WAVELET NEURAL NETWORK FOR 2D OBJECT CLASSIFICATION

Hong PAN, Li-Zuo JIN, Xiao-Hui YUAN, Si-Yu XIA, Jiu-Xian LI and Liang-Zheng XIA

School of Automation, Southeast University, Nanjing, China, 210096
Email: {enhpan, lzxia}@seu.edu.cn

ABSTRACT

In this paper, a wavelet neural network (WNN)-based approach for invariant 2D object classification is proposed. The method employs the WNN characterizing the singularities of the object curvature representation and performing the classification at the same time and in an automatic way. The discriminative time-frequency attributes of the singularities on the object boundary are firstly captured by the continuous wavelet transform (CWT) and then stored by the WNN as its initial scale-translation parameters. These parameters are trained to the optimum status during the learning stage. Thus, only a few convolutions at the optimum scale-translation grids are involved during the classification, which makes our method suitable for real-time recognition tasks. Compared with the artificial neural network (ANN)-based approach preceded by a wavelet filter bank with fixed scale-translation parameters as well as the traditional methods like Fourier descriptors and moment invariants, our scheme demonstrates the best discrimination performance under various noisy and affine conditions.

***Index Terms*—**Wavelet neural network, continuous wavelet transforms, curvature representation, object recognition

1. INTRODUCTION

Wavelet transform is efficient in representing and detecting local features of 2D objects due to the spatial and frequency localization property of wavelet bases [1]. With the detection of local features, an object can easily be recognized. Many new algorithms based on wavelet transform have been developed to solve object recognition problems [2-3]. However, there are some shortcomings for the feature matching of these algorithms. In order to ensure the reliability of the matching results, they all require an enormous number of scales to construct the time-frequency features at various scales during the classification process. Each scale corresponds to convolving the signal with a wavelet function; hence a large number of convolutions are needed for these algorithms. It makes them computationally inefficient.

ANN is good at tasks of pattern matching and classification. In order to further extend the ability of wavelet transform to represent complex patterns and reduce the computational burden of the wavelet-based classification methods, a unique network called the wavelet neural network, which integrating the wavelets into the ANN was proposed [4]. The WNN offers a good compromise between robust implementations resulting from the redundancy characteristics of wavelets and the capability of ANN in learning from examples, so it has been widely used in various areas [5-6]. However, the application of the WNN to object recognition is a relatively new approach. To the best of our knowledge, no research results on this topic have been published until now. Many problems such as the selection of wavelet function, determination of the network structure and initialization of the scale-translation parameters of WNN are yet to be solved.

In this paper, an efficient WNN-based approach for the 2D object classification is proposed. The method employs the WNN characterizing singularities of the object's curvature representation and performing the object classification at the same time and in an automatic way. The WNN can be considered as an expanded perceptron in which the neurons of the first layer are replaced by wavelet nodes. These nodes, due to the variable time-frequency resolution of the wavelet transform, allow the detection of singularities on the boundary as well as the extraction and selection of a small number of meaningful features from its sample points. The obtained features are then regarded as inputs to the subsequent neurons used as a classifier. The remarkable feature of our method is that the local attributes of each pattern can be effectively represented by a limited number of scale-translation parameters after a careful analysis to the CWT modulus extrema in time-frequency plane. It is worth noting that this wavelet analysis can be conducted offline, which does not increase the complexity of the algorithm. The discriminative scale-translation features of each pattern are stored in the WNN as the initial scale-translation parameters of the wavelet nodes. These parameters are then adjusted to the optimum status at the training stage. Thus, instead of matching features by convolving the signal with wavelet functions at a larger number of scales, the computational burden is significantly reduced in our method with only a few convolutions involving at the optimum scale-translation grids during the classification, which makes it suitable for real-time recognitions.

2. WAVELET NEURAL NETWORK

The feature extraction and representation properties of the wavelet transform can be merged into the structure of the ANN to further extend the ability to approximate complicated patterns. A typical structure suggested by these considerations is the WT-ANN which consists of a preprocessing wavelet-based filter banks and an ANN [7-8]. The performance of this WT-ANN architecture is highly dependent on both the values of the translation and scale parameters that characterize the preprocessing wavelet-based filter bank. These values are chosen by users during the design stage according to their knowledge and experience. Once the parameters are determined, no adjustment is possible during the learning stage of the ANN.

