華岡工程學報第二十一期 民國九十六年六月 中國文化大學工學院
Hua Kang Journal of Engineering Chinese Culture University vol.21 (2007) P.93-98

Connectivity Strategies for Second-order Neural Networks Applied to Rotation Invariant Compact Disk Recognition

Tai-Ning Yang[a], Chih-Jen Lee[b], Chun-Jung Chen[c], Shi-Jim Yen[d]

Department. of Computer Science, Chinese Culture University [a, b, c]
Department. of Computer Science and Information Engineering, National Dong-Hua University [d]

Abstract

A second-order neural network is designed to be invariant to changes in rotation. Rotation invariance is achieved through a special arrangement of the network structure. The training set only requires one view of each target object. We describe the weight sharing strategy and present a compact disk recognition neural network illustrating its usefulness. The simulation results show that the proposed neural network can distinguish between the target compact disks independent of the transformation in rotation.
Keywords: Neural networks,Invariant recognition,Rotation invariant features

應用在辨識任意角度旋轉光碟片的二階神經網路連結策略

楊泰寧, 李志仁, 陳俊榮
中國文化大學資訊科學系

顏 士 淨
國立東華大學資訊工程學系

摘要

本研究提出一個新的二階神經網路，我們適當安排此網路的結構，使其能達到旋轉不變的特性，所以此神經網路可以用來辨識任意角度旋轉的物件，最特別是我們在訓練過程並不像大多數其它學者的方式需要同一物件許多旋轉樣本，此網路的訓練集中每一物件只需一版本，實驗顯示我們的方法能成功用來辨識任意角度旋轉的光碟片。

關鍵詞: 神經網路，不變性辨識，旋轉不變特徵

1. Introduction

The invariant object recognition is an important topic in the pattern recognition society. In addition to the traditional statistical approaches, various neural networks such as the necognitron [1], backprop-trained first-order neural networks [2], and higher-order neural networks [3]-[6] have been applied to this problem. Because the invariance is built into the structure directly, the higher-order neural networks alleviate the disadvantage of larger training set and slower training time in the other neural networks approach. The popular structure of the higher-order neural network is the third-order neural network which is translation, rotation, and scale invariance.

The value of an output node in a third-order neural network is $y = f(\sum_l \sum_m \sum_n w_{lmn} x_l x_m x_n)$, where f is a nonlinear threshold function and x_l is the value of the l-th input node. As shown in Fig. 1, the output is determined by the weighted sum of the products from the combined triplets of the input pixels. The weights are shared by the combined triplets which define similar triangles. As shown in Fig. 2, the triangle formed by connecting pixels (a,b,c) is similar to the triangle formed by (d,e,f) such that $w_{abc} = w_{def}$.

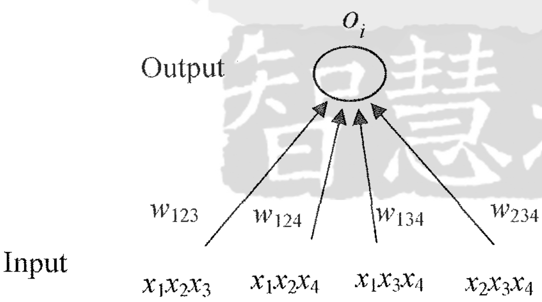

Figure 1. A third-order neural network.

Using this weight sharing rule, the third-order neural network extracts all triangles that are geometrically similar. It is clear that the distribution of the similar triangles included in a two-dimensional object is invariant over the translation, rotation, and scale operations, so that the third-order neural network based on the above weight sharing strategy is translation, rotation, and scale invariant.

An example is given in Fig. 2 where the two face images are the input of the third order neural network. The outputs are the same because the distribution of the similar triangles in the face image is invariant under the translation, rotation, and scale operations. The method for partitioning the set of triangles is explained in Perantonis and Lisboa's paper [4] and a coarse coding technique for reducing the input field size is proposed by Spirkovska and Reid [5].

Since the first-order neural network based on the back-propagation learning rule and the third-order neural network have been proposed for the invariant recognition, an interesting subject is the design of an invariant second-order neural network. The reason why there is no invariant second-order neural network is that the product of two pixels represents a line-segment and all line-segments may overlap with each other when applied with translation, rotation, and scale operations. Since the translation and scale invariance may be achieved in the preprocessing stage, we focus on constructing a rotation invariant neural network. In this paper, we propose a second-order neural network based on a specially designed weight sharing strategy which encodes the rotation invariance in the architecture. The proposed second order feature was successfully applied in the recognition of simple binary patterns such as printed character images [7].

