SONAR TARGET RECOGNITION USING RADIAL BASIS FUNCTION NETWORKS

B.Yegnanarayana, H.M.Chouhan, and C.Chandra Sekhar
Department of Computer Science and Engineering,
Indian Institute of Technology, Madras, India.

ABSTRACT

In this paper, we consider the problem of active sonar target classification based on the targets' material composition using a Radial Basis Function (RBF) network. Sonar target responses were measured under controlled laboratory conditions in a laboratory tank. Spherical targets of different material composition were used. An important task in the design of RBF networks is the appropriate choice of the RBF centers. In this paper, we propose a Karhunen-Loeve (KL) expansion based approach for centre selection. Results on the classification performance of the RBF network trained using the KL expansion based training proc dure are provided.

I. INTRODUCTION

The problem of detection and classification of submerged targets is of considerable interest in sonars. Active sonar target recognition is based on sonar returns received from a target. In the past, both time-domain [1,2] as well as frequency-domain [1,3] techniques have been used. At the present time, there is a growing interest in the use of neural networks for the automatic recognition of sonar targets [4-6].

Previous theoretical and experimental studies [1-3] on the forward scattering of sound waves by targets of different size, shape and material composition show that information about the target's acoustic (dependent on the material composition) and geometric (dependent on the size, shape and structure) characteristics is buried in its sonar return. The target echo is composed of multiple components which are due to the target's local features (such as reflection coefficient and the radius of curvature at the specular reflection point) and global features (such as shape, composition, and size). The early part of the echo is due to specular reflection of the incident pulse. The other components of the echo can be grouped into two sets. One set of arrivals is due to the Franz waves which propagate in the surrounding medium and on the outside of the scatterer. The time separation between the specularly reflected arrival and the Franz wave arrival is related to the target dimension. The other set of waves is due to the creeping waves which travel on the inside of the scatterer. The characteristics of the creeping waves are dependent on the elastic properties of the target. The Franz wave components are highly attenuated and are of limited classification value. But, since the elastic creeping waves are strongly coupled to the surrounding medium their amplitude in the sonar return is significant. Thus it is easier to extract information regarding the composition of the target from the sonar echo. Often, sonar signals of finite duration are used to interrogate the target. For the typical sonar target sizes used, it may not be possible to separate the different components of the target return in the time domain. Alternatively, one can use the target's frequency response to derive the target information. The nulls in the frequency response correspond to the creeping waves. Thus we see that the crucial step is to extract target features from the target return. The sonar target recognition problem may then be posed as that of pattern recognition.

Recently, there has been a resurgence of interest in the use of neural networks for pattern recognition due to the fact that multilayer feedforward neural networks can realize complex nonlinear decision functions. Among the several neural classifiers that have been suggested in the past, the multilayer perceptron network (MLP) trained using backpropagation (BP) algorithm, MLP-BP, has enjoyed a wide popularity. MLP-BP networks have been used in several applications including sonar target classification [4]. While the capabilities of the multilayered perceptron have been studied widely and understood now, the performance of the BP training algorithm is far less than satisfactory. Two important limitations of the BP algorithm are: (i) it is a gradient based procedure and hence is likely to get stuck in a local minimum, giving a suboptimal solution and (ii) it requires a lot of training time involving multiple passes. To overcome the first drawback, several alternate global optimum seeking algorithms such as simulated annealing, genetic algorithm, diffusion algorithm, etc., have been suggested. Though these algorithms can provide a global solution, they are computationally more intensive than BP.

In this paper, we present the results of our experiments carried out to study the active sonar target classification performance of a Radial Basis Function (RBF) network. RBF networks have been fruitfully used in a number of applications such as speech recognition [7] and channel equalization [8]. Since training the RBF network entails solving only a set of linear equations for every node in the output layer, RBF network training requires much lesser training time and is not plagued with the problem of local minima.

In the next section we detail the design and training of a RBF network for the purpose of sonar target classification. In Section III, we describe the experiments carried out to measure the sonar target echoes and the feature extraction procedure. Finally, target classification results are discussed in Section IV.

