A Robust Modular Wavelet Network Based Symbol Classifier

A.K. Mishra, P.W. Fieguth, D.A. Clausi

University of Waterloo, Vision and Image Processing Group
akmishra, pfieguth, dclausi@uwaterloo.ca

Abstract. This paper presents a robust automatic shape classifier using modular wavelet networks (MWNs). A shape descriptor is constructed based on a combination of global geometric features (modified Zernike moments and circularity features) and local intensity features (ordered histogram of image gradient orientations). The proposed method utilizes a supervised modular wavelet network to perform shape classification based on the extracted shape descriptors. Affine invariance is achieved using a novel eigen-based normalization approach. Furthermore, appropriate shape features are selected based on the inter- and intra-class separation indices. Therefore, the proposed classifier is robust to scale, translation, rotation and noise. Modularity is introduced to the wavelet network to decompose the complex classifier into an ensemble of simple classifiers. Wavelet network parameters are learned using an extended Kalman filter (EKF). The classification performance of proposed approaches is tested on a variety of standard symbol data sets (i.e., mechanical tools, trademark, and Oriya numerals) and the average classification accuracy is found to be 98.1% which is higher compared to other shape classifier techniques.

1 Introduction

The automatic classification and interpretation of shapes is a problem of great interest in a variety of significant areas of computer vision, notably in content based image retrieval [1, 2], target recognition, the handling of industrial parts for product assembly [3], characterization of biomedical images, quality specifications of natural products, and in agronomy [4]. The efficacy of a classifier mainly depends upon three factors: 1) the object representation scheme, 2) the selection of appropriate features, and 3) the classification scheme. Clearly, in an ideal, desirable setting a selected feature vector should have low dimensionality, be invariant to image manipulations (translation, rotation, scaling), and be insensitive to noise. At the same time, we seek a classifier with high accuracy and modest computational complexity.

A great many techniques have been developed for the description of objects, among them image projections [3], Fourier descriptors [5], template matching [6], Hough transform [7], medial axis transform [1], and Zernike moments [8]. There are a number of drawbacks in these techniques. Very briefly, moments [9] are not orthogonal bases, therefore higher order moments are very sensitive to noise. The Fourier transform [10] has inherent problems of long computation time, noise sensitivity and coordinate initialization. The Hough transform is computationally intensive for representing anything but the simplest curves. Zernike moments [9] provide good results but are sensitive to noise and scale-change.

In addition to effective feature choices, we also require an effective classifier to infer and classify shape on the basis of the selected features. Similar to features, there exist a wide variety of techniques for shape classification, including template based methods [6], syntactic and structural methods, and decision theory analysis and neural networks [5, 11]. Of these choices, the neural network has attracted a great deal of attention because of its high recognition / generalization capability and its ability to handle noisy data. The standard back-propagation supervised learning algorithm for neural networks is straightforward and provides high accuracy, however it suffers from local minima and long training time.

Although at first glance unrelated, there is a growing interest in the combining of neural and wavelet methods. Wavelet theory [12–14] is a rapidly developing branch of mathematics which has offered very efficient algorithms for approximating, estimating, analyzing and compressing nonlinear functions, not entirely unlike a neural network's ability to be considered a universal nonlinear function approximator. Because of the similarities between the discrete wavelet transform and single hidden-layer feedforward neural networks, Pradhan et al. [13] have combined wavelet methods and neural networks for power signal classification tasks.

In this paper, we propose a modular wavelet classifier based on a novel approach to feature selection and object representation, motivated from function-approximation principle of neural networks. The innovations of this paper include the use of a modular wavelet network, rather than conventional neural networks, the combining of global geometric and local intensity features for object representation. Furthermore, a extended Kalman filter approach is used to train the modular wavelet network.

The paper is organized as follows. Section 2 describes the overall approach of the proposed method. Experimental results are presented and discussed in Section 3, with conclusions and future work discussed in Section 4.

2 The Proposed Approach

The primary goal of the proposed method is to design a supervised shape-based classifier for the purpose of symbol recognition. To achieve this goal, the proposed method attempts to address issues associated with noise, affine distortions, shape discrimination, and efficiency. The proposed method is composed of three main processes: feature extraction, feature selection, and classification using a modular wavelet network.

Very briefly, the image noise is reduced using a 5×5 low-pass Gaussian mask. The denoised image is segmented by adaptive thresholding to separate the object from its background. From the segmented image, the proposed method extracts the magnitude of modified Zernike moments [10], circularity features [4], and image gradient orientation histograms [15]. The class separability [16] of each feature element is computed to select appropriate features. This feature vector is then taken as the input for the modular wavelet network. Each of these three steps is described in detail, below.

