Rotation Invariant Categorization of Visual Objects Using Radon Transform and Self-Organizing Modules

Andrew P. Papliński

Monash University, Australia
Clayton School of IT
Andrew.Paplinski@monash.edu

Abstract. The Radon transform in combination with self-organizing maps is used to build the rotation invariant systems for categorization of visual objects. The first system has one SOM per the Radon transform direction. The outputs from these directional SOMs that represent positions of the winners and related post-synaptic activities, form the input to the final categorizing SOM. Such a network delivers robust rotation invariant categorization of images rotated by angles up to around 12^o. In the second network the angular Radon transform vectors are combined together and form the input to the categorizing SOM. This network can correctly categorized visual stimuli rotated by up to 30^o. The rotation invariance can be improved by increasing the number of Radon transform angle, which has been equal to six in our initial experiments.

Keywords: Radon transform, Self-organizing maps, Rotation invariant vision.

1 Introduction

Radon transform has a long history of application in computer tomography, and relatively recently has been applied in a variety of image processing problems. Typically, Radon transform is used in conjunction with other transforms, wavelet and Fourier included. Magli et al. [1] and Warrick and Delaney [2] seem to initiate the use of Radon transform in combination with wavelet transform. More recently, a similar combination of transforms has been used in rotation invariant texture analysis [3,4], and in shape representation [5]. Other approach to rotation invariant texture analysis uses Radon transform in combination with Fourier transform [6]. Chen and Kégl [7] consider feature extraction using combination of three transforms: Radon, wavelet and Fourier. In [8], texture classification is performed by using a feature descriptor based on Radon transform and an affine invariant transform. Miciak [9] describes a character recognition system based on Radon transform and Principal Component Analysis. Hejazi et al. [10] present discrete Radon transform in rotation invariant image analysis. Close to our considerations are object identification problems discussed by Hjouj and Kammler in [11].

K.W. Wong et al. (Eds.): ICONIP 2010, Part II, LNCS 6444, pp. 360–366, 2010.

In the above papers the reader can find many variants of detailed description of Radon transform and its properties. Here, we can only reiterate the basic fact that Radon transform, $R(\theta, r)$, is composed of sums of pixels along the line that crosses the visual object under the angle θ at the distance r from the origin. It can be noted that Radon transform of the image rotated be a known angle, θ, can be easily inferred from the transform of the un-rotated image. This property makes Radon transform attractive in rotation invariant vision systems.

In this paper we use combination of Radon transform and Self-Organizing maps [12]. Our original idea was related to the way in which human vision could possibly recognize rotated characters as in a process of reading. However, the presented solutions can be used in a variety of systems of rotation invariant categorization of visual objects. We discuss two networks of self-organizing modules that perform the above task in different way.

2 One SOM Per Direction

We start with the system presented in Figure 1in which there is one dedicated self-organizing module, Dir, per Radon transform direction. The image is presented at the receptive field, RF, and is randomly sampled at the points symbolically indicated as green dots. Each line crossing the receptive field symbolizes the 'dendritic' summation of image pixels implementing a single point of Radon transform for a given line, (θ, r). In the example of Figure 1 Radon transform is calculated for $m = 6$ angles, along the $n = 8$ lines, hence, dimensionality of each vector $\mathbf{x}_D$ is $n = 8$, whereas the number of self-organizing modules, Dir, is equal to $m = 6$, that is, the number of Radon transform directions, θ.

In our particular computational examples presented below the diameter of the receptive field, RF, is $n = 75$ pixels normalised into a unity circle. We use letters of the Latin alphabet in 28-point font as the set of the test images. Each directional self-organizing module contain a randomly generated number of neurons approximately equal to πr^2, where r is selected to be equal to 16. Hence that number of neurons varies around 804. Each module produces a 3-dimensional output $\mathbf{y}_D$:

$$\mathbf{y}_D = g(\mathbf{x}_D) \tag{1}$$

where $\mathbf{x}_D$ and $\mathbf{y}_D$ represent input and output signals, respectively, and $g(\cdot)$ describes the Winner-Takes-All function of $\mathbf{W_D x_D}$ which produces a 2-D positional vector $\mathbf{v}$ and related postsynaptic activity d. Such 3-D outputs can be thought of as low dimensional signatures, or labels, specific for each input to the self-organizing modules. In this we follow our other works [13,14,15].

