

Recognizing targets from infrared intensity scan patterns using artificial neural networks

Tayfun Aytaç
TÜBİTAK ÇEKAE/İLTAREN
Şehit Yzb. İlhan Tan Kışlası
Ümitköy, TR-06800 Ankara
Turkey

Billur Barshan
Bilkent University
Department of Electrical and Electronics
Engineering
Bilkent TR-06800 Ankara
Turkey
E-mail: billur@ee.bilkent.edu.tr

Abstract. This study investigates the use of simple, low-cost infrared sensors for the recognition of geometry and surface type of commonly encountered features or targets in indoor environments, such as planes, corners, and edges. The intensity measurements obtained from such sensors are highly dependent on the location, geometry, and surface properties of the reflecting target in a way that cannot be represented by a simple analytical relationship, therefore complicating the localization and recognition process. We employ artificial neural networks to determine the geometry and the surface type of targets and provide experimental verification with three different geometries and three different surface types. The networks are trained with the Levenberg–Marquardt algorithm and pruned with the optimal brain surgeon technique. The geometry and the surface type of targets can be correctly classified with rates of 99 and 78.4%, respectively. An average correct classification rate of 78% is achieved when both geometry and surface type are differentiated. This indicates that the geometrical properties of the targets are more distinctive than their surface properties, and surface determination is the limiting factor in recognizing the patterns. The results demonstrate that processing the data from simple infrared sensors through suitable techniques can help us exploit their full potential and extend their usage beyond well-known applications. © 2009 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3067874]

Subject terms: simple infrared detectors; target classification; target differentiation; artificial neural networks; optimal brain surgeon; pattern recognition.

Paper 080450R received Jun. 8, 2008; revised manuscript received Nov. 20, 2008; accepted for publication Nov. 24, 2008; published online Jan. 30, 2009.

1 Introduction

Artificial neural networks have been employed efficiently as pattern classifiers in numerous applications.¹ These classifiers are nonparametric and make weaker assumptions on the shape of the underlying distributions of input data than traditional statistical classifiers. Therefore, they can prove more robust when the underlying statistics are unknown or the data are generated by a nonlinear system. Performance of neural network classifiers is affected by the choice of parameters of the network structure, training algorithm, and input signals, as well as parameter initialization.^{2,3} In this article, we achieve robust target differentiation by processing the data acquired from simple infrared detectors using multilayer artificial neural networks.

Target differentiation is of considerable interest for intelligent systems that need to interact with and operate in unknown environments autonomously. Such systems rely on sensor modules, which are often their only available source of information. Because the resources of such systems are limited, the available resources should be used in the best way possible. It is desirable to maximally exploit the capabilities of lower-cost sensors before more costly and sophisticated sensors with higher resolution and higher resource requirements are employed. This can be achieved

by employing better characterization and physical modeling of these sensors as well as processing the data they provide using suitable and effective techniques.

Although ultrasonic sensors have been widely used for object detection and ranging,⁴ they are limited by their large beam width and the difficulty of interpreting their readings due to specular, higher-order, and multiple reflections from surfaces. In addition, many readily available ultrasonic systems cannot detect objects up to 0.5 m, which corresponds to their blank-out zone. Therefore, in performing tasks at short distances from objects, use of inexpensive, practical, and widely available sensors such as simple infrared detectors are preferable to employing ultrasonic sensors or more costly laser and vision systems. Furthermore, in a sensor-fusion framework, infrared sensors would be perfectly complementary to these systems, which are not suitable for close-range detection.⁵⁻⁹ Infrared detectors offer faster response times and better angular resolution than ultrasonic sensors and provide intensity readings at nearby ranges (typically from a few centimeters up to a meter). The intensity versus range characteristics are nonlinear and dependent on the properties of the surface and environmental conditions. Consequently, a major problem with the use of simple infrared detectors is that it is often not possible to make accurate and reliable range estimates based on the value of a single intensity return, because the return depends on both the geometry and surface properties of the

encountered object. Likewise, the surface properties and the geometry of the target cannot be deduced from simple intensity returns without knowing its position and orientation.

Due to single intensity readings not providing much information about the target properties, recognition capabilities of infrared sensors have been underestimated and underused in most of the earlier work. To achieve accurate results with these sensors, their nonlinear characteristics can be analyzed and physically modeled based on experimental data. In addition, the data acquired from such simple infrared sensors should be processed effectively through the use of suitable techniques. Armed with such characterization, modeling, and suitable processing techniques, their potential can be more fully exploited and their usage can be extended beyond simple tasks such as counting and proximity detection. The aim of this study is to explore the limits and maximally realize the potential of these simple sensors so that they can be used for more complicated tasks such as differentiation, recognition, clustering, docking, perception of the environment, and mapping. For this purpose, we employ artificial neural networks to classify targets with different geometries and surface properties. We can differentiate a moderate number of targets and/or surfaces commonly encountered in indoor environments using a simple infrared system consisting of one emitter and one detector. The results indicate that by processing the data acquired from such simple infrared sensors effectively through the use of suitable techniques, substantially more information about the environment can be extracted than is commonly achieved with conventional usage.

This paper is organized as follows. In Section 2, we provide a literature survey of related earlier work on infrared sensing. A brief review of our recent work on differentiation with ultrasonic and infrared sensors is given in Section 3. Descriptions of the artificial neural network structure, training, and pruning algorithms are given in Section 4. In Section 5, experimental verification is provided. The infrared sensor, the experimental setup, and the procedure used in the experiments are described. Results for geometry and surface-type differentiation are presented. Concluding remarks are made and directions for future research are given in the last section.

2 Background on Infrared Sensing

The use of infrared sensors in the pattern recognition area has been mostly limited to the recognition or detection of features or targets in conventional 2-D images.¹⁰ Examples of work in this category that also involve the use of neural networks mostly focus on automatic target recognition (see Ref. 11 for a survey). In Ref. 12, the back-propagation algorithm is used for detecting small-sized targets in infrared images with highly cluttered backgrounds. Probabilistic neural networks are used to discriminate between aircraft and flares for infrared imaging seekers for counter-counter measurement purposes.¹³ Intensity and shape features of aircraft and flares are used as input to the network. A correct differentiation rate of 98% is achieved. In Refs. 14 and 15, the authors propose a multistage target detector for localization of targets in infrared images. In Ref. 16, a neural-network-based point target detection method is proposed in

single-frame infrared images with highly cluttered backgrounds. Other applications where infrared images are given as input to neural networks include target tracking,¹⁷ automatic vehicle detection,¹⁸ and face identification for biometric systems.¹⁹

We note that the position-invariant target differentiation achieved in this paper is different from such operations performed on conventional images in that here we work not on direct “photographic” images of the targets obtained by some kind of imaging system but rather on angular intensity scans obtained by rotating a point sensor. The targets we differentiate are not patterns in a 2-D image but rather objects in space, exhibiting depth, whose geometry and surface characteristics we need to identify. For this reason, position-invariant pattern recognition achieved in this paper is different from such operations performed on conventional images.

Besides infrared cameras that produce 2-D images, simple infrared sensors that usually consist of a single emitter/detector pair have also been used in safety and security systems, process control, machine vision, and robotics and automation.²⁰ More specifically, they have been used in simple object and proximity detection, counting, distance and depth monitoring, floor sensing, door detection, monitoring doors/windows of buildings and vehicles, light curtains for protecting an area, position control, and obstacle/collision avoidance.

In Ref. 21, the properties of a planar surface at a known distance have been determined using the Phong illumination model. Using this information, the infrared sensor employed has been modeled as an accurate range finder for surfaces at short ranges. A number of commercially available infrared sensors are evaluated in Ref. 22. Reference 23 describes a passive 2-D infrared array capable of identifying the locations of the people in a room. Infrared sensors have also been used for automated sorting of waste objects made of different materials.²⁴ In Ref. 25, an infrared system that can measure distances up to 1 m is described. Reference 26 deals with optical determination of depth information. In Ref. 27, simulation and evaluation of the recognition abilities of active infrared sensor arrays is considered for autonomous systems using a ray-tracing approach.

In Ref. 28, the authors develop a novel range estimation technique which is independent of surface type, because it is based on the position of the maximum intensity value instead of surface-dependent absolute intensity values. An intelligent feature of the system is that its operating range is made adaptive based on the maximum intensity of the detected signal.

