocessing, and input these invariant feature values to a neural network. But in this paper, we input images to neural network directly.

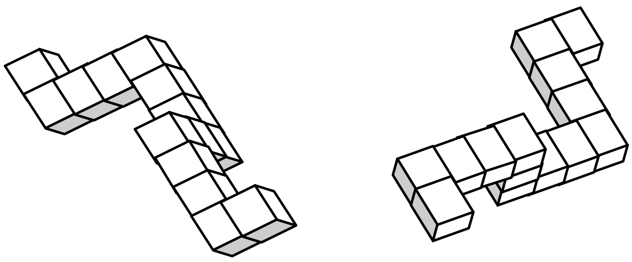

Fig.1 Image recognition using mental rotation.

III. Rotation Learning by Random Images

We detect the rotation angle from rotated input images by one 3-layer neural network. 3 layers are input (2 images), hidden, and output (direction) layer. Fig.2 shows this network model. The number of units in an input layers is 13x13(pixels)x2(images). 13x13 is corresponding to the size of input images. As the normal input images, random images with 13x13 pixels shown in Fig.3 are employed) the system does not learn general rotation function when using limited number of character patterns [7]). 16 rotated versions of each pattern at every 22.5 degrees around the image center as a rotation center were generated, and employed as the teaching patterns at each for rotated angle. To prevent the edge of the image from cutting by the image window according to the rotation, random pixels were set only in the central 9x9 pixels of the image. The gray level value of the input binary pattern part is "1", and the background part is "0".

The number of units in a hidden layer is 64 or 100, and the number of units in an output layer is 1+8 or 1+16. 1 unit is corresponding to matched/unmatched images answer ("0/1"), and 8 or 16 units are corresponding to 8 or 16 directions (each 45 or 22.5 degree) of the difference angle of input 2 images.

First, the network was trained 500,000 set times. An 1 set means training of all combination of the image (for example, 8 type rotated image 1 x 8 type rotated image 2 = 64 combination inputs, or, 16x16=256 inputs. Input image 1 and image 2 use different image at the probability of 50%). The results were evaluated by random image inputs (same as training) and alphabet image inputs.

Fig.4 shows an example of results when the number of hidden layer units is 100 and the number of output layer units is 1+8, and Fig.5 shows an example of results when the number of output layer units is 1+16.

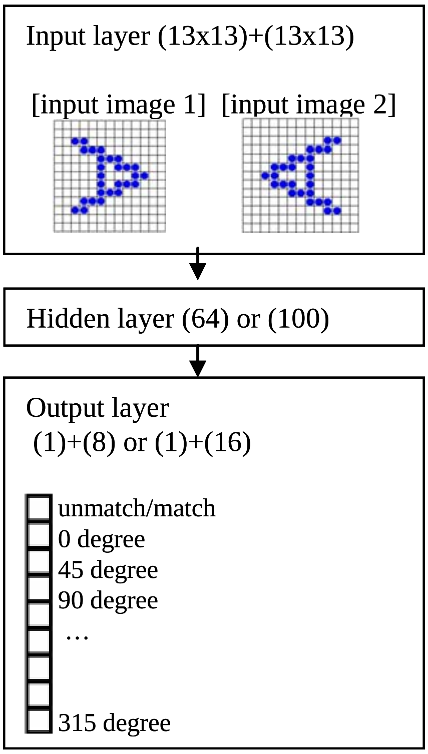

Fig.2 Three layer mental rotation network.

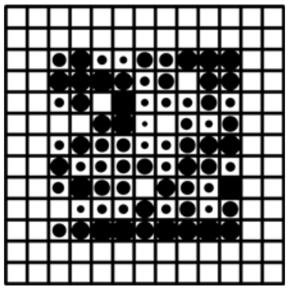

Fig.3 Example of input random image.