In Figure 2A we show the result of training one of the directional maps, Dir, namely, the 60^o map. Each map is excited with n-dimensional vectors ($n = 75$ in our example) representing the value of Radon transform for a given direction, θ. Each Dir map encapsulates directional similarities of the letters capturing features characteristic for a given direction. The 3-dimensional signatures, or labels, $\mathbf{y}_D$, generated by directional maps are then applied to the combined map, TrImg. An input vector $\mathbf{x}_T$ is of dimensionality $3m$, where m is the number of

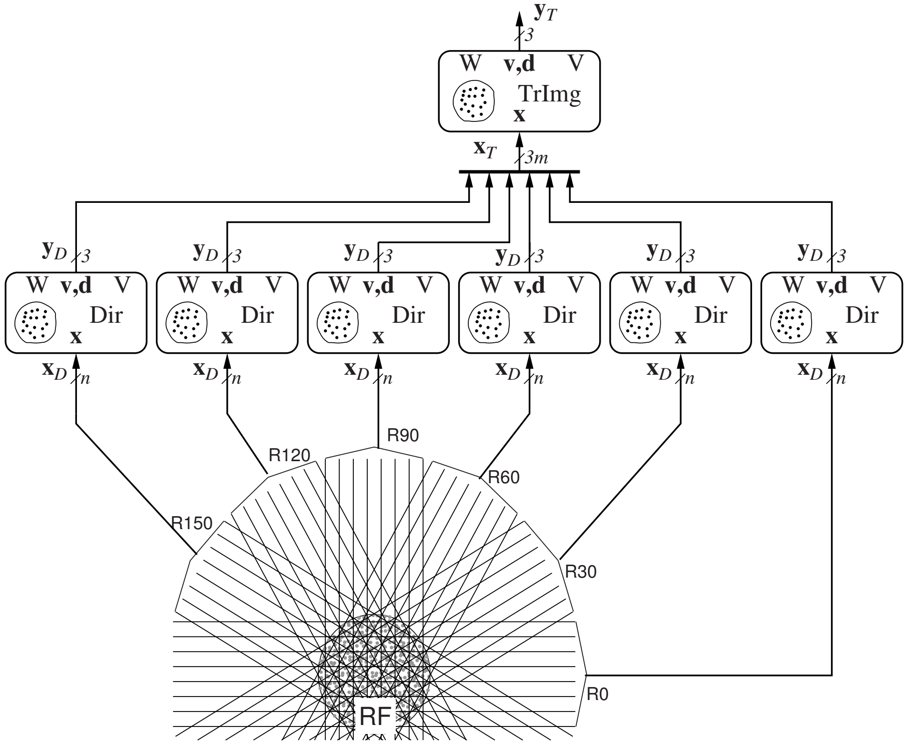

Fig. 1. The categorization network with one self-organizing module per Radon transform direction

Radon transform directions. More formally, for the combined module, TrImg, we can write:

$$\mathbf{y}_T = g(\sum_{i=1}^{m} W_{Ti}\mathbf{y}_{Di}) \tag{2}$$

After training we obtain a combined map as presented in Figure 2B. It can be observed that the combined map captures, as expected, visual similarities between letter.

Now we test the responses of the trained network to the rotated images of letters. The results are presented in Figure 3. We rotate the letters by 2^o angles varying from -12^o to 12^o as indicated in the map. Firstly, it can be noted that majority of the rotated letters are correctly clustered. The quality of the clustering is indicated by the right-hand side plots. The upper plot gives a relative confidence level as measured by the inner products of respective weights and input vectors (see sec. 4 for details). Since these are unity vectors, the maximum of the inner products is equal to 1. The bottom-left plot gives the average size of the rotated letter clusters. Again the radius of the neuronal circle is unity, which gives an idea about the size of the clusters.