In the thesis work in Ref. 29, infrared sensors are used for position estimation. Reflectance from spherical objects is modeled by considering the position, orientation, and the characteristics of the emitter and detector, the position, size, and reflectivity of the spherical object, and the intensity of the reflected light. A 3-D position estimation of objects is performed with the help of a nontouch screen. A 2-D position estimation is implemented using an electrically powered wheelchair whose motion is controlled by head movements.

3 Review of Our Recent Work on Target Differentiation

In our recent work, we have been differentiating target types using different sensing modalities. First, we investigated the processing of ultrasonic signals using neural networks for robust differentiation of commonly encountered features in indoor environments.³⁰ We showed that neural networks can differentiate more target types employing only a single sensor node with a higher correct differentiation rate (99%) than achieved with previously reported techniques (61–90%) employing multiple sensor nodes. In Refs. 30 and 31, we investigated the preprocessing of input ultrasonic signals to neural networks using various signal transformations. In these works, only the differentiation of geometry types of targets was employed.

Encouraged by the successful differentiation rate achieved with ultrasonic sensors, we next attempted to perform differentiation with infrared sensors. We first employed a rule-based approach which is based on extracting empirical rules by inspecting the nature of the infrared intensity scans.³² Second, we employed a template-based approach, based on comparing the acquired infrared intensity scans with previously stored templates acquired from targets with different properties. This approach relies on the distinctive natures of the infrared intensity scans and requires the storage of a complete set of reference scans of interest. We considered the following different cases: targets with different geometrical properties but made of the same surface material,³³ targets made of different surface materials but of the same planar geometry,³⁴ and targets with both different geometry and surface properties,³⁵ generalizing and unifying the results of Refs. 33 and 34. In parametric surface differentiation reported in Ref. 36, only the reflection coefficients obtained using a physical reflection model are considered as parameters and used in the differentiation process, instead of using the complete infrared intensity scans as in the previous differentiation approaches. Finally, we extended the parametric surface differentiation approach proposed in Ref. 36 to the differentiation of both the geometry and surface type of the targets using statistical pattern recognition techniques.³⁷

4 Artificial Neural Networks (ANNs)

In this study, we consider the use of ANNs to identify and resolve parameter relations embedded in infrared intensity scan patterns acquired from target types of different geometry, possibly with different surface properties, for their differentiation in a robust manner. This is done in two stages, where the first stage consists of determining the target geometry and the second stage involves determining the surface type of the target.

Multilayer ANNs consist of an input layer, one or more hidden layers to extract progressively more meaningful features, and a single output layer, each comprised of a number of units called neurons. The model of each neuron includes a smooth nonlinearity, which is called the activation function. Due to the presence of distributed nonlinearity and a high degree of connectivity, theoretical analysis of ANNs is difficult. These networks are trained to compute the boundaries of decision regions in the form of connection weights and biases by using training algorithms. Per-

formance of ANNs is affected by the choice of parameters related to the network structure, training algorithm, and input signals, as well as parameter initialization.³

4.1 ANN Structure and Parameters

The ANN used in this paper consists of one input, one hidden, and one output layer, with 160, 10, and 3 neurons, respectively. The numbers for the input and hidden layers both include the bias values of 1. The hyperbolic tangent function of the form $\varphi(v) = (1 - e^{-2v}) / (1 + e^{-2v})$, illustrated in Fig. 1, is used as the activation function for all the neurons. The output neurons can take continuous values between -1 and 1 , and the decision at the output is made based on a maximum selection scheme, also known as the winner-take-all approach. The structure of the ANN is given in Fig. 2.

It is important to determine the optimal network structure with respect to the correct differentiation rate and network complexity. The infrared scan patterns are provided as input to the ANN after being downsampled by 10 to reduce the complexity of the network (the number of connection weights between the input and hidden layers). This sampling rate is chosen such that the patterns preserve their shapes and no identifying information is lost. Our different trials indicate that inclusion of more samples of the original scan patterns does not improve the differentiation accuracy. Fully connected ANNs are trained starting with different initial conditions, different weight factors, and different numbers of neurons in the hidden layer. Both modular and nonmodular structures have been considered.

4.2 Training the Neural Network

Two training algorithms are employed to begin with, namely, back-propagation algorithm (BPA) and Levenberg–Marquardt algorithm (LMA). A set of training patterns is presented to the network. The aim is to minimize the sum-squared-error criterion function between the resulting signal at the output and the desired signal:

$$\mathcal{E}(w) = \frac{1}{2N} \sum_{i=1}^N \sum_{j=1}^3 (d_{ij} - z_{ij})^2 \quad (1)$$

where w is the weight vector, d_{ij} and z_{ij} are the desired and actual output values for the i 'th training pattern and the j 'th output, and N is the number of training patterns.

With the BPA, the error criterion function is minimized with a gradient-descent procedure. Because the results of training with BPA were not satisfactory, training of the ANN is performed by the LMA, which is more robust and converges in a shorter time than BPA. LMA is a very fast training algorithm based on a damped Gauss–Newton method for the solution of the nonlinear least-squares problems. The original description of the LMA is given in Refs. 38 and 39. The application of the LMA to neural network training is described in Refs. 3 and 40. This algorithm appears to be the fastest method for training moderate-sized feedforward neural networks. The main drawback of the LMA is that it requires the storage of some matrices that can be quite large for certain problems. The memory requirements are proportional to the number of weights in the network.

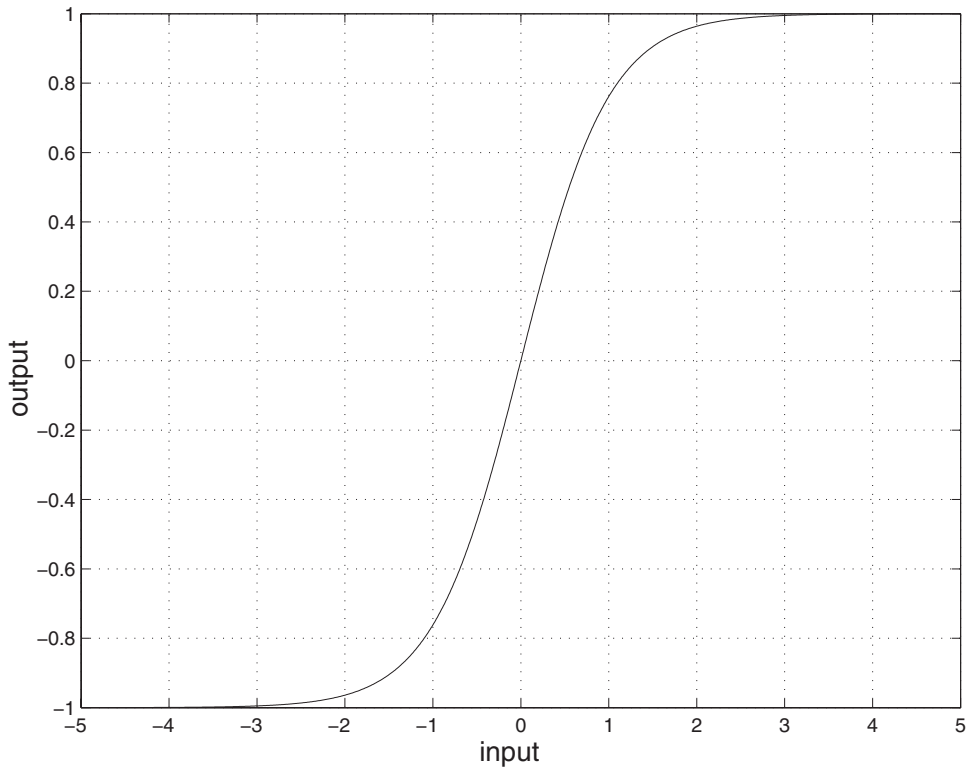


Fig. 1 The hyperbolic tangent function, used as the activation function for the ANN.

The Hessian matrix $H \triangleq \partial^2 \varepsilon / \partial w^2$ involved in neural-network minimization problems is often ill-conditioned, making the problem harder to solve. The LMA is suitable for such problems, because it is designed to approach second-order training speed without having to compute the Hessian matrix. When the error criterion function has the form of a sum of squares [as in Eq. (1)], the Hessian matrix can be approximated as $H \cong J^T J$ and the gradient can be computed as $J^T e$, where $J \triangleq \partial \varepsilon / \partial w$ is the Jacobian matrix that contains the first derivatives of the network errors with respect to the weights and biases and e is a vector of net-

work errors. The Jacobian matrix can be computed through a standard back-propagation technique⁴⁰ that is much less complex than computing the Hessian.