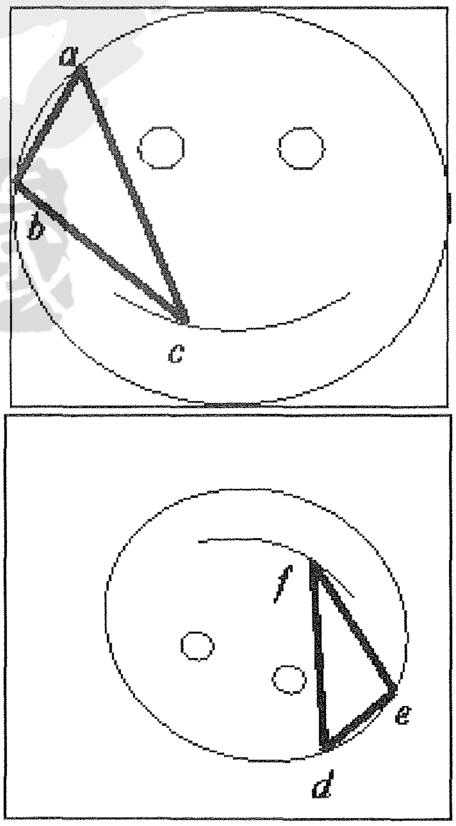

Figure 2. Two similar triangles formed by connecting two triplets of pixels.

2. Rotation invariant second-order neural network

The output of a node in the proposed second-order neural network is given by $y = f(\sum_l \sum_m w_{lm} x_l x_m)$,where f is a nonlinear threshold function and x_l is the value of the l-th input node. Fig. 3 is an example of the second-order neural network. It is simpler than the third-order neural network because the complexity of the weights number is reduced from $O(n^3)$ to $O(n^2)$ where n is the number of input pixels.

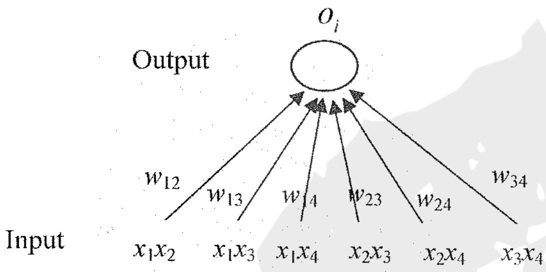

Figure 3. A second-order neural network.

The combined pair of the two pixels represents a line-segment. We use Fig. 4 to explain the proposed connectivity strategy, a weight sharing rule. Since the line-segment connecting the a-th and the b-th pixels may overlap the line-segment connecting the c-th and the d-th pixels after rotation operation. We let them share the same weight, that is $w_{ab} = w_{cd}$. In this way, the number of weights is reduced. For instance, if $w_{12} = w_{24}$ and $w_{13} = w_{23}$ then the second-order neural network in Fig. 3 is transformed to be the network in Fig. 5.

The ideas behind the weight sharing rules in the second-order and the third-order neural networks are similar. In the third-order neural network, the triangles which may overlap after translation, rotation and scale operations share the same weight, while in the proposed second-order neural network, the line-segments which may overlap after rotation operation share the same weight. In the following, we explain the method of partitioning the set of line-segments. Assume that there are $L_1 * L_2 * A$ different weights, where L_1, L_2, and A are positive integers. That is, we reduce the number of weights from $\binom{n}{2}$ to $L_1 * L_2 * A$ by partitioning the line-segments into $L_1 * L_2 * A$ different classes, where n is the number of pixels. Each weight is represented by an ordered triplet from $(0,0,0)$ to $(L_1 - 1, L_2 - 1, A - 1)$.

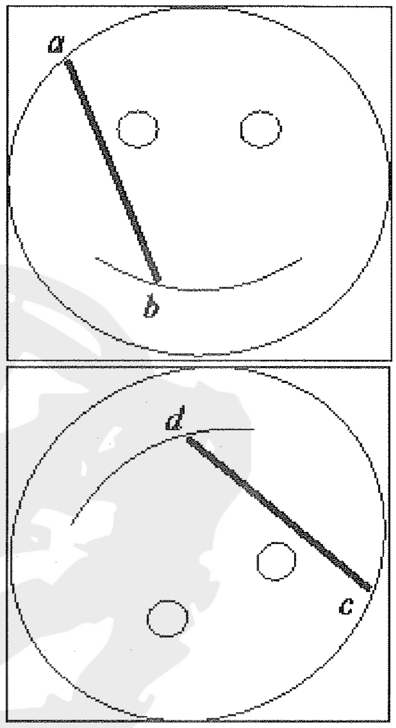

Figure 4. Two line-segments formed by connecting two tuples which share the same weight.

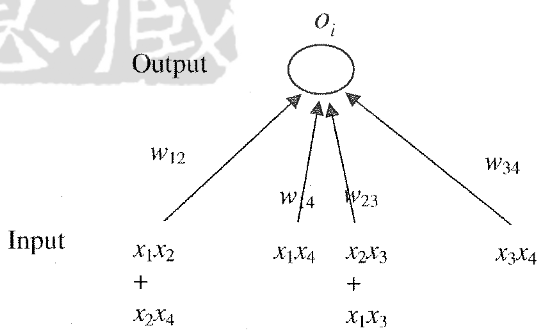

Figure 5. A second-order neural network with weight sharing, $w_{12} = w_{24}$ and $w_{13} = w_{23}$.