To overcome the limitation, the WNN was proposed, in which the classical training of only the connection weights of the ANN is extended to an additional tuning of the scale-translation parameters depending upon the correct classification. The WNN can be considered as an extended perceptron consisting of two parts. The first part contains wavelet nodes in which the classical sigmoidal functions are substituted by the wavelet basis functions. They act as preprocessing units for singularity detection and feature extraction. The classification is performed by the second part, a traditional single-layer or multi-layer perceptron. During the training stage, the WNN is able not only to learn arbitrarily complex decision regions defined by the connection weights, but also to look for those parts of the time-frequency plane that are suited for a more reliable classification of the input signals.

3. THE PROPOSED METHOD

The goal of our research is to design a 2D object recognition approach which utilizes the WNN to extract the optimum scale-translation features of the object boundary curvature representation and to perform the classification automatically. The extracted scale-translation features are outputs of the wavelet nodes which correspond to the optimum CWT coefficients generated by convolving the normalized curvature function with wavelet basis functions at specific scale-translation grids. These optimum scale-translation grids characterize singularities of the curvature representation in time-frequency plane. Thus, by mapping singularities of the curvature representation to dominant points on the object boundary, the time-frequency characteristics of irregular structures along the object boundary are perfectly described by the wavelet nodes with optimum scale-translation parameters. In addition, the number of wavelet nodes and the initial values of the scale-translation parameters can also be determined from the local extrema evolution map of the CWT modulus.

Our approach comprises four steps: 1. The boundary of a 2D object is firstly extracted and resampled such that the boundary data records are scale normalized. 2. Conduct boundary transforms to obtain the curvature representation and make the representation invariant to rotation. 3. Feed the invariant curvature representation into the WNN and calculate the wavelet coefficients by convolving it with the wavelet nodes which are tuned into the optimum scale-translation grids during the learning stage. 4. Classify the input object into the class whose representative boundary exhibits the maximum similarity to that of the input object.

Since the information on the shape of a curve is often concentrated at dominant points having high curvature, the curvature representation [9] of a curve plays an important role in image analysis. The curvature representation itself is invariant to translation. Therefore, only scale and rotation normalization should be taken into account. In this work, the boundary is resampled and normalized into 1024 equally distant points. The rotation-invariant curvature representation can be achieved by searching curvature values along the boundary and selecting the point with the largest curvature value as the middle point of the boundary. Finally, we normalize the curvature sequence by dividing its maximum absolute value, as shown in Eq.(1).

$$C(t) = \frac{c(t)}{\max\limits_{0 \le t \le T-1} |c(t)|} \tag{1}$$

where $c(t)$ and $C(t)$ denote the original and the normalized curvature function, respectively. T denotes the length of the normalized boundary.

4. CONSTRUCTION OF THE WNN

4.1. General architecture

For a multi-class recognition task, like our application, it is advisable to use a system with the multi-network single output (MNSO) structure. MNSO divides the whole WNN into several clusters. Each cluster consisting of one sub-WNN is responsible for the discrimination of the objects belonging to a single class. Outputs of all clusters are compared in a comparison layer which selects the maximum value as the output of the WNN. The input object is then classified into the class represented by the cluster with the maximum output. As shown in Fig.1, in our system, the WNN comprises several clusters, i.e. sub-WNN1~sub-WNNk, where k is the number of the object classes. Each sub-WNN is established by a three-layer structure. The first layer contains wavelet nodes with scaled and translated versions of the wavelet function as activation functions. In the other two layers, traditional sigmoid neurons are used. The input of each sub-WNN is the normalized curvature function of the object boundary. The real Morlet wavelet $\psi(t) = \cos(1.75t)e^{-t^2/2}$ is used as the kernel function of the wavelet nodes for all sub-WNNs in that it provides the best recognition performance.

Classification performance of the WNN greatly depends on its structure. As will be discussed, the decision of the number of wavelet nodes and the initialization of the scale-translation parameters of wavelet nodes rely on the spectral content characterizing the boundary of the class and can be solved in a reasonable manner by analyzing its normalized curvature function in time-frequency plane. In the following parts, let us take the sub-WNN1 and airplanes shown in Fig.2 as an example to illustrate the design procedure of the WNN.

4.2. Determination of the number of wavelet nodes

As discussed in section 3, the characteristics of dominant points on the boundary can be adaptively explored by the wavelet nodes with the optimum scale-translation parameters. Thus the number of wavelet nodes is uniquely decided by the number of explored dominant points. On the other hand, the singularities of a curvature representation can be detected from the local extrema of its wavelet transform. The evolution across scales of the local wavelet extrema specifies the location and the local shape of these singularities that correspond to the dominant points of the object boundary. We can thus not only detect the dominant points but also characterize them with the scale-translation parameters of the wavelet nodes.