The LMA uses the following Newton-like update, where a diagonal matrix is added to the approximation to the Hessian matrix:

$$w_{k+1} = w_k - [J^T J + \mu I]^{-1} J^T e. \tag{2}$$

Here, w_k is the weight vector at the k 'th iteration and I is the identity matrix. The size of the diagonal matrix added to the Hessian matrix is adjusted with the learning-rate parameter μ . LMA is a hybrid of gradient-descent and Gauss-Newton relaxation methods. Large values of μ produce parameter update increments primarily along the negative gradient direction (gradient descent), while small values result in updates governed by the Gauss-Newton method. When the scalar μ is zero, this is simply the Gauss-Newton method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. The Gauss-Newton method is faster and more accurate near an error minimum, so the aim is to shift toward this method as quickly as possible. Thus, the value of μ is chosen adaptively to produce a downhill step. The method used in the implementation of the LMA differs from that proposed in Ref. 38 in that the size of the elements of the diagonal matrix added to the approximated Hessian is adjusted according to the size of the ratio between actual decrease and predicted decrease in the error function.^{41,42} If this ratio is greater than 0.75, the parameter μ is halved; if the ratio is smaller than 0.25, μ is doubled. Next, it is checked if there is a decrease in the value of the error

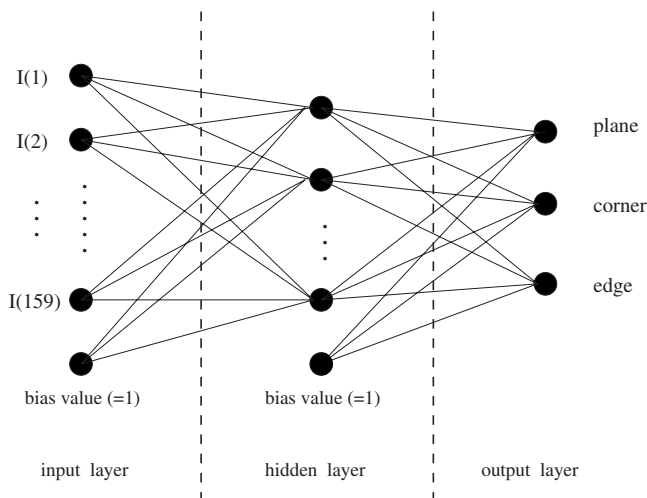


Fig. 2 The structure of the ANN before pruning.

criterion function defined in Eq. (1). If the error function decreases for the given step size and direction, then the iteration is performed, and the value of μ is not changed. Because the value of μ is adjusted adaptively, its initial value is not particularly critical and only influences the initial convergence rate; if it is too large, the algorithm takes small steps, and if it is too small, the algorithm increases it until small enough steps are taken.

The training process is terminated if a certain precision goal is reached or if the specified maximum number of iterations (100,000) is exceeded, whichever occurs earlier. The latter case occurs very rarely. The acceptable error level is set to a value of 10^{-3} .

When the ANNs are trained with the LMA, a weight decay factor is used for regularization.⁴³ Weight decay adds a penalty term to the error function to penalize large weights that can cause excessive variance at the output. The weight decay factor in our implementation is chosen as 10^{-4} . By choosing a sufficiently small factor, one can reduce the average generalization error. If the weight-decay factor is too small ($<10^{-4}$), it will take a long time to converge to the desired accuracy. For greater weight-decay factors, the ANN may not converge to the desired accuracy.

4.3 Pruning with Optimal Brain Surgeon

After training the fully connected network, the network structure is further optimized by pruning the weights. Pruning is a well-known method for determining the number of hidden layer neurons in feed-forward neural networks. After training with a relatively large number of hidden layer neurons, some of the neurons and weights are possibly eliminated according to some criterion. The network is trained until it has the minimum number of weights and hidden-layer neurons for a given error tolerance level.

Pruning can be done by eliminating the weights having the smallest magnitudes, but the resulting ANNs may not be optimal as the weights with smaller magnitudes may be important for the training.¹ In this paper, the optimal brain surgeon (OBS) technique⁴⁴ is employed for finding the optimal network structure. OBS is a sensitivity-based weight-pruning technique that makes use of second-order approximation of the error criterion function $\mathcal{E}(w)$, defined in Eq. (1), for evaluating the effect of the weights on the training error. The network is trained to a local minimum, and the weight resulting in the smallest increase in the training error is pruned each time. OBS not only removes the weights but also readjusts the remaining weights optimally.

The Taylor series expansion of the error criterion function around a local minimum is given by

$$\delta\mathcal{E} = \left(\frac{\delta\mathcal{E}}{\delta w}\right)^T \cdot \delta w + \frac{1}{2} \delta w^T \cdot \frac{\delta^2\mathcal{E}}{\delta w^2} \cdot \delta w + \mathcal{O}(\|\delta w\|^3). \quad (3)$$

Because it is assumed that the network has converged at least to a local minimum of the error function, the first term in Eq. (3) is approximately zero. The third and higher-order terms, represented by the last term in the equation, are neglected. Using the definitions of the Jacobian and the Hessian matrices from Section 4.2, the equation can be rewritten as

$$\delta\mathcal{E} = \mathbf{J}^T \cdot \delta w + \frac{1}{2} \delta w^T \cdot \mathbf{H} \cdot \delta w + \mathcal{O}(\|\delta w\|^3) \quad (4)$$

$$\cong \frac{1}{2} \delta w^T \cdot \mathbf{H} \cdot \delta w. \quad (5)$$

The elimination of a single weight, say the q 'th weight w_q , can be formulated as a constrained optimization problem with the constraint

$$\mathbf{u}_q^T \cdot \delta w + w_q = 0, \quad (6)$$

where \mathbf{u}_q is the unit vector along the q 'th direction (a unit vector with all components except the q 'th one being zeros). Thus, the objective of OBS is to minimize Eq. (5) subject to the above constraint

$$\min_q \left\{ \min_{\delta w} \left(\frac{1}{2} \delta w^T \cdot \mathbf{H} \cdot \delta w \right) \mid \mathbf{u}_q^T \cdot \delta w + w_q = 0 \right\}. \quad (7)$$

We form the Lagrangian

$$L = \frac{1}{2} \delta w^T \cdot \mathbf{H} \cdot \delta w + \lambda (\mathbf{u}_q^T \cdot \delta w + w_q), \quad (8)$$

where λ is a Lagrange multiplier, and solve the problem by the method of Lagrange multipliers to get

$$\delta w = - \frac{w_q}{[\mathbf{H}^{-1}]_{qq}} \mathbf{H}^{-1} \cdot \mathbf{u}_q, \quad (9)$$

where δw is the optimal weight change and $[\mathbf{H}^{-1}]_{qq}$ is the q 'th diagonal element of the inverse Hessian matrix. After the weight w_q is pruned, the resulting change in error is

$$L_q = \frac{1}{2} \frac{w_q^2}{[\mathbf{H}^{-1}]_{qq}}, \quad (10)$$

where L_q is called the ‘‘saliency’’ of weight q —the increase in the error that results when the weight w_q is deleted. OBS procedure is outlined below:⁴⁵

1. Train a reasonably large network to minimum error.
2. Compute \mathbf{H}^{-1} .
3. Find the q that gives the smallest saliency and compute L_q .
4. If L_q is much less than a preset error bound, delete the q 'th weight and proceed to step 5; otherwise, proceed to step 6.
5. use the q from step 3 to update all the weights using L_q .
6. If no more weights can be eliminated without a large increase in training error, retrain the network.