In Fig. 6, the distances between the image center and the two pixels are calculated. dis_i and dis_j represent the distances and $dis_i \le dis_j$. θ is the angle from the i-th pixel to the j-th pixel counter-clockwise. This line-segment whose existence indicated by the value of $x_i x_j$ will use the weight with the order $\left(\frac{dis_i * L_1}{dis_{max}+\delta}, \frac{dis_j * L_2}{dis_{max}+\delta}, \frac{\theta * A}{2\pi+\delta}\right)$, where δ is a small positive number and dis_{max} is the maximum distance between all pixels and the center.

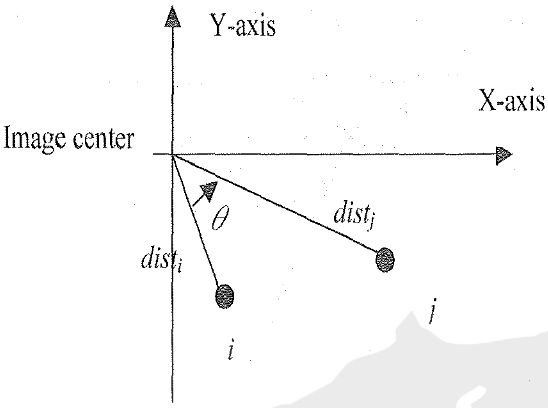

Figure 6. The parameters for determining which weight the combined product $x_i x_j$ share.

Note that in the invariant higher-order neural networks, the connections for the equivalence of the weights are constructed before the presence of any input data. Since the variance caused by the rotation operation disappears after the process of weight sharing which is built directly in the network architecture, the second-order neural network only requires to learn one view of each target object. Thus the training time and the training set size are reduced.

3. Compact disk recognition experiments

Since all standard 12-cm compact disks have the same radius, it is practical to construct a recognition system for distinguishing these disks. In the preprocessing, each compact disk pattern is normalized to be $512*512$ patterns with 256 gray levels. After the edge detection algorithm with Sobel approximation [8], the images is transformed to be binary patterns. Fig. 7 is a sample of the compact disk image. We label the compact disk with xy, where x and y are integers from 0 to 9. In this way, there are 100 targets. We set L_1, L_2, and A to be 8 respectively.

Since the higher-order neural networks are able to provide nonlinear separation, we use a simple perceptron-like rule to train the proposed network. The same rule was used for training the third-order neural network in Spirkovska and Reid's paper [5]. Each compact disk owns ten testing patterns with different random orientations. The system reaches 100% recognition rate. This corresponds to the recognition rate of the third-order neural network in Spirkovska and Reid's experiment [5].

For comparison, we also use the popular ring projection feature in the experiment [7]. The system using ring projection feature only reaches 88% recognition rate. The main reason of the lower rate is it could not distinguish between the compact disks labeling with permutations of same characters such as 12 and 21.

Figure 7. One sample image of the target compact disk labeled with 12.

4. Conclusions

The most important advantage of the proposed second-order neural network is that the rotation invariance is incorporated into the network architecture and only one view of each target pattern is required. Using the proposed weight sharing rule, the number of weights is reduced while the rotation invariance is achieved. The simulation results about the compact disk recognition indicate our model has the merits of smaller training set and quick training time in comparison with the existing approach.

References

[1] K. Fukushima. Analysis of the process of visual pattern recognition by the necognitron. Neural Networks. vol. 2, pp.413–420(1989).

[2] S. E. Troxel. The use of neural networks in PSRI recognition. in *Proc. Joint Int. Conf. Neural Networks,* San Diego, CA, July 24-27, pp.593–600(1988).

[3] L. Spirkovska , M. B. Reid, Connectivity strategies for higher-order neural networks applied to pattern recognition. in *Proc. Joint Int. Conf. Neural Networks,* San Diego, CA, June 18-21., pp.121–126(1990).
[4] S. J. Perantonis , J. G. Lisboa. Translation, rotation, and scale invariant pattern recognition by high-order neural networks and moment classifier. IEEE Trans. Neural Net., vol. 3, no. 2. pp.241–251(1992).
[5] L. Spirkovska , M. B. Reid. Coarse-coded higher-order neural networks for PSRI object recognition. IEEE Trans. Neural Net, vol. 4, no. 2. pp.276–283(1993).
[6] C.Liu , J. H. Kim , R. W.Dai. Multiresolution locally expanded HONN for handwritten numeral recognition. Pattern Recognition Lett. vol. 18, pp.1019–1025(1997).
[7] T.N Yang , S.D. Wang. A rotation invariant printed Chinese character recognition system. Pattern Recognition lett. vol. 22, pp.85–95 (2001).
[8] K.W. Pratt. Digital image processing. Second Edition. John Wiley & Sons, Inc. pp.497–517(1991).