

## Neural Network-Based Target Differentiation Using Sonar for Robotics Applications

Billur Barshan, Birsal Ayrulu, and Simukai W. Utete

**Abstract**—This study investigates the processing of sonar signals using neural networks for robust differentiation of commonly encountered features in indoor robot environments. The neural network can differentiate more targets with higher accuracy, improving on previously reported methods. It achieves this by exploiting the identifying features in the differential amplitude and time-of-flight (TOF) characteristics of these targets. Robustness tests indicate that the amplitude information is more crucial than TOF for reliable operation. The study suggests wider use of neural networks and amplitude information in sonar-based mobile robotics.

**Index Terms**—Artificial neural networks, evidential reasoning, learning, majority voting, sensor data fusion, sonar sensing, target classification, target differentiation, target localization, ultrasonic transducers.

### I. INTRODUCTION

Neural networks have been employed efficiently as pattern classifiers in numerous applications [11]. These classifiers make weaker assumptions on the shape of the underlying distributions of input data than traditional statistical classifiers and can prove more robust when the underlying statistics are unknown or the data is generated by a nonlinear system.

The motivation behind the use of neural network classifiers in sonar or radar systems is the desire to emulate the remarkable perception and pattern recognition capabilities of humans and animals [1], [21], [23]. A comparison between neural networks and standard classifiers for radar-specific emitter identification is provided by Willson [27]. An acoustic imaging system that combines holography with multilayer feedforward neural networks for three-dimensional object recognition is proposed in [25]. A neural network which can recognize three-dimensional cubes and tetrahedrons using sonar is described in [9]. Neural networks have also been used in the classification of sonar returns from undersea targets [10], [21] and for the localization of cylinders in air [19].

Sonar is a very useful and cost-effective mode of sensing for mobile robots. This paper investigates the use of neural networks to process sonar signals encountered in target differentiation and localization applications for indoor environments. The pattern recognition capability of neural networks allows differentiation of more targets with increased accuracy by exploiting the identifying features in the differential amplitude and time-of-flight (TOF) characteristics of the reflected signals. The robustness of the network performance to partial removal of the input information has been investigated, demonstrating that the network is robust to different failure modes, and indicating that the amplitude information is more crucial than TOF for reliable target differentiation and localization. A comparison with previously investigated approaches indicates improved performance.

Manuscript received June 23, 1999; revised February 14, 2000. This paper was recommended for publication by Associate Editor A. Bicchi and Editor V. Lumelsky upon evaluation of the reviewers' comments. This paper was presented in part at the IEEE International Conference on Robotics and Automation, San Francisco, CA, April 2000.

B. Barshan and B. Ayrulu are with the Department of Electrical Engineering, Bilkent University, TR-06533 Bilkent, Ankara, Turkey.

S. W. Utete is with the Robotics Research Group, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, U.K.

Publisher Item Identifier S 1042-296X(00)06952-4.

The most common sonar ranging system is based on TOF which is the time elapsed between transmission and reception of a pulse. Differential TOF models of targets have been used by several researchers [7], [16], [17]. Systems using only qualitative information [12], combining amplitude, energy, and duration of the echo signals along with TOF information [2], [3], [8], or exploiting the complete echo waveform [13] have also been considered. In addition, several researchers have investigated rejecting or handling crosstalk and higher order reflections [5], [6].

Accurate target classification can be achieved by using multitransducer pulse/echo systems and employing both amplitude and TOF information. However, a major problem with using the amplitude information of sonar signals is that the amplitude is very sensitive to environmental conditions. For this reason, and also because the standard electronics typically provide only TOF data, amplitude information is seldom used. In the present paper, neural networks are used to process amplitude and TOF information so as to reliably handle the target classification problem.

### II. BACKGROUND ON SONAR SENSING

In the commonly used TOF systems, an echo is produced when the transmitted pulse encounters an object and a range value  $r = ct_o/2$  is produced when the echo amplitude first exceeds a preset threshold level  $\tau$  back at the receiver at time  $t_o$ . Here,  $t_o$  is the TOF and  $c$  is the speed of sound in air (at room temperature,  $c = 343.3$  m/s).

In general, it is observed that the echo amplitude decreases with increasing range and azimuth  $\theta$  (Fig. 1). The echo amplitude falls below  $\tau$  when  $|\theta| > \theta_o$ , which is related to the aperture radius  $a$  and the resonance frequency  $f_o$  of the transducer by  $\theta_o = \sin^{-1}(0.61c/af_o)$  [28].

With a single stationary transducer, it is not possible to estimate the azimuth of a target with better resolution than  $2\theta_o$ . In our system, two identical acoustic transducers  $a$  and  $b$  with center-to-center separation  $d$  are employed to improve the angular resolution. Each transducer can operate both as transmitter and receiver and detect echo signals reflected from targets within its own *sensitivity region* [Fig. 1(a)]. Both members of the sensor configuration can detect targets located within the *joint sensitivity region* [Fig. 1(b)].