In our implementation, when 5% of the weights are pruned, the network is retrained within a maximum of 50 iterations. (Retraining can also be done each time one of the weights is pruned. However, this is a very time-consuming process.) At each retraining step, the ANN is tested with the test data and the error and the corresponding weights are stored. The pruned network resulting in the minimum test error is chosen as the optimal one and is

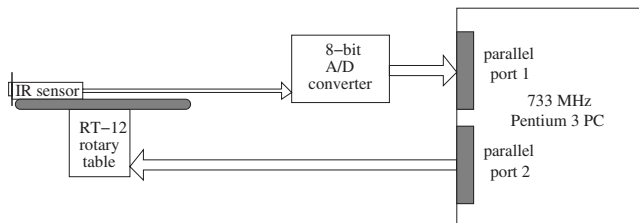


Fig. 3 Block diagram of the experimental setup.

retrained with the remaining weights, but this time with zero weight decay factor. In the implementation of LMA and OBS, the ANN-based system-identification toolbox is used.^{41,46}

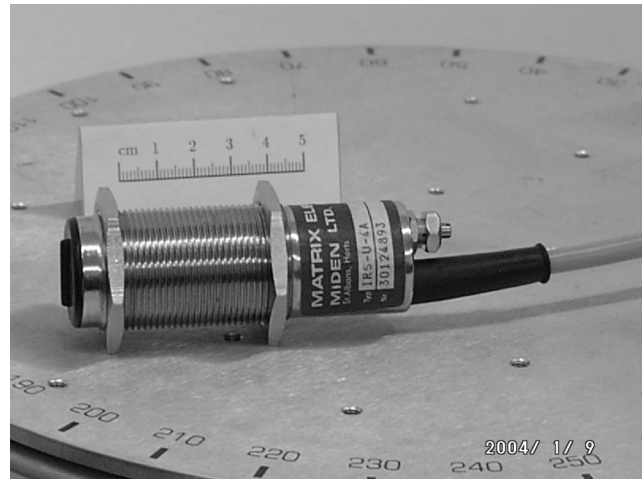
5 Experimental Verification

5.1 Experimental Setup

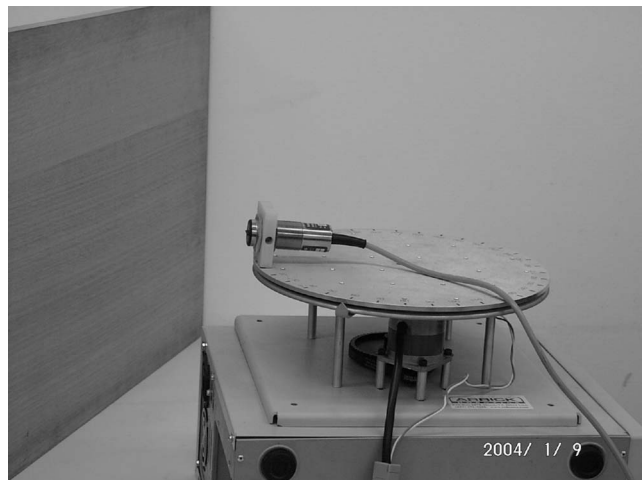
We verified our neural network implementation experimentally with patterns acquired by a simple infrared sensor from different target types. The infrared sensor⁴⁷ used in this study consists of an emitter and detector and works with 20–28 V dc input voltage, and it provides an analog output voltage proportional to the measured intensity reflected off the target. The analog signal is digitized using an 8-bit microprocessor-compatible A/D converter chip having a conversion time of 100 μ s and interfaced to the parallel port of a computer (Fig. 3). The detector window is covered with an infrared filter to minimize the effect of ambient light on the intensity measurements. Indeed, when the emitter is turned off, the detector reading is essentially zero. The sensitivity of the device can be adjusted with a potentiometer to set the operating range of the system. The infrared sensor [see Fig. 4(a)] is mounted on a 12-in. rotary table⁴⁸ to obtain angular intensity scans from these targets. A photograph of the experimental setup can be seen in Fig. 4(b). The target primitives employed in this study are a planar surface, a 90 deg corner, and a 90 deg edge, whose cross sections are given in Fig. 5, and each with a height of 120 cm. The horizontal extents of all targets are large enough that they can be considered infinite, and thus edge effects need not be considered. They are covered with materials of different surface properties. In this study, we used aluminum, white cloth, and Styrofoam packaging material as different surface types.

5.2 Differentiation of Geometry Types with ANN

Reference scan patterns are collected for each geometry-surface combination with 2.5-cm distance increments, from their nearest to their maximum observable ranges, at $\theta = 0$ deg. These are shown in Fig. 6, where the different curves on the same graph correspond to intensity scan patterns acquired at different distances for a particular target type. These scan patterns are the original scans, not their downsampled versions used as training inputs to the ANN. Scans of corners covered with white cloth and Styrofoam packaging material [Fig. 6(e) and 6(f)] have a triple-humped pattern (with a much smaller middle hump) corresponding to the two orthogonal constituent planes and their intersection. The intensity scan patterns for corners covered with aluminum [Fig. 6(d)] have three distinct saturated



(a)



(b)

Fig. 4 (a) The infrared sensor and (b) the experimental setup used in this study.

humps. Notice that the return signal intensities saturate at an intensity corresponding to about 11 V output voltage.

The training set consists of 147 sample scan patterns, 60 of which correspond to planes, 49 of which correspond to corners, and 38 of which correspond to edges. The number of scans for each geometry is different. This is because the targets have different reflective properties and each target is detectable over a different distance interval determined by its geometry and surface properties (see Fig. 7). We have

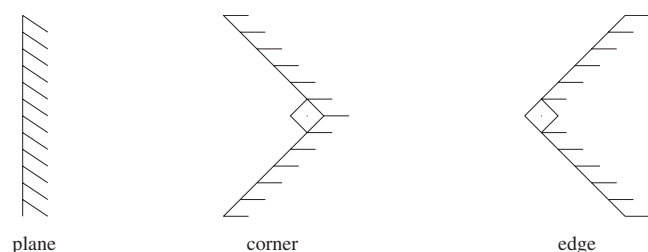


Fig. 5 Target geometries differentiated in this study.

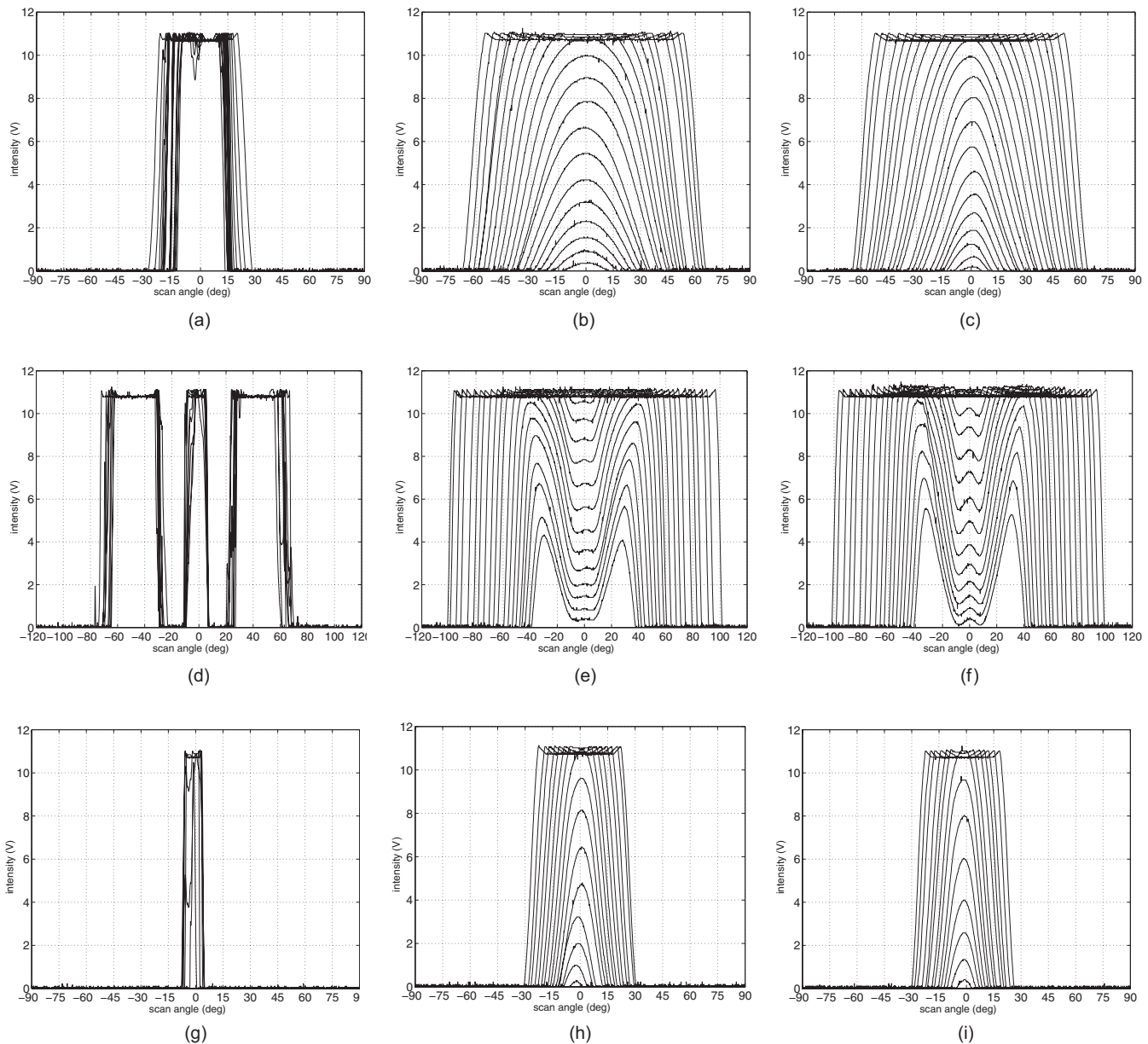


Fig. 6 Intensity scans for targets (first row, plane; second row, corner; third row, edge) covered with different surfaces (first column, aluminum; second column, white cloth; third column, Styrofoam) at different distances.

chosen to acquire the training scan patterns in a uniformly distributed fashion over the detectable range for each target. The input weights are initialized randomly. The ANN resulting in the highest correct differentiation rate on the training and test sets has 10 hidden-layer neurons in fully connected form. Initially, OBS is not used for pruning the network, so that this ANN does not have the optimal structure.