2.1 Feature extraction

The first step in the proposed method is to extract a combination of Zernike moments, image gradient orientation histograms, and circularity features. The basic idea behind combining Zernike moments with image gradient orientation histograms is that Zernike moments are, in general, global features rather than local features. As such, Zernike moments are not well-suited for recognizing symbols which are partially occluded. On the other hand, local symbol appearance and shape characteristics within an image are well-captured using image gradient orientation histograms. Furthermore, the roundness of an object (essential for trademark images) is well-captured by circularity features. Therefore, combining these features provide a more complete representation of the symbol for classification purposes. One issue with using features extracted from Zernike moments is that such features are invariant to rotation, but not to scale and translation. Therefore, an image normalization step is crucial for robust shape classification.

Image Normalization To produce Zernike moment features that are scale and translation invariant, a new image normalization method is proposed as follows. The boundary of the object is obtained using adaptive thresholding based on Otsu's method [17]. The eigenvector of the second-order covariance matrix is then determined for the extracted boundary. Finally, the length of the major axis is normalized by the magnitude of the largest eigenvector (lm). The objective of this normalization is to have the length of the major axis of the reconstructed image be a constant λ. Considering a scale factor $a = (\lambda/lm)$, the scale invariance is achieved by transforming the original image $f(x, y)$ into a new normalized image $f(x/a, y/a)$. To achieve both scale and translation invariance, the final normalized image $g(x, y)$ can be defined as follows:

$$g(x, y) = f(x', y') = f(x/a + \bar{x}, y/a + \bar{y}) \tag{1}$$

where $\bar{x}$ and $\bar{y}$ define the centroid of the original image.

Zernike Moments [8, 4] Zernike moments are orthogonal and are widely applied in image recognition. The complex Zernike moment is derived from Zernike polynomials as shown in (2), (3):

$$V_{nm}(x, y) = V_{nm}(\rho \cos(\theta), \rho \sin(\theta)) = R_{nm}(\rho) \exp(jm\theta) \tag{2}$$

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!(n+|m|)/2-s)!(n-|m|)/2-s)!} \rho^{n-2s} \quad (3)$$

where ρ is the radius from (x, y) to the shape centroid, θ is the angle between ρ and x axis, n and m are integers and subject to $n-|m| =$ even number, $|m| \leq n$. Zernike polynomials are a complete set of complex-valued orthogonal function over the unit disk $x^2 + y^2 \leq 1$. The moments are shown in (4):

$$A_{nm} = \begin{Bmatrix} \frac{n+1}{\pi} \sum_x \sum_y g(x,y) V^*_{nm}(x,y), & x^2 + y^2 \leq 1 \\ 0 & Else \end{Bmatrix} \quad (4)$$

where $*$ denotes a complex conjugate. Due to the constraint of $n - |m|$ be an even number and $m < n$, there are $n/2$ repetition of moments in each order n. Since Zernike basis functions uses the unit disk as their domain, this disk must be specified before moments are calculated.

Circularity as a global shape feature [4] We considered circularity as another important shape feature which is given by O. Considering the object as a circular disk of radius ρ, we can write O in terms of A_{40} and A_{20} as shown in (6) (7), (8).

$$A_{40} = \frac{5}{\pi} \int_0^{2\pi} \int_0^{\rho(\theta)} (6\rho^4 - 6\rho^2 + 1)\rho \partial \rho d\theta \quad (5)$$

$$A_{20} = \frac{5}{\pi} \int_0^{2\pi} \int_0^{\rho(\theta)} (2\rho^2 - 1)\rho \partial \rho d\theta \quad (6)$$

If we consider the object as a perfect circular disk of radius ρ, then O can be defined as follows:

$$\frac{A_{40}}{A20} = \frac{5}{3} \left[\frac{\int_0^x (6\rho^4 - 6\rho^2 + 1)\partial \rho}{\int_0^x (2\rho - 1)\partial \rho} \right] = \frac{5}{3} \quad (7)$$

$$\mathrm{O} = \left(\frac{3}{5}\right) \cdot \frac{A_{40}}{A_{20}} \quad (8)$$

where O is the circularity of shape. As the circularity of the shape increases, the circularity measure approaches one. Therefore, for a perfect circle, the value O will be one and for any irregular shapes will have a circularity value less than one.

Image gradient orientation histogram as a feature vector: A histogram of the edge direction is used to represent the local shape characteristics of the symbol. The image gradient information is extracted from the normalized image using a Sobel gradient operator. The corresponding gradient orientations are quantized into 36 bins of 10 degrees each. The image gradient orientation histogram of the normalized image is invariant to translation and scaling. To achieve rotational invariance, the histogram is sorted in accordance with the magnitude of its probability density function.