II. RADIAL BASIS FUNCTION NETWORK

RBF classifier is a feedforward mapping network that can be used to form complex decision domains just as the multilayer perceptron. The network architecture is shown in Fig.1. RBF networks can be regarded as three layer networks with a single hidden layer. The hidden units, instead of evaluating a weighted sum of their inputs (as in the case of the MLP), encode the inputs by computing how close they are to the centres of the receptive fields. To do this, each hidden unit i has an activation function of the form $\phi(\mathbf{x}_p - \mathbf{c}_i)$, where ϕ is an appropriately chosen basis function, $\mathbf{x}_p$ is the pth input pattern vector and $\mathbf{c}_i$ is a vector representing the ith RBF centre. Assuming the output layer to be linear, the activation of the jth output node is obtained as

$$y_{jp} = \sum_{i=1}^{N} w_{ij}\, \phi(\mathbf{x}_p - \mathbf{c}_i), \tag{1}$$

where N is the number of nodes in the hidden layer and w_{ij} is the connection weight between the ith hidden node and the jth output node. Letting

$$h_{ip} = \phi(\mathbf{x}_p - \mathbf{c}_i), \tag{2}$$

we can write Eq.(1) as

$$y_{jp} = \mathbf{h}_p \mathbf{w}_j, \tag{3}$$

where $\mathbf{h}_p = [h_{1p}\ h_{2p} \ldots h_{Np}]$ and $\mathbf{w}_j = [w_{1j}\, w_{2j} \ldots w_{Nj}]^T$ (T denotes transpose). For a network with M output nodes, the network output vector $\mathbf{y}_p$ can be expressed as

$$\mathbf{y}_p = \mathbf{h}_p \mathbf{W}, \tag{4}$$

where $\mathbf{W} = [\mathbf{w}_1 | \mathbf{w}_2 | \ldots | \mathbf{w}_M]$, $\mathbf{y}_p = [y_{1p}\ y_{2p} \ldots y_{Mp}]$ and $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_M$ are the weight vectors for the different output nodes. For P input patterns, stacking $\mathbf{h}_p$ in $\mathbf{H}$, i.e., $\mathbf{H} = [\mathbf{h}_1 | \mathbf{h}_2 | \ldots | \mathbf{h}_P]^T$, the weight vectors can be obtained as

$$\mathbf{W} = \mathbf{H}^+ \mathbf{Y}, \tag{5}$$

where $\mathbf{Y} = [\mathbf{y}_1 | \mathbf{y}_2 | \ldots | \mathbf{y}_P]^T$ and + denotes the pseudo-inverse. In practice then, one can obtain the weights using the singular value decomposition or an iterative method such as the LMS algorithm.

Note that the data centres $\mathbf{c}_i$ and their number N, the basis functions ϕ, and the parameters of the basis functions, if any, are design parameters that need to be chosen carefully, depending on the application on hand, to obtain the required performance. Thus two important issues that need to be addressed are: (i) which basis function is suitable for a classifier design and how to choose its parameters, if any, and (ii) how to choose the data centres $\mathbf{c}_i$ and their number N.

With regard to the choice of the RBF activation functions, the use of Gaussians, thin-plate-splines and multiquadratic functions has been reported. For the classifier application, addressed herein, we have chosen the Gaussians as in [8].

For Gaussian RBFs,

$$h_{ip} = \phi(\mathbf{x}_p - \mathbf{c}_i) = \exp\left[-\sum_{n=1}^{L}\left[\frac{(x_{pn} - c_{in})^2}{2\sigma_{in}^2}\right]\right] \tag{6}$$

requires the specification of the spread of the Gaussian functions, σ_{in}^2. While Tsoi [9] made an ad hoc choice, Renals [7] chose the spread based on a distance measure depending on the type of the input feature vector.