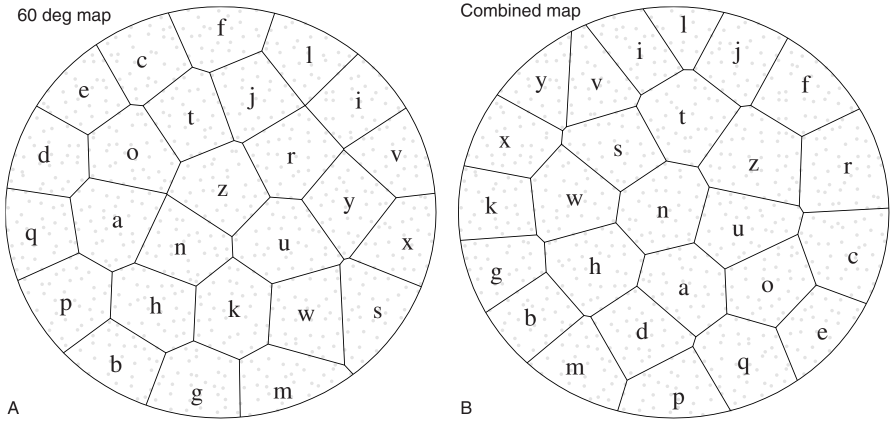

Fig. 2. The result of training maps. A: A directional 60^o Dir map. B: A combined categorization map.

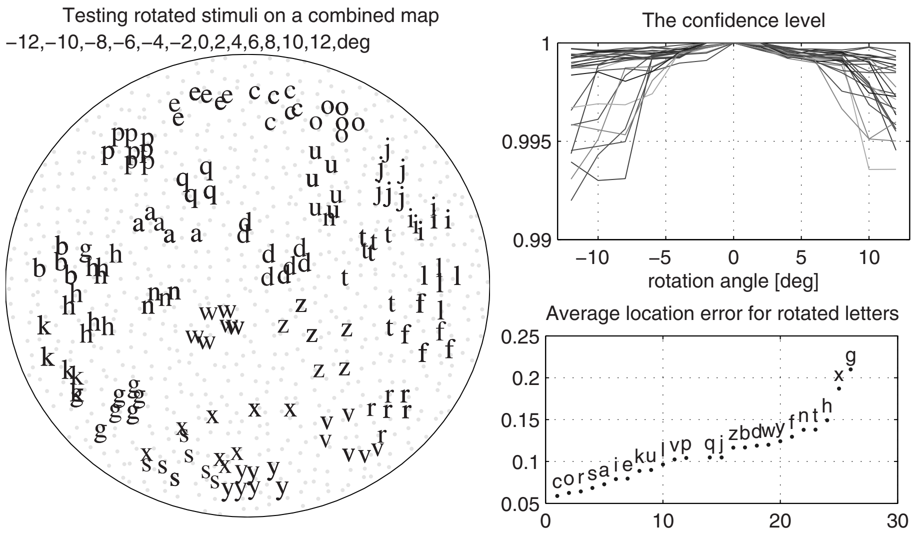

Fig. 3. Categorization of rotated letter by the first network

Although it is encouraging that the above network deliverers categorization invariance for relatively small angles, it would be interesting to find a solution that was invariant to relatively large rotation angles. One possibility is presented in the next section. With reference to eqn (2) and Figure 1 we note that if we rotate the image by the Radon transform angle, it is equivalent to shifting responses, $\mathbf{y}_{Di}$, between the directional self-organizing modules. This gives the potential of the perfect rotation invariant behaviour. This potential is lost, however, when we multiply $\mathbf{y}_{Di}$ by the segments of the weight matrix W_{Ti} that have been trained for non-rotated images.

3 The Single-SOM Network

In this solution we have replaced the directional SOMs by simple summations, as shown in Figure 4. The circular 'dendrites' symbolize the summations of the respective Radon transform rays. The resulting vectors $\mathbf{x}_S$ are of dimensionality $n = 75$ (8 in Figure 4), and are inputs to a single categorizing SOM, TrImg. More formally, we can write:

$$\mathbf{x}_S = \sum_{i=1}^{m} R(\theta_i, \mathbf{r}) \ , \quad \mathbf{y}_S = g(W_T \mathbf{x}_S) \tag{3}$$

where $R(\boldsymbol{\theta}, \mathbf{r})$ is a Radon transform matrix for a given image and vectors $\boldsymbol{\theta}, \mathbf{r}$ of all possible angles and lines, respectively. Summation over all Radon transform angles does remove directional sensitivity, but, unfortunately, ignores the