In order to build the evolution map of the local wavelet extrema across scales, the CWT of the normalized curvature function of a representative sample of the class is evaluated. Fig. 3(b) depicts the projection of the CWT coefficients of the curve shown in Fig. 3(a), i.e. the normalized curvature function of the airplane from class-1, on scale-translation plane. Then, we relate each wavelet modulus extrema at the scale a to a wavelet modulus extrema at the scale (a+1) which is as close as possible and with the same sign. Thus, we can obtain a sequence of modulus extrema lines (extrema skeletons) that propagate across scales up to the finest scale. Each of extrema skeleton represents a potential or false singular point of the curvature function from class-1, because boundary noise may also introduce modulus extrema. However, the wavelet amplitude of the boundary singularity decreases when the scale decreases, while the wavelet amplitude of the boundary noise increases on average when the scale decreases [1]. At this point, we remove any extrema skeleton whose amplitude increases on average when the scale decreases and retain those extrema skeletons whose amplitude decreases when the scale decreases. Most of the boundary noise can be discarded during this selection. Because of the noise stochastic nature, there may be some extrema

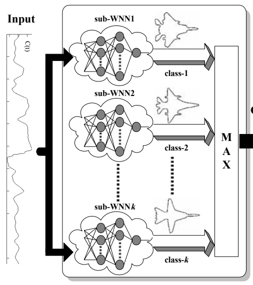


Figure 1. Structure of the proposed WNN

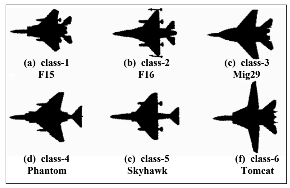
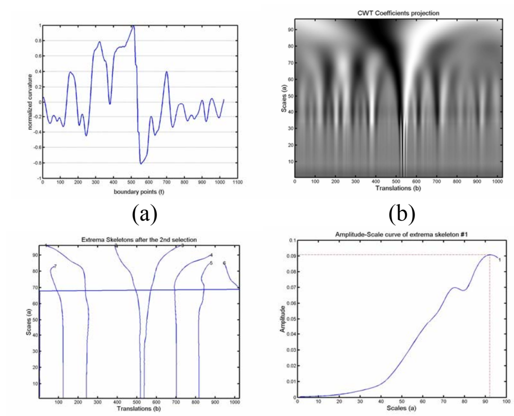


Figure 2. Six classes airplanes for classifications

skeletons which represent the same trends as the extrema skeletons due to the actual boundary singularities, can not be got rid of by the above elimination. In order to attain a complete reduction of noise influence, a further selection has to be performed by introducing a length threshold L to the extrema skeletons. Usually, important singularities can be detected at a wide range of scale levels and their extrema skeletons cover almost all scale levels. On the contrary, local extrema of the noise often concentrate at low scale levels and may disappear quickly as the scale increases, which probably results in extrema skeletons with short length. Therefore, among the remaining extrema skeletons after the first selection, we neglect all extrema skeletons whose length is shorter than the length threshold L. These are extrema skeletons that are mostly influenced either by the noise or by small ripples of the object boundary which do not carry significant information. As a result, by omitting the extrema skeletons due to the noise and non-significant singularities, only the dominant singularities with long extrema skeletons will be kept. After the second selection, each remaining extrema skeleton characterizes a dominant point on the object boundary. Fig.3(c) shows the pruned extrema skeleton map after two selections (L = 80). An extrema skeleton contains important features of the relevant dominant point, such as the location of the point and representative scale-translation parameters where the maximum wavelet modulus appears. In our sub-WNNs, each detected extrema skeleton due to a dominant point is characterized by two wavelet nodes, i.e. the L-Node which represents the location of the point, and the ST-Node which describes the representative scale-translation features of the point. Thus, the number of wavelet nodes in a sub-WNN is twice the number of the dominant extrema skeletons of the object boundary from a certain class. In particular, as suggested in Fig.3(c), there are 7 extrema skeletons detected in the pruned extrema skeleton map of class-1, so the number of wavelet nodes required in the sub-WNN1 is 14. For other sub-WNNs, the number of wavelet nodes can also be determined in a similar way.