The question of RBF centre selection has been addressed by Tsoi [9] and Chen et al. [8]. One approach to the centre selection is to choose centre for every data or a subset of data in the data set. Such an approach not only requires that a large number of centres be chosen to adequately sample the input data space, but also has the disadvantage that if the centres are close, it may lead to near linear dependency of those centres and in turn to numerical illconditioning problem. Similar ad hoc schemes were also adopted by Renals [10] and Chen et al. [8]. Another approach is to find clusters of the input data and locate the RBF centres at the cluster centres [11]. Recently Chen et al. [8] have proposed the use of orthogonal least squares method for selecting the centres. Further they suggested the use of AIC criteria for selecting the number of centres.

In this paper, we shall adopt an approach based on Karhunen-Loeve (KL) expansion of the feature vector covariance matrix, for RBF centre selection. Our approach is as follows. Let

$$\mathbf{x}_{mk} = (x_{k1}, x_{k2}, \ldots, x_{kL}), \tag{7}$$

denote the kth feature vector due to the kth input pattern vector from the mth class. The feature data covariance matrix of all the K feature vectors belonging to the mth class can be obtained as

$$\mathbf{X}_m = (1/K)\sum_{k=1}^{K}(\mathbf{x}_{mk} - \bar{\mathbf{x}}_m)(\mathbf{x}_{mk} - \bar{\mathbf{x}}_m)^T \tag{8}$$

where $\bar{\mathbf{x}}_m$ is the mean vector of the mth class. The covariance of feature data from all M classes is given by

$$\mathbf{X} = (1/M)\sum_{m=1}^{M}\mathbf{X}_m . \tag{9}$$

The eigen decomposition of $\mathbf{X}$ can be expressed as

$$\mathbf{X} = \mathbf{E}\Lambda\mathbf{E}^T , \tag{10}$$

where Λ is a diagonal matrix with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_L$ of $\mathbf{X}$ and $\mathbf{E}$ is the corresponding eigenvector matrix. Further, we will assume that the eigenvalues are arranged in non-increasing order. The significance of this decomposition in the context of pattern recognition can be explained as follows. The eigenvector corresponding to the largest eigenvalue captures most of the characteristics of the input feature data. The other eigenvectors vectors capture less information compared to the first depending on the magnitude of their corresponding eigenvalues. Furthermore, since the eigenvectors are orthonormal they capture mutually exclusive information. Thus we may choose the eigenvectors as our RBF network centers.

The other question is : How many of the L eigenvectors are essential to capture most of the information in the input feature vectors in order to correctly identify their class membership. It must be noted that for an M class problem, at least M RBF centers are necessary. In order to reduce the size of the network, we may choose only a subset of the eigenvectors. The smaller the subset of eigenvectors chosen, the larger will be the mean square error in representation. Since our emphasis here is on the correct classification rather than representation of the input feature vectors we may choose a smaller subset. One possible approach to choosing a subset of the eigenvectors is to use a threshold on the eigenvalues to select the eigenvectors. An alternative approach is to use Aikaike's Information Criteria (AIC). We have used the simple threshold scheme in our experiments.

III. DATA COLLECTION AND FEATURE EXTRACTION

Sonar target responses were measured under controlled conditions in a laboratory tank of dimension 10m x 3m x 2m. Spherical targets of different material composition, namely, wood, aluminium, iron and rubber were used. Since our emphasis was on classification based on the material composition, all targets used were of the same size. Though the tank walls are anechoic, the transmitter, the target and the receiver were positioned at the centre of the tank to minimize any reflections from the side walls. A circular disc type transmitter of 12^o beamwidth at its resonant frequency of 110 kHz was used. Gated five-cycle sinewave signals of centre frequency 110 kHz were amplified and fed to the transmitter. Echoes from the targets were received using a B&K 8103 omnidirectional transducer. The received echoes were digitized and stored in a TEK 2220 storage oscilloscope, which were then transferred over the GPIB to an IBM-PC compatible for further processing.