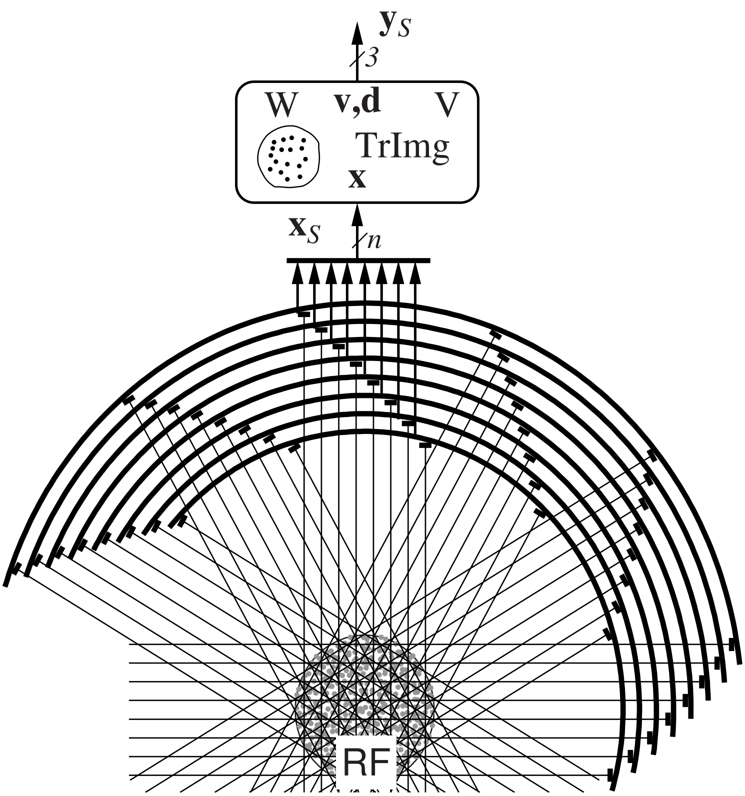

Fig. 4. The categorization network with summed Radon transform and a single SOM

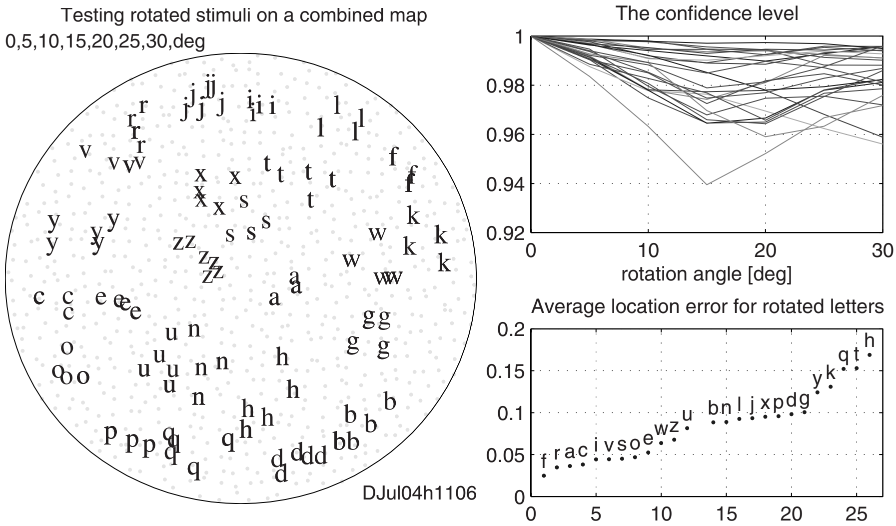

Fig. 5. Categorization of rotated letter by the first network

richness of the image details contained in the full Radon transform. It is, however, expected that the network will correctly classify images rotated by larger angles that in the network of Figure 1. The results of testing the behaviour of the network of Figure 4 are presented in Figure 5. This times we rotate the test images by angles varying from $\mathbf{5^o}$ to $\mathbf{30^o}$. Comparing with the results presented in Figure 3 we note that visually the letters are also well clustered despite of larger rotation angles. This is reflected in the left hand side bottom plot of the average location error, and is confirmed by the higher values of the confidence level presented in the upper plot.