We test the ANN with infrared intensity scans acquired by situating targets at randomly selected distances r and azimuth angles θ (see Fig. 8) and collecting a total of 194 test scans, 82 of which are from planes, 64 from corners, and 48 from edges. The targets are randomly located at azimuth angles varying from -45 to 45 deg from their nearest to their maximum observable ranges in Fig. 7.

(Note that the test scans are collected for random target positions and orientations, whereas the training set was collected for targets at equally spaced ranges at $\theta=0$ deg)

When a test scan is obtained, first, we estimate the angular position of the target as follows. Assuming the observed scan pattern is not saturated, we check whether it has multiple humps or not. If so, it is a corner, and we find the angular location of the central hump and the corresponding intensity value. If not, we find the angular location of the maximum, denoted θ_{MAX} , and again the corresponding intensity value. If the observed scan pattern is saturated so that there are multiple maxima, we find its center of gravity [(COG) as described below] instead of the maximum value. A corner scan is considered saturated when its central intensity enters the saturation region, not

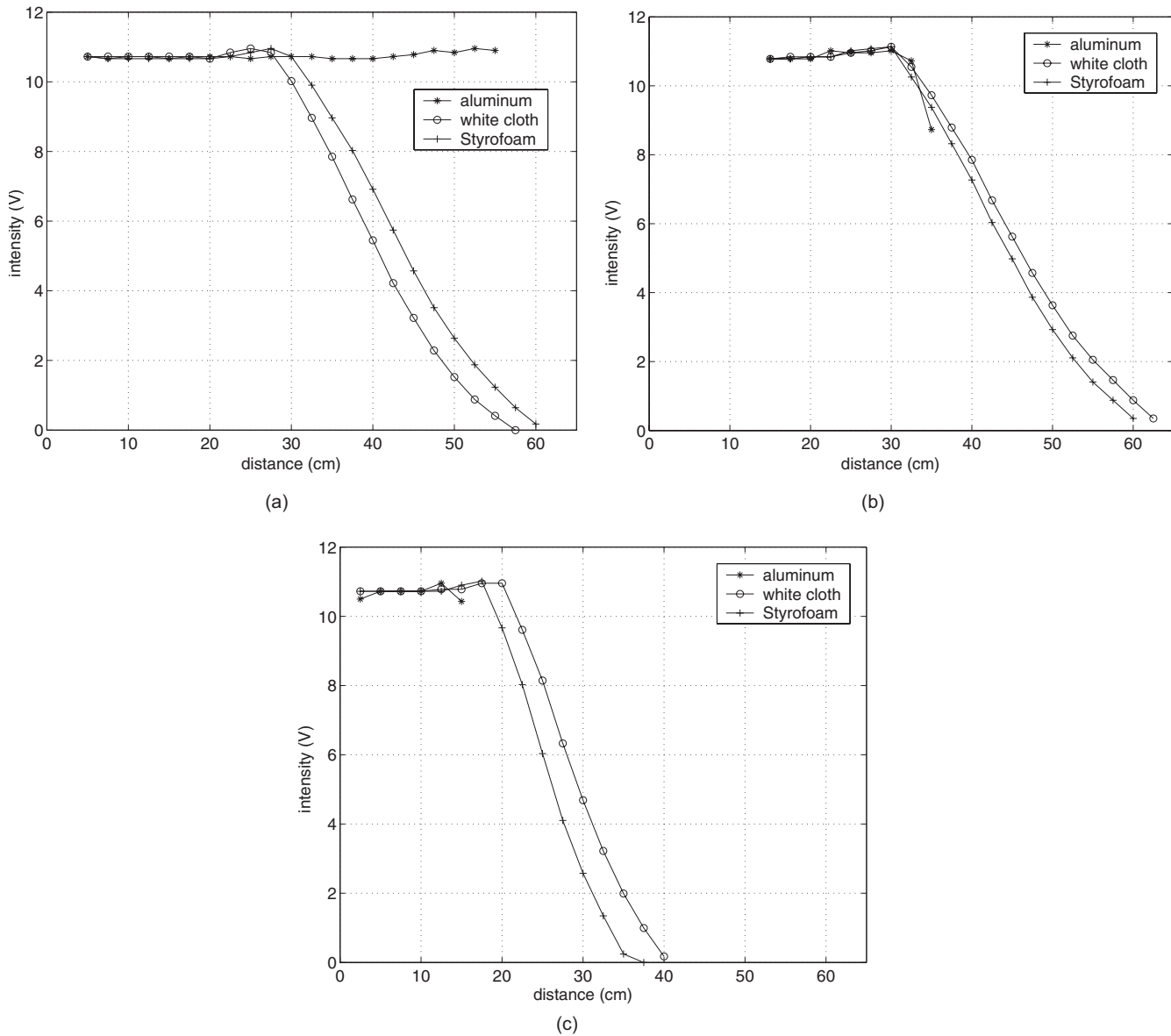


Fig. 7 COG intensity value versus distance curves for different targets: (a) plane, (b) corner, and (c) edge.

the side humps, because it is the former value that is relevant for our method. These angular values can be directly taken as estimates of the angular position of the target.

Alternatively, the angular position can be estimated by finding the COG of the scan as follows:

$$\theta_{\text{COG}} = \frac{\sum_{i=1}^n \alpha_i I(\alpha_i)}{\sum_{i=1}^n I(\alpha_i)}, \quad (11)$$

where n is the number of samples in the angular scan. Ideally, the COG and maximum intensity estimates would be equal due to the symmetry of the scanning process, but in practice they differ by a small amount. We consider the use of both alternatives when tabulating our results. Plots of the COG intensity (the intensity value at θ_{COG}) of each scan in Fig. 6 as a function of the distance at which that scan was obtained is provided in Fig. 7. As seen in the figure, the

detectable ranges of different target types fall within the interval [2.5, 62.5] cm. The absolute azimuth estimation of errors over all test targets is provided in Table 1. It can be concluded that the accuracy of the two azimuth estimation methods is comparable.

The test scans are shifted by the azimuth estimate, then downsampled by 10, and the resulting scan patterns are given as input to the ANN, which is trained using the LMA algorithm. The algorithm is used in the batch mode, where the network parameters are updated after processing the whole input data. The confusion matrix for the COG case is tabulated in Table 2. Corners are always correctly identified and not confused with the other target types due to the distinctive nature of their scans. Planes are confused with edges at six instances out of 82, and similarly, edges are confused with planes in five cases out of 48. An overall

Table 1 Absolute azimuth estimation errors over all test targets (P: plane, C: corner, E: edge; AL: aluminum, WC: white cloth, ST: Styrofoam).

Method	P			C			E			Average error
	AL	WC	ST	AL	WC	ST	AL	WC	ST	
Maximum (deg.)	0.9	2.3	0.8	2.4	1.7	1.3	1.1	2.0	1.7	1.6
COG (deg.)	0.9	1.0	0.8	2.4	1.4	1.1	1.2	2.2	2.3	1.5

correct differentiation rate of 94.3% is achieved. Second, to observe the effect of the azimuth estimation method, we used the maximum values of the unsaturated intensity scans. The results are presented in Table 3. In this case, both corners and edges are always correctly differentiated. However, seven of the test scans from planar surfaces are incorrectly differentiated as edges. Six of these are covered with aluminum, whose intensity scans are saturated. An overall correct differentiation rate of 96.4% is achieved, which is better than that obtained using COG due to the improvement in the classification of edges.