2.2 Feature selection

The features described in the previous section do not necessarily convey any intuitive meaning for classification [16]. Furthermore, the dimensionality of the feature vector may be too high if all of the features are used in their entirety. Therefore, a systematic evaluation of the features is important for constructing a meaningful shape descriptor. A systematic and thorough analysis of the features based on inter- and inter-class separation [16] is used in designing the proposed shape descriptor.

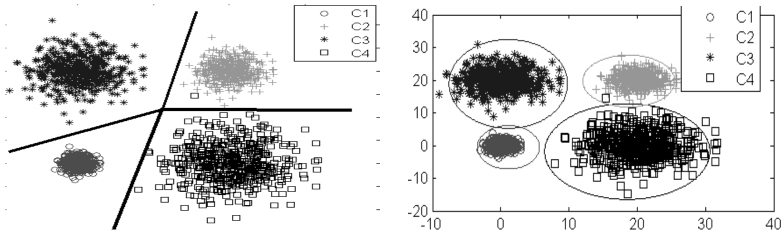
2.3 Modular wavelet network

Using the extracted shape descriptor, the symbol is classified using a modular wavelet network. The concept of modular wavelet networks is motivated by the fundamental principle of function approximations. To separate a large number of classes from each other using a conventional neural network requires a highly non-linear complex separability function. The classification boundary for a 4-class problem using a conventional multi-layer perceptron (MLP) network is shown in Fig 1(a). The proposed modular wavelet network decomposes the n-class problem into n two-class problems using n simple wavelet classifiers. Each classifier separates its assigned class from the rest of the $n-1$ classes. The classification boundary using the modular wavelet network is shown in Fig 1(b). Each ellipse in the figure shows a classification boundary which separates the assigned class from the three other classes.

One module of the modular wavelet network is shown in Fig 2, which consists of input, output and hidden layers. In the hidden layer, the input vector is first translated by a vector T and then dilated by a vector d. The radial wavelet function used in this network is a Mexican hat wavelet. The response of each hidden unit is scaled by its connecting weights (w) to the output units and then summed to produce the final output. The network output in the above structure is as:

$$y = \sum_{j=1}^{J} w_j \psi \left[d_j (X - T_j) \right] + \theta \tag{9}$$

where X, T and d are the input vector, the translation and the dilation parameter respectively, w and θ are weight and the addition parameter in the network to take care of nonzero mean function respectively, and the wavelet function ψ is given by:



(a) (b)

Fig. 1. Illustration of classification boundary for conventional neural network and the proposed modular wavelet network for a simulated cluster. (a) Classification boundary (thick black line) using conventional neural network. (b) Classification boundary using Modular wavelet network. Each ellipse shows the classification boundary of a class vs rest three other classes

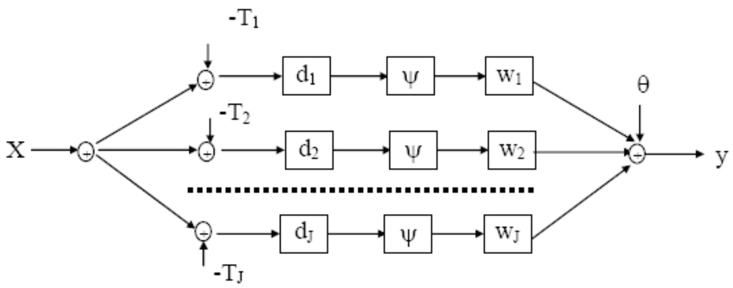


Fig. 2. One Module of Modular Wavelet network (Fig 3)with translation, dilation and wavelet function.

$$\psi(X^{'}) = (1 - \left\|X^{'}\right\|^{2})e^{\frac{\left\|x^{'}\right\|^{2}}{2}} \tag{10}$$

where $X^{'} = d(X - T)$. The complete classifier for the n-class problem is shown in Fig 3. The input shape descriptor is feed into each of the n wavelet classifier modules. Each classifier module provides a binary output to indicate whether the symbol belongs to the assigned class. The outputs of all classifier modules are fed into a simple gating module, which aggregates the results to determine the final class of the symbol. The network parameters θ, w, T and d are learned using the EKF-based training approach [13]. The EKF approach is well-suited for the proposed wavelet network as wavelet networks are non-linear in nature.