4 Some Implementation Remarks

All input vectors applied to the self-organizing modules, e.g., $\mathbf{x}_{\boldsymbol{D}}$ in eqn (1), or $\mathbf{x}_{\boldsymbol{S}}$ in eqn (3), are normalised and projected on the unity hypersphere by adding one additional dimension. Similarly, all weight vectors are kept on the unity hypersphere. As a result of such an arrangement the inner products of weight and input vectors are equal to the cosine of the angle between such vectors. Working with unity vectors makes it possible to use the dot-product learning law [12] which speeds up the training.

The neuronal lattice is organized in such a way that each neuron is assigned a random position inside a unity circle (see Figure 2). By adding third dimension the position vectors are projected on a 3-D unity sphere. All position vectors are stored in the position matrix $\boldsymbol{V}$ of dimension $\boldsymbol{N}\times 3$, where $\boldsymbol{N}$ is the total number of neurons.

5 Conclusion

We have described preliminary investigation of two networks combining Radon transform and self-organizing maps that are used in categorization of rotated images. Radon transform is easy to implement since it involves only summation of image pixel values along the set of parallel lines crossing the image under a specified set of angles. We have shown that such networks can produce a degree of rotation invariance that can be attractive both in image processing tasks and in analysis of aspects of human vision.

References

1. Magli, E., Presti, L.L., Olmo, G.: A pattern detection and compression algorithm based on the joint wavelet and Radon transform. In: Proc. IEEE 13th Int. Conf. Dig. Sig. Proc., pp. 559–562 (1997)
2. Warrick, A., Delaney, P.A.: Detection of linear features using a localized Radon transform with a wavelet filter. In: Proc. ICASSP, pp. 2769–2772 (1997)
3. Jafari-Khouzani, K., Soltanian-Zadeh, H.: Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms. IEEE Trans. Img. Proc. 14(6), 783–795 (2005)
4. Yu, G., Cao, W., Li, Z.: Rotation and scale invariant for texture analysis based on Radon transform and wavelet transform. In: Proc. 3rd ICPCA, pp. 704–708 (2008)
5. Yao, W., Pun, C.M.: Invariant shape representation by Radon and wavelet transforms for complex inner shapes. In: Proc. IEEE Int. Conf. Inform. Autom., pp. 1144–1149 (2009)
6. Xiao, S.S., Wu, Y.X.: Rotation-invariant texture analysis using Radon and Fourier transforms. J. Phys.: Conf. Ser. 48, 1459–1464 (2007)
7. Chen, G., Kégl, B.: Feature extraction using Radon, wavelet and Fourier transform. In: Proc. IEEE Int. Conf. Syst. Man and Cybernetics, pp. 1020–1025 (2007)
8. Liu, G., Lin, Z., Yu, Y.: Radon representation-based feature descriptor for texture classification. IEEE Trans. Img. Proc. 18(5), 921–928 (2009)
9. Miciak, M.: Character recognition using Radon transformation and principal component analysis in postal applications. In: Proc. Int. Multiconf. Comp. Sci. Info. Tech., pp. 495–500 (2008)
10. Hejazi, M., Shevlyakov, G., Ho, Y.S.: Modified discrete Radon transforms and their application to rotation-invariant image analysis. In: Proc. IEEE Workshop Mult. Sig. Proc., pp. 429–434 (2006)
11. Hjouj, F., Kammler, D.W.: Identification of reflected, scaled, translated, and rotated objects from their Radon projections. IEEE Trans. Img. Proc. 17(3), 301–310 (2008)
12. Kohonen, T.: Self-Organising Maps, 3rd edn. Springer, Berlin (2001)
13. Papliński, A.P., Gustafsson, L., Mount, W.M.: A model of binding concepts to spoken names. In: Proc. 17th Int. Conf. Neural Inf. Proc., Sydney (submitted, 2010)
14. Chou, S., Papliński, A.P., Gustafsson, L.: Speaker-dependent bimodal integration of Chinese phonemes and letters using multimodal self-organizing networks. In: Proc. Int. Joint Conf. Neural Networks, Orlando, Florida (2007)
15. Papliński, A.P., Gustafsson, L.: Feedback in multimodal self-organizing networks enhances perception of corrupted stimuli. In: Sattar, A., Kang, B.-h. (eds.) AI 2006. LNCS (LNAI), vol. 4304, pp. 19–28. Springer, Heidelberg (2006)