At the next step, the ANN is pruned with the OBS technique. Initially, there are $1633 = (159 + 1) \times 10 + (10 + 1) \times 3$ weights in the network in fully connected form. 1600 of these weights are between the input layer and the hidden layer, and 33 of them are between the hidden layer and the output layer. The plot of the sum-squared error $\mathcal{E}(w)$ [Eq. (1)] with respect to the number of remaining weights during pruning is shown in Fig. 9 for the training and test phases. In this figure, the errors evolve from right to left. The minimum error is obtained on the test set when 263 weights remain. The eliminated weights are set to zero. As the number of weights is decreased beyond 263, both the training and test errors increase rapidly due to the elimination of too many weights. If 263 weights are kept, the cor-

responding number of hidden-layer neurons is still 10. Thus, although 84% of the weights have been pruned, none of the hidden-layer neurons have been eliminated. The connectivity and the structure of the pruned network is illustrated in Fig. 10.

Using the weights resulting in the smallest test error, we retrained the network again with the LMA but this time with zero weight-decay factor. The ANN converges in seven iterations to an error of 0.00033. The output values of the pruned network are shown in Fig. 11. The first 64 test scans are from corners, the next 48 from edges, and the last 82 from planes. The target geometry corresponding to the maximum output value is selected. The differentiation results for the optimized network are given in Table 4. An overall correct differentiation rate of 99.0% is achieved. Therefore, apart from optimizing the structure of the ANN by eliminating the unnecessary weights, pruning the network results in improved geometry differentiation.

5.3 Differentiation of Surface Types with ANN

In the second stage, we consider differentiating the surface types of the targets assuming their geometries have been identified. The same network structure and the same procedure used in geometry differentiation is employed in surface type classification.

For the training set, all surface types are correctly differentiated for each type of geometry. For the test set, the confusion matrix for the three geometries and surfaces is given in Table 5. When the geometry is planar, an average correct differentiation rate of 80.5% is achieved. Aluminum planes are always correctly classified. The surface types of the corners are correctly classified with a rate of 85.9%. All corners covered with aluminum are correctly differentiated due to their distinctive features. The worst classification rate (64.6%) is achieved for edges due to the narrower base

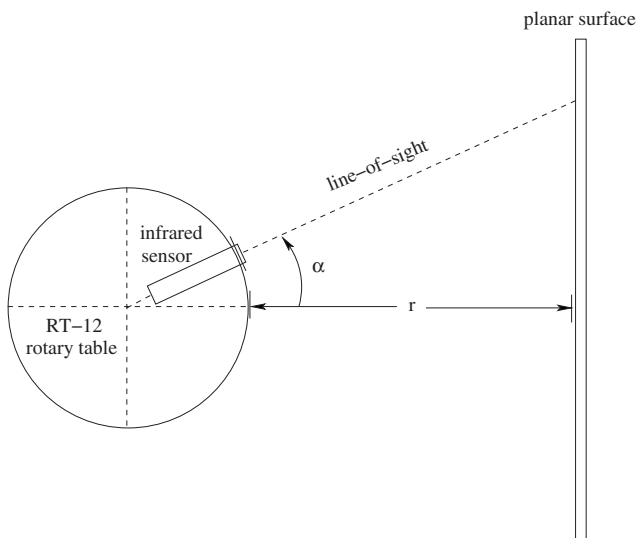


Fig. 8 Top view of the experimental setup and the related geometry. The emitter and detector windows are circular with 8 mm diameter and center-to-center separation of 12 mm. (The emitter is above the detector.) Both the scan angle α and the surface azimuth θ are measured counter-clockwise from the horizontal axis.

Table 2 Confusion matrix for ANN before pruning: COG-based azimuth estimation.

Target	Differentiation result			Total
	P	C	E	
P	76	—	6	82
C	—	64	—	64
E	5	—	43	48
Total	81	64	49	194

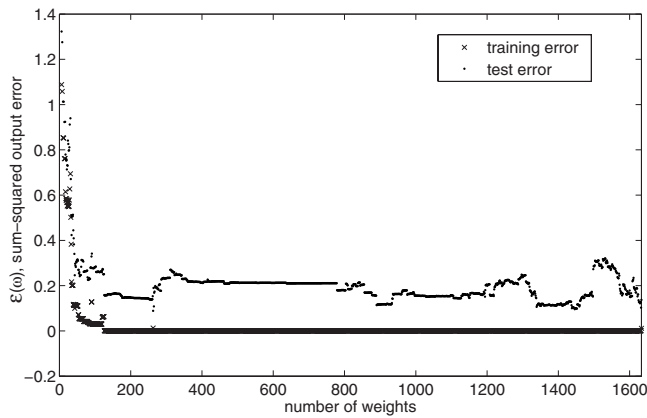


Fig. 9 Training and test errors while pruning the ANN with OBS.

width of the scans from edges. Edges covered with white cloth are not confused with Styrofoam packaging material. However, edges covered with Styrofoam are incorrectly classified as edges covered with white cloth with a rate of 72.2%. In instances where two or three outputs of the neural network are equal, the surface type is classified as unidentified. An overall correct differentiation rate of 78.4% is achieved for all surface types. It can be concluded that the geometry type most often confused with others is the edge. The corner geometry and the aluminum surface type are the most distinctive. The surface differentiation rate is not as high as expected because of the similarity of the intensity scans acquired from white-cloth- and Styrofoam-covered surfaces. We have also experimented with surface materials such as wood, white-painted matte wall, black cloth, and white, brown, and violet paper. However, the correct differentiation rates obtained with these surface types were about the same.

Table 3 Confusion matrix for ANN before pruning: maximum intensity-based azimuth estimation.

Target	Differentiation result			Total
	P	C	E	
P	75	—	7	82
C	—	64	—	64
E	—	—	48	48
Total	75	64	55	194

Once the geometry and the surface type of the target are identified, its range can be estimated by using linear interpolation on the appropriate curve in Fig. 7 using the intensity value at the azimuth estimate. This way, the localization process can be completed.

6 Conclusions

In this study, the intensity scan patterns acquired by a simple infrared sensor are processed using an artificial neural network for robust target differentiation. Both geometry and surface-type differentiation are considered. The input signals to the network are the infrared intensity scan patterns obtained from different target types by scanning them with the sensor. Two different methods are considered for estimating the azimuth of the targets. Although the two methods are comparable in terms of estimation accuracy, the maximum intensity method, in general, gives better results than the COG technique when the scans are processed by the neural network. The neural network is trained with the LMA and pruned with the OBS technique to achieve the

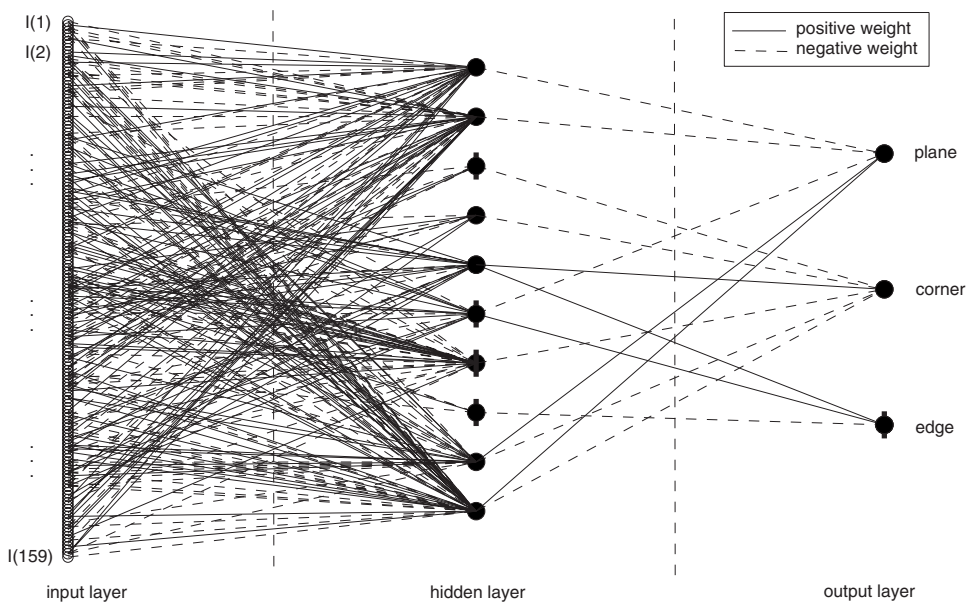


Fig. 10 The structure of the ANN after pruning with OBS. Positive weights are represented by a solid line, while a dashed line represents a negative weight. A bias is represented by a vertical line through the neuron.