The acquired target responses were of very high signal-to-noise ratio (SNR). Several sets of target responses were measured from which it was possible to identify a reference echo for each of the targets (Fig.2). To these reference echoes, white Gaussian noise was added to simulate noisy sonar returns with 20dB, 25dB, and 30dB. Out of 200 such noisy sonar returns synthetically generated for each of the targets, 60% were used for training the network and the rest were used for testing the performance of the network.

The feature vectors for input to the network were obtained using the frequency domain characterization. Each of the target responses was Fourier transformed using a 2048-point FFT. The squared magnitude spectrum of the sonar return was divided by the squared magnitude spectrum of the 5-cycle sinewave input to obtain the transfer function of the targets. Furthermore, in order to reduce the dimensionality of the feature vector, we divided the frequency interval into 15 bins of equal size. The average energy in each of the bins was calculated. Each of the resulting vectors was normalized and used as input to the network.

IV. RESULTS AND DISCUSSION

Classification experiments were carried out using a RBF network with a one-out-of-N type of output coding. The network was a three layer feed-forward network. The input layer had 15 nodes corresponding to the 15 elements of the input feature vector. The output layer had 4 nodes to represent the four different sonar target classes. The centers of the RBF network were chosen using the KL transform based procedure detailed in the previous section. In all the experiments it was found adequate to use just 4 nodes with the centres being the eigenvectors corresponding to the first four dominant eigenvalues. The widths of the Gaussians were chosen using a n-nearest neighbour rule.

The dissimilarity measures, shown in Table 1, are indicative of the degree of difficulty of the sonar target classification problem under study. Different inter-class centroid separations are given in Table 1a and the intra-class variations are given in Table 1b. As expected the dissimilarity measures diminish with a decrease in the SNR and also the intra-class variations increase with a decrease in the SNR. Also the between class separation for the target pairs (aluminum-wood) and (rubber,cork) is less compared to the target pairs (aluminum,rubber), (aluminum,cork), (wood,rubber) and (wood,cork). This is again as expected from the targets' acoustic properties since rubber and cork are both acoustically soft materials.

The results of the classification experiments, for the different SNRs, are given in Table 2. The values given represent the average performance of the classifier, averaged over trials with different sets of synthetically generated noisy data. The classifier performance as a function of number of training samples, for 20dB SNR, is shown in Fig.3. While in the 25dB SNR case, near exact performance could be achieved only after training with 50 samples, for 30 dB SNR similar performance could be achieved with much fewer samples. These performance results corroborate with the observations made from the dissimilarity measures.

In conclusion, we note the following. The choice of Gaussian nodes seems to be appropriate for a classifier design. The training method suggested herein is robust and did not break down due to any numerical illconditioning problems in any of the cases studied herein. Furthermore, the satisfactory performance of the network with the minimal number of nodes required for the sonar target classification problem is indicative of the fact that most of the features of the input data are captured in the first few significant KL components. The advantage of much shorter training time required to train the RBF network in addition to its real-time classification capabilities (shared by other neural networks) makes it a very useful tool for sonar target classification. Ofcourse, in contrast to the batch mode training scheme suggested herein, it would be more useful if a suitable adaptive on-line learning algorithm can be developed.

ACKNOWLEGEMENT

This work was supported by the Department of Electronics, Government of India.

REFERENCES

[1] C.W.Horton, "A review of reverberation, scattering, and echo structure," J.Acoust.Soc.Am.51, 1049-1061 (1972).

[2] J.F.Hoffman, "Classification of spherical targets using likelihood ratios and quadrature components," J.Acoust.Soc.Am.49, 23-30 (1971).

[3] P.Chestnut, H.Landsman and R.W.Floyd, "A sonar target recognition experiment," J.Acoust.Soc.Am.66, 140-147 (1979).

[4] R.P.Gorman and T.J.Sejnowski, "Learned classification of sonar targets using a massively parallel network," IEEE Trans. Acoust.,Speech,Signal Processing 36, 1135-1140 (1988).

[5] H.L.Roitblat, P.W.B.Moore, P.E.Natchtigall, R.H.Penner and W.W.L.Au, "Dolphin echolocation : Identification of return echoes using a counter propagation network," Proc. Intl. Joint Conf. on Neural Networks, vol 1, 295-300 (1989).