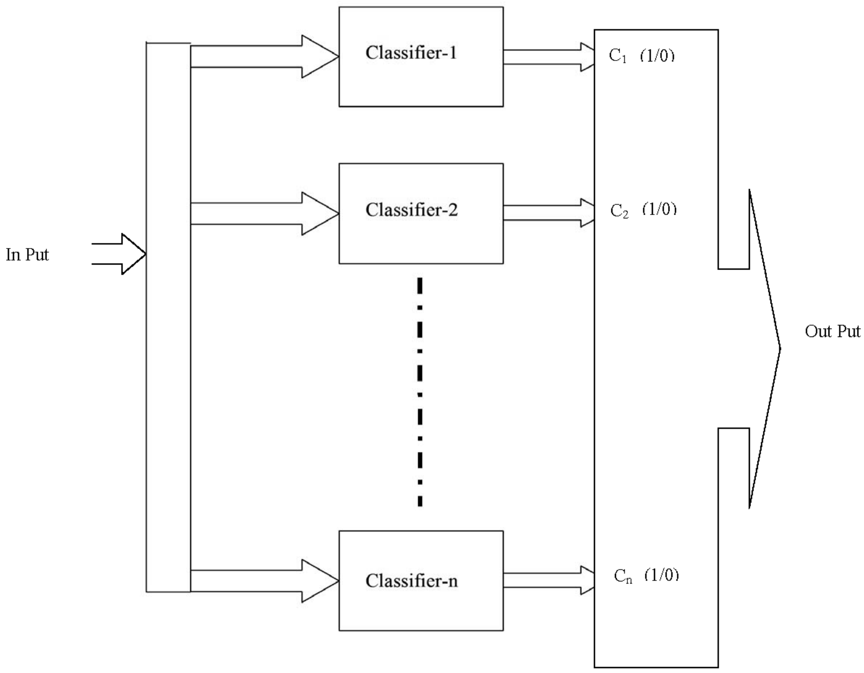


Fig. 3. The Proposed Modular Wavelet network classifier. Each module classifies one class from the rest of the $n-1$ classes.

3 Results and Discussions

3.1 Data sets

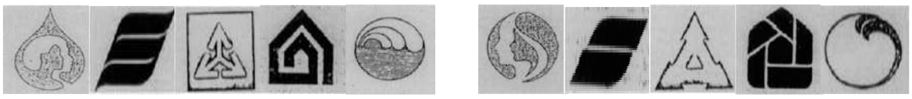
The classification performance of the proposed method is tested on 10 trademarks [1] (Fig 4), 12 general mechanical tools [3] (Fig 5) and hand written Oriya numerals [18] (Fig 6, the script for numerals used by the people of eastern India. Each image of trade mark and mechanical tool are scaled five times and at each scale 10 orientations with an interval of 15 degrees is generated. At each orientation, five noise levels (5 dB to 15 dB SNR) are added to the images to test performance under noise. Therefore, a total of 2500 trademark images and 3000 mechanical tool images are generated for testing purposes. Further, we have collected 100 Oriya numerals from 10 Oriya writers.

3.2 Classification Results

Normalization for all data is carried out by considering $\lambda = 100$. A total of 1500 trademark, 1800 mechanical tool, and 40 Oriya numeral images are used for training the modular wavelet networks.

Table 1. Average percentage classification accuracy of proposed method compared to other methods across three test case images.

-	Trademark [1]	Mechanical tool [3]	Oriya numerals
Proposed approach	100	96	98
Jain98 [1]	70	-	-
Roy06 [18]	-	-	92
Hu00 [3]	-	98	-



(a) (b)

Fig. 4. Trademark images [1]. It is important to note that each image pair (a-b) are considered too similar by the USPTO. The overall classification accuracy of the proposed method on these images are found to be 96%.

As mentioned earlier, a class separation test is conducted on the described features to determine the shape descriptor. From this experiment it was found that the magnitudes of 14 highest energy moments are of greatest importance for classification purposes. For the features selected for the proposed shape descriptor, we calculated the class separability (η) of modified Zernike moments as 94.04% , circularity as 97.04%, and gradient orientation histogram features as 69%. This shows that the features selected for the shape descriptor are sufficient for classification purposes. For the modular wavelet network, 10, 12 and 10 classifications modules were used for the trademark image, mechanical tool and Oriya numerals respectively.

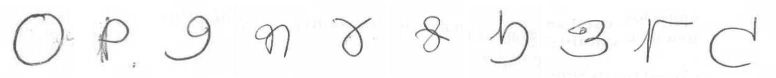
The accuracy of the proposed method is found to be 100%, 98.5% and 99% for the training data sets of the trademark, mechanical tool and Oriya numerals, respectively. The average classification accuracies of the proposed methods compared to other Jain98 [1], Hu00 [3] and Roy06 [18] on test data sets are provided on Table 1. Based on these results, it can be seen that the proposed classification algorithm can be very effective for the purpose of symbol recognition using real-world data.