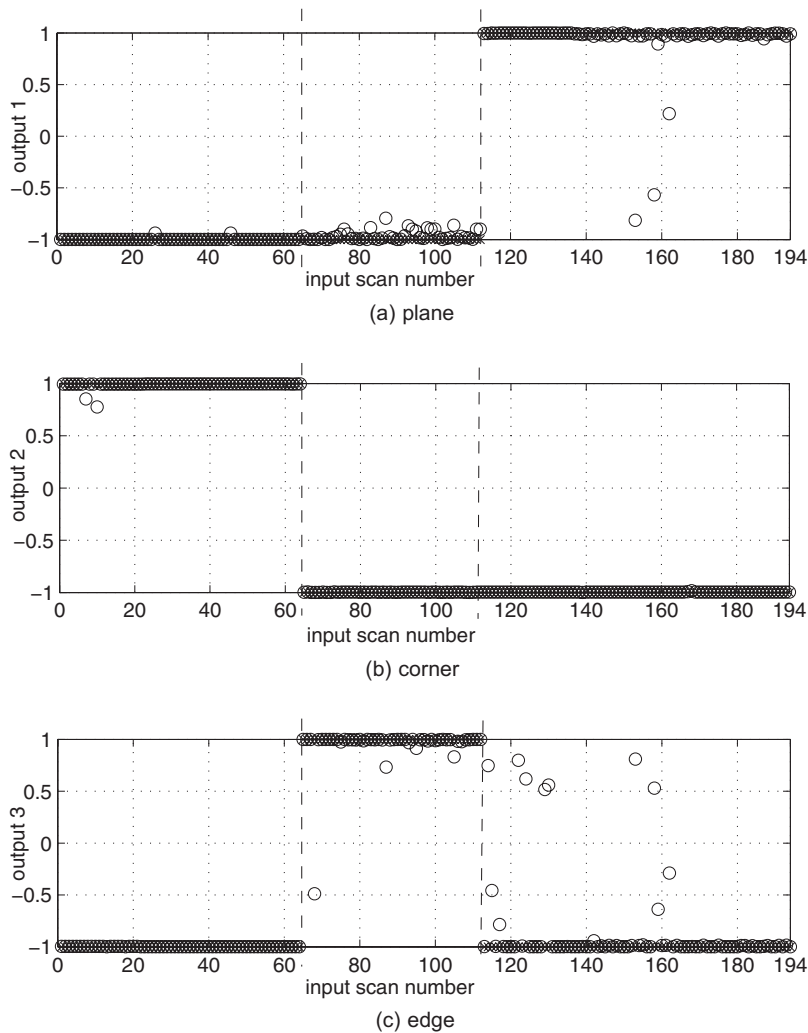


Fig. 11 ANN outputs for the optimal network: (a) plane, (b) corner, and (c) edge.

optimal network structure. Besides optimizing the structure, pruning results in improved classification. A modular approach is adopted where first the geometry of the targets is determined, followed by the surface type. Geometry and the surface type of the targets is classified with overall correct differentiation rates of 99 and 78.4%, respectively. Correct surface differentiation rate is not as high as geometry differentiation due to the similarity of the intensity

scans of some surface types having the same geometry. The results indicate that the geometrical properties of the targets are more distinctive than their surface properties, and surface determination is the limiting factor in differentiation. In this paper, we have demonstrated target differentiation for three target geometries and three different surface types. Based on the data we have collected, the differentiation results, and our previous works,^{35,36} it seems possible to increase the vocabulary of different geometries, provided they are not too similar. However, the same cannot be said for the number of different surfaces. For a given total number of distinct targets, increasing the number of surfaces and decreasing the number of geometries will in general make the results worse. However, decreasing the number of surfaces and increasing the number of geometries will in general improve the results.

Table 4 Confusion matrix for ANN after pruning: maximum intensity-based azimuth estimation.

Target	Differentiation result			Total
	P	C	E	
P	80	—	2	82
C	—	64	—	64
E	—	—	48	48
Total	80	64	50	194

The results reported here represent the outcome of our efforts to explore the limits of what is achievable in terms of identifying information with only a simple emitter/detector pair. Such simple sensors are usually put to much lower information-extracting uses. This work demonstrates that infrared intensity scans contain sufficient information to differentiate target geometries and surface types reason-

Table 5 Confusion matrix for three geometries and three surface types (P: plane, C: corner, E: edge, U: unidentified; AL: aluminum, WC: white cloth, ST: Styrofoam).

			Detected									
			P			C			E			U
			AL	WC	ST	AL	WC	ST	AL	WC	ST	
Actual	P	AL	24	—	—	—	—	—	—	—	—	—
		WC	—	23	6	—	—	—	—	—	—	—
		ST	—	9	19	—	—	—	—	—	—	1
	C	AL	—	—	—	22	—	—	—	—	—	—
		WC	—	—	—	—	14	8	—	—	—	—
		ST	—	—	—	—	1	19	—	—	—	—
	E	AL	—	—	—	—	—	—	8	—	—	2
		WC	—	—	—	—	—	—	—	19	—	1
		ST	—	—	—	—	—	—	—	13	4	1

ably well, and neural networks are capable of resolving this identifying information. When coupled with appropriate processing and recognition techniques, simple infrared sensors can be used to extract substantially more information about the environment than such devices are commonly employed for. This allows the possible applications to go beyond relatively simple tasks such as simple object and proximity detection, counting, distance and depth monitoring, floor sensing, position measurement, and obstacle/collision avoidance. The demonstrated system would find application in areas where recognition of patterns hidden in infrared scans or signals is required. Some examples are system control based on optical signal detection, identification, and clustering. Intelligent autonomous systems, such as mobile robots, that need to perform map-building, navigation, docking, and obstacle-avoidance tasks in unknown environments is another application area. Industrial applications where different materials/surfaces must be identified and separated may also benefit from this approach.

Given the attractive performance-for-cost of infrared-based systems, we believe that the results of this study will be useful for engineers designing or implementing infrared systems and researchers investigating algorithms and performance evaluation of such systems. While we have concentrated on infrared sensing in this paper, artificial neural networks can also be used with other sensing modalities such as radar and sonar, where the targets are characterized by complex signatures.

A possible direction for future work is to identify more generally shaped targets (such as a vase or a bottle) by acquiring several scan patterns from each target obtained at different heights and providing them as input to a neural network or other type of classifier. Techniques for fusing the input patterns from multiple sensing modalities, such as ultrasonic, infrared, and laser systems, can be developed for improved robustness in target differentiation. For example,

multiple neural networks can be trained, each specialized on one type of sensing modality. The decisions made by the individual networks can be fused or combined afterward using sensor-fusion techniques.

Acknowledgment

This work was supported in part, by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under Grant No. EEEAG-105E065.

References

1. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., Wiley Interscience, Hoboken, NJ (2000).
2. C. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford (1995).
3. M. T. Hagan, H. B. Demuth, and M. H. Beale, *Neural Network Design*, PWS, Boston, MA (1996).
4. J. D. Tardos, J. Neira, P. M. Newman, and J. J. Leonard, "Robust mapping and localization in indoor environments using sonar data," *Int. J. Robot. Res.* **21**(4), 311–330 (2002).
5. A. M. Flynn, "Combining sonar and infrared sensors for mobile robot navigation," *Int. J. Robot. Res.* **7**(6), 5–14 (1988).
6. H. M. Barberá, A. G. Skarmeta, M. Z. Izquierdo, and J. B. Blaya, "Neural networks for sonar and infrared sensors fusion," in *Proc. 3rd Int. Conf. on Information Fusion*, Vol. 2, pp. 18–25, International Society for Information Fusion, Fairborn, OH (2000).
7. V. Genovese, E. Guglielmelli, A. Mantuano, G. Ratti, A. M. Sabatini, and P. Dario, "Low-cost, redundant proximity sensor system for spatial sensing and color-perception," *Electron. Lett.* **31**(8), 632–633 (1995).
8. A. M. Sabatini, V. Genovese, E. Guglielmelli, A. Mantuano, G. Ratti, and P. Dario, "A low-cost composite sensor array combining ultrasonic and infrared proximity sensors," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Vol. 3, pp. 120–126 (1995).
9. I. D. Kelly and D. A. Keating, "Flocking by the fusion of sonar and active infrared sensors on physical autonomous mobile robots," in *Proc. 3rd Int. Conf. on Mechatronics and Machine Vision in Practice*, pp. 1–4, IEE, London (1996).
10. F. T. S. Yu and S. Jutamulia, Eds., *Optical Pattern Recognition*, Cambridge University Press, Cambridge (1998).
11. M. W. Roth, "Survey of neural network technology for automatic target recognition," *IEEE Trans. Neural Netw.* **1**(1), 28–43 (1990).