[6] T.B.Hayley, "Applying neural networks for automatic active sonar classification," Proc. IEEE Intl. Conf. on Pattern Recognition, 41-44 (1990).

[7] S.Renals and R.Rohwer, "Phoneme classification experiments using radial basis functions," Proc. Intl. Joint Conf. on Neural Networks, vol 1, 461-467 (1989).

[8] S.Chen, C.F.N.Cowan and P.M.Grant, "Orthogonal least squares learning algorithm for radial basis function networks," IEEE Trans. Neural Networks,302-309 (1991).

[9] A.C.Tsoi, "Multilayer perceptron trained using radial basis functions," Electronics Letters 25, 1296-1297 (1989).

[10] S.Renals, "Radial basis function network for speech pattern classification," Electronics Letters 35, 437-439 (1990).

[11] S.Beck, L.Deuser, R.Still, and J.Whitely, "A hybrid neural network classifier of short duration acoustic signals," IJCNN-91, I-119 - I-123 (1991).

Table 1a. Inter Class Separation

30 dB	AL	CR	RU	WD
AL	---	0.7616	0.9653	1.2232
CR	0.7616	---	0.7139	1.0393
RU	0.9653	0.7139	---	0.7900
WD	1.2232	1.0393	0.7900	---
25 dB	**AL**	**CR**	**RU**	**WD**
AL	---	0.7747	0.9489	1.2217
CR	0.7747	---	0.7117	1.0337
RU	0.9489	0.7117	---	0.7845
WD	1.2217	1.0337	0.7845	---
20 dB	**AL**	**CR**	**RU**	**WD**
AL	---	0.8255	0.8297	1.1989
CR	0.8255	---	0.8556	1.0226
RU	0.8297	0.8556	---	0.8730
WD	1.1989	1.0226	0.8730	--

Table 1b. Inter Class Variance

	AL	CR	RU	WD
30 dB	0.0490	0.0284	0.0496	0.0290
25 dB	0.1394	0.0812	0.1395	0.0953
20 dB	0.3839	0.2271	0.3990	0.2351

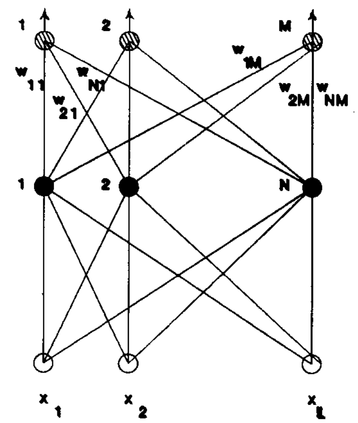

Fig 1. Schematic of RBF network

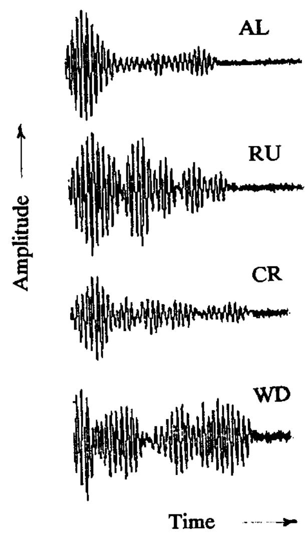

Fig 2. Backscattered pulses from different targets.
(AL) aluminium (RU) rubber (CR) cork (WD) wood

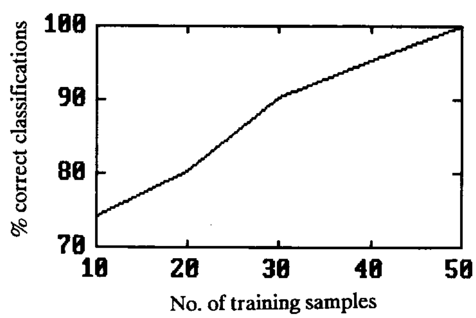

Fig 3. Network learning curve.