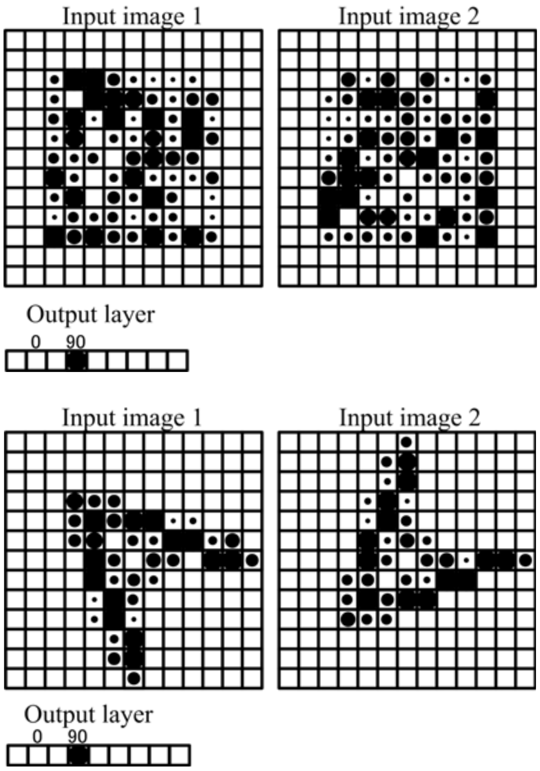

Fig.4 Results for random (upper) and alphabet (lower) images (hidden-layer:100, output-layer:9).

IV. ROTATION LEARNING FOR NOISE ADDED IMAGES

In this section, we train and evaluate using images with added noise. Input image 1 and input image 2 are not in perfect matching any more. The added noise is the Gaussian distribution σ^2 ($\sigma^2 = \alpha \times \sigma^2_{max}$: α is the random scale of 0.0 ~ 1.0, and the variance σ^2_{max} = 0.03, 0.10, and 0.30. Input image 1 and input image 2 use different α at the probability of 50%). Input image pixel values of added noise were rounded within the range from 0.0 to 1.0. The network was trained 500,000 set times, as same as section 3.

Fig.6 shows an example of results when the number of hidden layer units is 100 and the number of output layer units is 1+8, and Fig.7 shows an example of results when the number of output layer units is 1+16. Table I and Table II show the recognition rate of unmatched images and correct rotation direction. These results suggest that the rotation recognition function works effectively as for the situation including the noise, and in artificial fixed pattern, for example, alphabet images.

Table III shows the training error (sum of square difference of network outputs and teaching signals) of hidden layer when input random images.

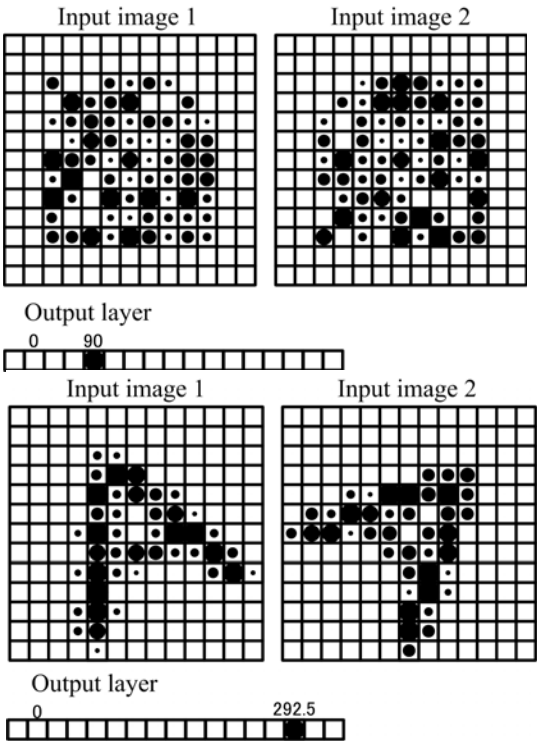

Fig.5 Results for random (upper) and alphabet (lower) images (hidden-layer:100, output-layer:17).

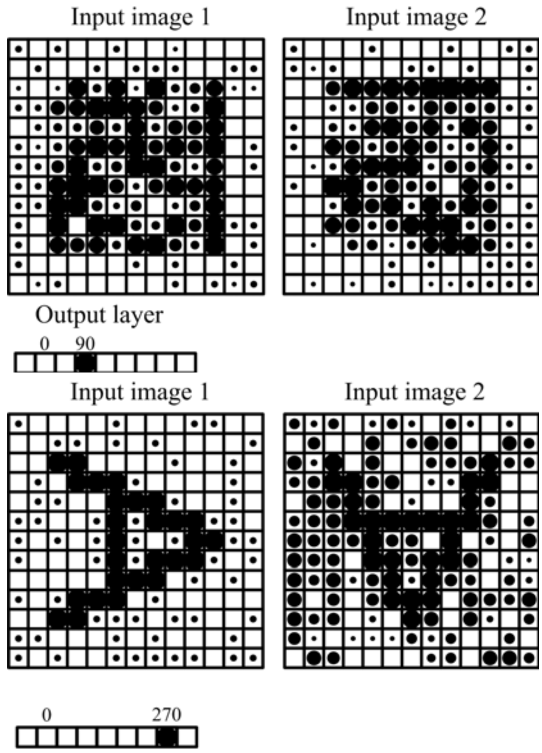

Fig.6 Noisy case results for random (upper) and alphabet (lower) images (hidden-layer:100, output-layer:9).