4.3. Initialization of scale-translation parameters

The scale-translation values of the wavelet nodes and the connection weights form the parameters of a sub-WNN. Connection weights can be randomly initialized between (-1,1), while the scale-translation parameters should be carefully initialized to match the time-frequency properties of the dominant point described by the wavelet node. Note that each extrema skeleton due to a dominant point is characterized by two wavelet nodes, i.e. L-Node and ST-Node. Thus, initial scale-translation values of these two wavelet nodes can be decided by the relevant extrema skeleton. For the L-Node, its scale parameter a_1 is always



(a) (b)

(c) (d)

Figure 3. (a) Normalized curvature curve of class-1, (b) CWT projection on the scale-translation plane, (c) Pruned extrema skeletons, (d) Amplitude-scale curve of the extrema skeleton #1

set to 1 and its translation parameter b_1, corresponding to the accurate position of the dominant point, is obtained by tracing the extrema skeleton when it intersects with the finest scale level a = 1. For the ST-Node, we build a wavelet amplitude-scale curve according to the extrema skeleton and find the scale level where the maximum amplitude occurs. The scale coordinate of the amplitude maxima serves as the scale parameter a_2 of the ST-Node, while the translation parameter b_2 can be initialized by the scale-translation mapping existing in the extrema skeleton. Fig.3(d) exhibits the amplitude-scale curve of the extrema skeleton #1, in which the maximum amplitude occurs at scale level a = 92. Combining information provided by the extrema skeleton #1, scale-translation parameters of the two wavelet nodes are initialized as (a_1 = 1, b_1 = 243) and (a_2 = 92, b_2 = 102). Initial scale-translation parameters of the rest wavelet nodes in the WNN can also be obtained similarly.

5. EXPERIMENTAL RESULTS

Airplanes belonging to six classes, i.e. class-1~class-6, as shown in Fig.2, are used as targets for classifications. In this experiment, 1980 samples suffering from geometry transform and random Gaussian noise comprise the test sets. 1320 samples are randomly selected to form the training sets for the learning of the WNN. For each sample from a certain class, test sets are formed by changing the size of the sample in various positions and orientations. In addition, these samples are corrupted with noise and distortion. Noise and distortion effects are introduced by adding random noise to the boundary points. The measure of noise in an object boundary is determined by the signal-to-noise ratio (SNR).

The efficiency of the proposed WNN-based recognition method has been compared with the methods based on an ANN preceded either by a wavelet filter bank (WT-ANN) or traditional shape descriptors, like the Fourier descriptors [10] (FD-ANN) and the moment invariants [11] (MI-ANN), which act as a preprocessing stage. The WT-ANN is characterized by the same structure of the related WNN. In particular, the input layer consists of wavelet nodes with the scale-translation parameters having the same values as those used in the initialization of the wavelet nodes of the WNN. A random initialization, on the other hand, has been preferred for any sigmoid neuron involved in the structure of the

Table 1. Comparison of recognition results of various classifiers based on different SNRs

classifiers	class-1	class-2	class-3	class-4	class-5	class-6
SNR = 50dB						
WNN	95.45%	97.58%	98.48%	96.97%	97.58%	100%
WT-ANN	93.94%	95.45%	98.18%	94.55%	96.97%	100%
FD-ANN	90.61%	92.12%	96.97%	93.33%	94.24%	98.48%
MI-ANN	90.91%	92.42%	97.27%	93.03%	93.94%	98.48%
SNR = 40dB						
WNN	94.24%	96.97%	98.48%	96.97%	96.97%	99.39%
WT-ANN	91.82%	93.94%	95.45%	93.03%	94.85%	96.97%
FD-ANN	86.97%	89.70%	94.55%	88.48%	91.52%	94.85%
MI-ANN	86.36%	92.42%	94.85%	89.39%	92.42%	95.45%
SNR = 30dB						
WNN	92.12%	94.55%	96.97%	93.94%	95.45%	98.79%
WT-ANN	87.58%	93.33%	93.64%	88.79%	92.42%	95.76%
FD-ANN	82.42%	87.27%	90.91%	83.64%	86.36%	93.94%
MI-ANN	83.33%	89.39%	91.52%	84.85%	87.88%	94.55%
SNR = 20dB						
WNN	89.39%	93.03%	95.45%	89.39%	93.33%	96.97%
WT-ANN	83.33%	89.70%	90.30%	84.55%	88.79%	91.82%
FD-ANN	74.24%	80.61%	83.94%	76.97%	81.52%	86.36%
MI-ANN	76.97%	83.33%	85.45%	79.39%	83.33%	88.48%
SNR = 10dB						
WNN	86.36%	89.39%	90.91%	84.85%	89.70%	94.55%
WT-ANN	78.48%	83.33%	83.64%	77.27%	81.82%	88.79%
FD-ANN	70.30%	74.24%	75.76%	68.79%	73.94%	80.61%
MI-ANN	72.73%	76.06%	77.27%	70.91%	75.76%	82.42%

hidden layer and the output layer. For the FD-ANN, a subset containing 10 lowest frequency Fourier coefficients of the object boundary is served as its inputs and the classifier uses a multi-layer perceptron network with one hidden layer. Moment invariants of an airplane are obtained from a binary image of the airplane where pixels inside the boundary are assigned intensity level 1 and pixels outside the boundary are assigned intensity level 0. We apply a set of moment invariants [11] to extract the regional features of airplanes and feed them into the MI-ANN.