12. M. V. Shirvaikar and M. Trivedi, "A neural network filter to detect small targets in high clutter backgrounds," *IEEE Trans. Neural Netw.* **6**(1), 252–257 (1995).
13. P. Cayouette, G. Labonte, and A. Morin, "Probabilistic neural networks for infrared imaging target discrimination," in *Automatic Target Recognition XI*, F. A. Sadjadi, Ed., *Proc. SPIE* **5094**, 254–265 (2003).
14. A. L. Chan, S. Z. Der, and N. M. Nasrabadi, "Multistage infrared target detection," *Opt. Eng.* **42**(9), 2746–2754 (2003).
15. A. L. Chan, S. Z. Der, and N. M. Nasrabadi, "Improved target detector for FLIR imagery," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. II, pp. 401–404 (2003).
16. P. Zhang and J. Li, "Neural-network-based single-frame detection of dim spot target in infrared images," *Opt. Eng.* **46**(7), 1–11 (2007).
17. C. Narathong and R. M. Inigo, "Neural network target tracker," *Proc. SPIE* **1294**, 110–117 (1990).
18. A. B. Correia Bento and R. Nunes, "Grouping multiple neural networks for automatic target recognition in infrared imagery," in *Automatic Target Recognition XI*, F. A. Sadjadi, Ed., *Proc. SPIE* **4379**, 124–135 (2001).
19. J. Bauer and J. Mazurkiewicz, "Neural network and optical correlators for infrared imaging based face recognition," in *Proc. 5th Int. Conf. on Intelligent Systems Design and Applications*, pp. 234–238 (2005).
20. H. R. Everett, *Sensors for Mobile Robots, Theory and Application*, A K Peters, Wellesley, MA (1995).
21. P. M. Novotny and N. J. Ferrier, "Using infrared sensors and the Phong illumination model to measure distances," in *Proc. IEEE Int. Conf. on Robotics and Automation*, Vol. 2, pp. 1644–1649 (1999).
22. L. Korba, S. Elgazzar, and T. Welch, "Active infrared sensors for mobile robots," *IEEE Trans. Instrum. Meas.* **43**(2), 283–287 (1994).
23. K. Hashimoto, T. Tsuruta, K. Morinaka, and N. Yoshiike, "High performance human information sensor," *Sens. Actuators, A* **79**(1), 46–52 (2000).
24. P. J. de Groot, G. J. Postma, W. J. Melssen, and L. M. C. Buydens, "Validation of remote, on-line, near-infrared measurements for the classification of demolition waste," *Anal. Chim. Acta* **453**(1), 117–124 (2002).
25. G. Benet, F. Blanes, J. E. Simó, and P. Pérez, "Using infrared sensors for distance measurement in mobile robots," *Rob. Auton. Syst.* **40**(4), 255–266 (2002).
26. J. J. Esteve-Taboada, P. Refregier, J. Garcia, and C. Ferreira, "Target localization in the three-dimensional space by wavelength mixing," *Opt. Commun.* **202**(1–3), 69–79 (2002).
27. B. Iske, B. Jäger, and U. Rückert, "A ray-tracing approach for simulating recognition abilities of active infrared sensor arrays," *IEEE Sens. J.* **4**(2), 237–247 (2004).
28. Ç. Yüzbaşıoğlu and B. Barshan, "Improved range estimation using simple infrared sensors without prior knowledge of surface characteristics," *Meas. Sci. Technol.* **16**(7), 1395–1409 (2005).
29. H. V. Christensen, "Position detection based on intensities of reflected infrared light," Ph.D. Thesis, Aalborg University, Denmark (2005).
30. B. Ayruolu and B. Barshan, "Neural networks for improved target differentiation and localization with sonar," *Neural Networks* **14**(3), 355–373 (2001).
31. B. Barshan and B. Ayruolu, "Fractional Fourier transform preprocessing for neural networks and its application to object recognition," *Neural Networks* **15**(1), 131–140 (2002).
32. T. Aytaç and B. Barshan, "Rule-based target differentiation and position estimation based on infrared intensity measurements," *Opt. Eng.* **42**(6), 1766–1771 (2003).
33. T. Aytaç and B. Barshan, "Differentiation and localization of targets using infrared sensors," *Opt. Commun.* **210**(1–2), 25–35 (2002).
34. B. Barshan and T. Aytaç, "Position-invariant surface recognition and localization using infrared sensors," *Opt. Eng.* **42**(12), 3589–3594 (2003).
35. T. Aytaç and B. Barshan, "Simultaneous extraction of geometry and surface properties of targets using simple infrared sensors," *Opt. Eng.* **43**(10), 2437–2447 (2004).
36. T. Aytaç and B. Barshan, "Surface differentiation by parametric modeling of infrared intensity scans," *Opt. Eng.* **44**(6), 067202 (2005).
37. B. Barshan, T. Aytaç, and Ç. Yüzbaşıoğlu, "Target differentiation with simple infrared sensors using statistical pattern recognition techniques," *Pattern Recogn.* **40**(10), 2607–2620 (2007).
38. K. Levenberg, "A method for the solution of certain problems in least squares," *Q. Appl. Math.* **2**, 164–168 (1944).
39. D. Marquardt, "An algorithm for least squares estimation on nonlinear parameters," *SIAM J. Appl. Math.* **11**(2), 431–441 (1963).
40. M. T. Hagan and M. Menhaj, "Training feed-forward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.* **5**(6), 989–993 (1994).
41. M. Nørgaard, "Neural network-based system identification toolbox, Version 1.1," Tech. Rep. No. 97-E-851, Department of Automation, Department of Mathematical Modeling, Technical University of Denmark, Denmark (1997).
42. M. Nørgaard, O. Ravn, N. K. Poulsen, and L. K. Hansen, *Neural Networks for Modelling and Control of Dynamic Systems*, Springer-Verlag, London (2000).
43. J. Sjöberg and L. Ljung, "Overtraining, regularization, and searching for minimum with application to neural nets," *Int. J. Control* **62**(6), 1391–1407 (1995).
44. B. Hassibi, D. G. Stork, and G. J. Wolf, "Optimal brain surgeon and general network pruning," in *Proc. IEEE Int. Conf. on Neural Networks*, pp. 293–300 (1993).
45. B. Hassibi and D. G. Stork, "Second-order derivatives for network pruning: optimal brain surgeon," in *Advances in Neural Information Processing Systems*, S. J. Hanson, J. D. Cowan, and C. L. Giles, Eds., Vol. 5, pp. 164–172, Morgan Kaufmann, San Mateo, CA (1993).
46. "The NNSYSID toolbox, Version 2," (<http://www.iau.dtu.dk/research/control/nnsysid.html>) (2003).
47. *IRS-U-4A Proximity Switch Datasheet*, Matrix Elektronik, AG, Oberehrendingen, Switzerland (1995).
48. "RT-12 rotary positioning table," Arrick Robotics, (<http://www.robotics.com/rt12/html>, (2002)).



Tayfun Aytaç received his BS in electrical engineering from Gazi University, Ankara, Turkey in 2000 and MS and PhD in electrical engineering from Bilkent University, Ankara, Turkey in 2002 and 2006, respectively. During his graduate studies, he was a research and teaching assistant at Bilkent University, Department of Electrical and Electronics Engineering. He joined TÜBİTAK UEKAE/ İLTAREN Research and Development Group in 2006, where he is currently a senior research scientist. His current research interests include infrared sensing, intelligent sensing, infrared imaging systems, automatic target recognition, target tracking and classification, sensor data fusion, and sensor-based robotics.



Billur Barshan received her BS in both electrical engineering and physics from Boğaziçi University, Istanbul, Turkey, and MS and PhD in electrical engineering from Yale University, New Haven, Connecticut, in 1986, 1988, and 1991, respectively. In 1993, she joined Bilkent University, Ankara, where she is currently a professor in the Department of Electrical and Electronics Engineering. Barshan is the founder of the Robotics and Sensing Laboratory in the same department. Her current research interests include sensor-based robotics, ultrasonic, optical, and inertial sensing, multisensor data fusion, and human motion analysis/classification.