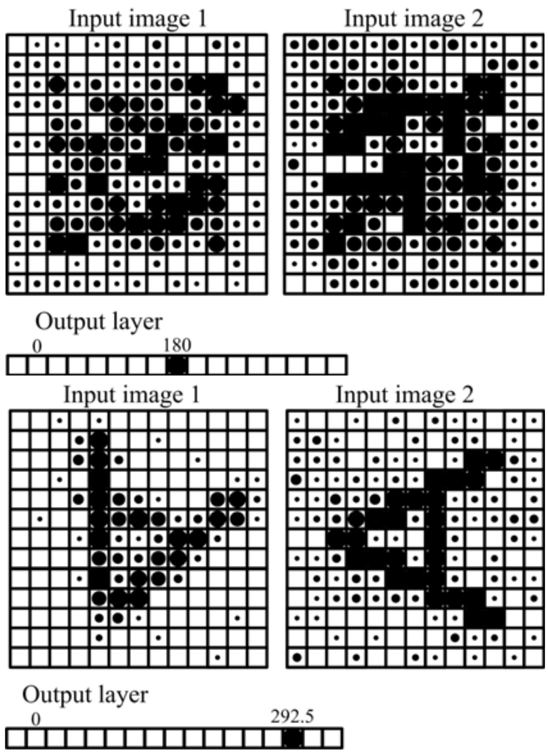

Fig.7 Noiy case results for random (upper) and alphabet (lower) images (hidden-layer:100, output-layer:17).

TABLE I. RECOGNITION RATE OF UNMATCHED IMAGES

Number of directions	16	
Number of hidden layer units	64	100
σ^2_{max} = 0.00	95.40%	97.80%
σ^2_{max} = 0.03	96.90%	98.60%
σ^2_{max} = 0.10	94.35%	97.90%
σ^2_{max} = 0.30	94.45%	95.50%

TABLE II. RECOGNITION RATE OF CORRECT ROTATION ANGLE

Number of direction	16	
Number of hidden layer units	64	100
σ^2_{max} = 0.00	44.50%	61.10%
σ^2_{max} = 0.03	57.70%	59.20%
σ^2_{max} = 0.10	52.65%	58.90%
σ^2_{max} = 0.30	55.25%	57.30%

TABLE III. TRAINING ERROR OF HIDDEN LAYER

	No noise σ^2_{max} = 0.00		Added noise σ^2_{max} = 0.03	
Number of hidden layer units	64	100	64	100
Square error	1.1449	0.8100	1.2766	0.7915

V. CONCLUSION

In this paper, we proposed and evaluated the image rotation recognition system that constructed by one 3-layer back-propagation neural network. This system inputs are 2 images and the output is a rotation angle (angle difference of inputs images, or answering "images are unmatched"). When inputs are noisy alphabetic images, the recognition rate of unmatched images is about 95%, and the recognition rate of correct rotation angle is about 60%. Then, we conclude that the mental rotation function can be created by neural network.

In future works, we want to make image size up from 13x13 to more, and rotation directions up from 16 to 24 directions, and, analyze characteristics of connection weights.

REFERENCES

[1] Parsons, Lawrence M., "Imagined spatial transformations of one's hands and feet," Cognitive Psychology 19: 178-241, 1987.

[2] Moore, D. S. and Johnson, S. P., "Mental rotation in human infants: A sex difference," Psychological Science, 19, pp.1063-1066, 2008.

[3] H. Yang, C.C. Guest, "Performance of back-propagation for rotation invariant pattern recognition", IEEE ICNN87, vol.4, pp.365-370, 1987

[4] Fukumi, M. Omatu, S. Takeda, F. Kosaka, T., "Rotation-invariant neural pattern recognition system with application to coin recognition," IEEE Trans. On Neural Networks, vol.3, No.2, pp. 272-279, 1992.

[5] Fukumi, M. Omatu, S. Nishikawa, Y, "Rotation-invariant neural pattern recognition system estimating a rotation angle," IEEE Trans. On Neural Networks, vol.8, No.3, pp. 568-581, 1997.

[6] Manuel Carcenac, "A modular neural network applied to image transformation and mental images," Neural Computing and Applications, vol.17, No.5-6, pp.549-568, 2008.

[7] S. Sankoda, T. Imura, H. Masuyama, Y. Sato, and S. Tamura, "Normalization of rotated pattern by neural network", The Journal of the Institute of Image Information and Television Engineers, vol.52, no.11, pp.1713-1723, 1998 (In Japanese)