The discrimination performances of the various classifiers are assessed and compared in Table I. From the obtained results, we can make the following observations. (1). Among all the classifiers, the WNN has always shown the best performance in terms of the classification accuracy. The WNN outperforms the WT-ANN due to the additional tuning of the scale and translation parameters of the wavelet nodes during the learning stage. While the scale-translation parameters of the WT-ANN can not adaptively adjust to the optimum status, once they are determined. On the other hand, both wavelet-based methods achieve better discrimination results than the FD-ANN and the MI-ANN. Fourier descriptors and moment invariants provide global features which are able to distinguish objects with different shapes, thus recognition results of the FD-ANN and the MI-ANN are comparable to that of the WT-ANN for class-3 and class-6. But global features are incapable of differentiating objects with similar shapes, which interprets the low recognition rates of the FD-ANN and the MI-ANN for class-1, class-2 and class-4, class-5. Moreover, it is also worth noting how good the behavior of the WNN is with respect to the increasing noise level. The performance degradation of the WNN influenced by the noise is always smaller than that characterizing other signal processing solutions. (2). For high noise levels (10dB and 20dB), the WT-ANN performs much better than the FD-ANN. This outcome can be justified by the capability of the wavelet transform to shift most of the noise power to the highest frequencies on time-frequency plane, which is absent in the Fourier transform-based approach. (3). The performance gap between the WNN and the WT-ANN increases as the SNR decreases. This result shows that the scale-translation parameters of the WNN are adapted to the optimum status during the learning process, which can provide a fair benefit over a non-optimized time-frequency analysis provided by the wavelet filter bank in the WT-ANN.

6. CONCLUSION

In this paper, we developed an efficient 2D object recognition method using the WNN. The normalized curvature function was applied to represent the shape of an object. By making a preliminary CWT analysis to the curvature representation of an object contour, we fully capture the local time-frequency attributes of the singularities on the object boundary. Thus the number of wavelet nodes and the initial scale-translation parameters of the WNN can be determined under the guidance of the knowledge about the time-frequency properties of the object boundary, without any ambiguity. Compared with the WT-ANN with fixed scale-translation parameters and traditional methods like Fourier descriptors and moment invariants, our scheme demonstrates the best recognition performance under various noisy and affine conditions.

7. REFERENCES

[1] S.Mallat and W.L.Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Information Theory*, vol.38, pp.617-643, 1992.

[2] O.M.Tieng and W.W.Boles, "Recognition of 2D object contours using the wavelet transform zero-crossing representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.19, pp. 910-916, 1997.

[3] M.I.Khalil and M.M.Bayoumi, "Invariant 2D object recognition using the wavelet modulus maxima," *Pattern Recognition Letters*, vol.21, pp.863-872, 2000.

[4] Q.H.Zhang and A.Benveniste, "Wavelet networks," *IEEE Trans. Neural Network*, vol.3, pp. 889-898, 1992.

[5] C.H.Lin, and C.H.Wang, "Adaptive wavelet networks for power-quality detection and discrimination in a power system*", IEEE Trans. Power Delivery*, vol.21, pp.1106-1113, 2006.

[6] A.Subasi, M.Yilmaz and H.R.Ozcalik, "Classification of EMG signal using wavelet neural network," *Journal of Neuroscience Methods*, vol.156, pp. 360-367, 2006.

[7] S.Osowski, and D.D.Nghia, "Fourier and wavelet descriptors for shape recognition using neural networks-a comparative study", *Pattern Recognition*, vol.35, pp.1949-1957, 2002.

[8] Z.L.Gaing, "Wavelet-based neural network for power disturbance recognition and classification", *IEEE Trans. Power Delivery*, vol.19, pp. 1560-1568, 2004.

[9] S.Abbasi, F.Mokhtarian, and J.Kittler, "Curvature scale space image in shape similarity retrieval," *Multimedia Systems*,vol.7, pp.467-476, 1999.

[10] C.T.Zahn and R.Z.Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Computer*, vol.21, pp. 269-281, 1972.

[11] M.K.Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Information Theory*, vol.8, pp,179-187, 1962.