4 Conclusions and Future Work

In this paper, a novel shape-based classifier using modular wavelet networks is presented for the purpose of symbol recognition. The proposed system consists of a novel shape descriptor based on global and local shape features and a wavelet shape classifier network. A novel modular wavelet network is employed for the purpose of shape classification, with EKF based training approach used

MT1 MT2 MT3 MT4 MT5 MT6 MT7 MT8 MT9 MT10MT11MT12

Fig. 5. Different mechanical tools [3]. The classification accuracy for these image using the proposed method is found to be 100%.



0 1 2 3 4 5 6 7 8 9

Fig. 6. Oriya numerals (0-9). The English equivalent corresponding to each Oriya numeral is provided below the images. The average classification accuracy of the proposed approach on Oriya numeral is found to be 98%.

to obtain faster convergence. The experimental results show that a high level of classification accuracy can be achieved by the proposed method using real-world data sets. Furthermore, the method is well-suited for parallel and online implementation due to modular concept. In the future, we plan on performing extensive testing using larger symbol data sets containing a wider variety of symbols. Furthermore, we intend on employing the above approach for the purpose of content based image and video retrieval.

5 Acknowledgements

This work has been supported in part by NSERC and by GEOIDE (Geomatics for Informed Decisions), an NSERC Centre of Excellence. We are grateful to Prof Asok Kumar Pradhan and Ayasknata Behera for assisting us in wavelet networks and EKF based training.

References

1. Jain, A., Vailaya, A.: Shape-based retrieval: A case-study with trademark image databases. Pattern Recognition **31**(9) (September 1998) 1369–1390
2. Badawy, O., Kamel, M.: Shape-based image retrieval applied to trademark images. (2004) 373–392
3. Hou, T., Pern, M.: A shape classifier by using image projection and a neural network. IJPRAI **14**(2) (2000) 225–242
4. Zhenjiang, M.: Zernike moment-based image shape analysis and its application. Pattern Recognition Letter **21**(2) (February 2000) 169–177

5. Osowski, S., Nghia, D.: Fourier and wavelet descriptors for shape recognition using neural networksa comparative study. Pattern Recognition **35**(9) (September 2002) 1949–1957
6. Choi, M., Kim, W.: A novel two stage template matching method for rotation and illumination invariance. Pattern Recognition **35**(1) (January 2002) 119–129
7. Tsai, D.: An improved generalized Hough transform for the recognition of overlapping objects. IVC **15**(12) (December 1997) 877–888
8. Khotanzad, A., H, H.Y.: Invariant image recognition by Zernike moments. IEEE Transaction on Pattern Analysis and Machine Intelligence **12**(5) (1990) 489–497
9. Wallin, A., Kubler, O.: Complete sets of complex Zernike moment invariants and the role of the pseudoinvariants. IEEE Transaction on Pattern Analysis and Machine Intelligence **17**(11) (November 1995) 1106–1110
10. Kim, W., Kim, Y.: A region-based shape descriptor using Zernike moments. **16**(1-2) (September 2000) 95–102
11. Daqi, G., Chunxia, L., Yunfan, Y.: Task decomposition and modular single-hidden-layer perceptron classifiers for multi-class learning problems. Pattern Recognition **40**(8) (2007) 2226–2236
12. Postalcioglu, S., Yasar, B.: Wavelet networks for nonlinear system modeling. Neural Computing and Applications **16**(4-5) (May 2000)
13. Pradhan, A., Routray, A., Behera, A.: Power quality disturbance classification employing modular wavelet network. Power Engineering Society General Meeting, Montreal, Que, IEEE (June 2006) 5
14. Liu, F., Luo, L.: Immune clonal selection wavelet network based intrusion detection. In: ICANN (1). (2005) 331–336
15. Postalcioglu, S., Yasar, B.: Gradient-based polyhedral segmentation for range images. Pattern Recognition Letter **24**(12) (August 2003) 2069–2077
16. Oh, S., lee, J., Ching, Y.S.: Analysis of class separation and combination of class-dependent features for handwriting recognition. IEEE Transaction on Pattern Analysis and Machine Intelligence **21**(10) (1999)
17. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man and Cybernetics **9**(1) (January 1979) 62–66
18. Roy, K., Pal, T., Pal, U., Kimura, F.: Oriya handwritten numeral recognition syste. In: ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition, Washington, DC, USA, IEEE Computer Society (2005) 770–774