The target primitives modeled in this study are *plane*, *corner*, *acute corner*, *edge*, and *cylinder* (Fig. 2). Since the wavelength ( $\lambda \cong 8.6$  mm at  $f_o = 40$  kHz) is much larger than the typical roughness of surfaces encountered in laboratory environments, targets in these environments reflect acoustic beams specularly, like a mirror. Specular reflections allow the single transmitting-receiving transducer to be viewed as a separate transmitter  $T$  and virtual receiver  $R$  [14]. Detailed physical reflection models of these target primitives with corresponding echo signal models are provided in [2]. Experimentally obtained amplitude and TOF characteristics of these target primitives are illustrated in Figs. 3 and 4. In these figures, the solid lines correspond to the average over eight data sets. The level of amplitude and TOF noise is also illustrated by plotting the  $\pm 3\sigma_A$  and  $\pm 3\sigma_t$  curves together with the average amplitude and TOF curves, respectively. Here,  $\sigma_A(\sigma_t)$  is the amplitude (TOF) noise standard deviation. For the target differentiation methods compared in the following sections, experimentally obtained data in similar form have been used.

### III. TARGET DIFFERENTIATION ALGORITHM

In this section, the target differentiation algorithm used in earlier work [2] is reviewed. This will not only be useful in motivating the structure of the inputs to be used in the neural network, but will also

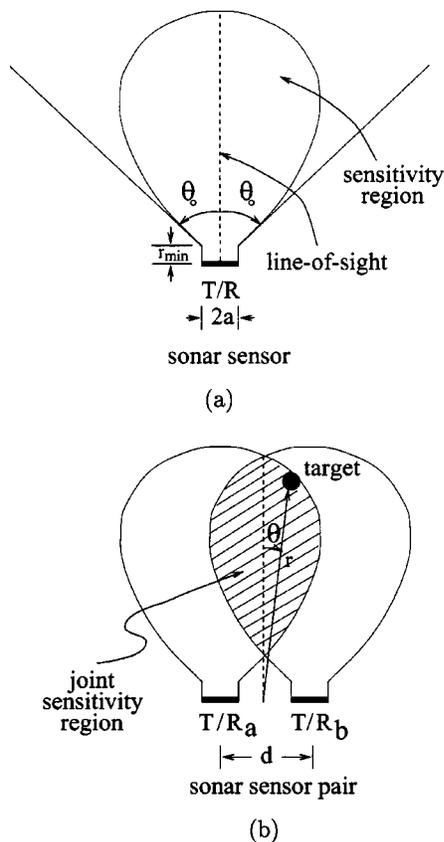


Fig. 1. (a) Sensitivity region of an ultrasonic transducer. (b) Joint sensitivity region of a pair of ultrasonic transducers. The intersection of the individual sensitivity regions serves as a reasonable approximation to the joint sensitivity region.

provide a basis for comparison of its performance. This classification algorithm has its roots in the plane/corner differentiation algorithm developed in the still earlier work by Barshan and Kuc [3], which is based on the idea of exploiting amplitude differentials in resolving target type (Fig. 3). In [2], the algorithm is extended to include other target primitives using amplitude and TOF differentials of Figs. 3 and 4. The extended algorithm may be summarized in the form of rules:

if  $[t_{aa}(\alpha) - t_{ab}(\alpha)] > k_t \sigma_t$  **and**  $[t_{bb}(\alpha) - t_{ba}(\alpha)] > k_t \sigma_t$  then **acute corner**  $\rightarrow$  exit  
 if  $[A_{aa}(\alpha) - A_{ab}(\alpha)] > k_A \sigma_A$  **and**  $[A_{bb}(\alpha) - A_{ba}(\alpha)] > k_A \sigma_A$  then **plane**  $\rightarrow$  exit  
 if  $[\max\{A_{aa}(\alpha)\} - \max\{A_{bb}(\alpha)\}] < k_A \sigma_A$  **and**  $[\max\{A_{bb}(\alpha)\} - \max\{A_{ab}(\alpha)\}] < k_A \sigma_A$  then **corner**  $\rightarrow$  exit  
 else **edge, cylinder or unknown**  $\rightarrow$  exit

Here,  $A_{aa}$ ,  $A_{ab}$ ,  $A_{ba}$ , and  $A_{bb}$  denote the maximum values of the sonar echo signals, and  $t_{aa}$ ,  $t_{ab}$ ,  $t_{ba}$ , and  $t_{bb}$  denote the TOF readings extracted from these signals. The first index in the subscript indicates the transmitter, the second denotes the receiver. The  $k_A$  ( $k_t$ ) is the number of  $\sigma_A$  ( $\sigma_t$ ) which is employed as a safety margin to achieve robustness in the differentiation process. Differentiation is achievable only in those cases where the differentials exceed  $k_A \sigma_A$  ( $k_t \sigma_t$ ). If this is not the case, a decision cannot be made and the target type remains unknown.

Two variations of this algorithm can be distinguished. The first takes into account the noise model and statistics to achieve robustness ( $k_A, k_t \neq 0$ ), whereas the second treats the data as noiseless

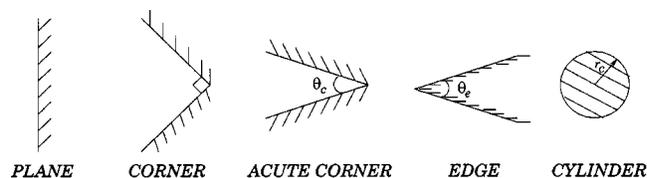


Fig. 2. Horizontal cross sections of the target primitives modeled and differentiated in this study.

( $k_A, k_t = 0$ ). Since the first version is more conservative, a lower rate of incorrect decisions is expected at the expense of a higher rate of undecidable target type. In the second case, there is no safety margin and consequently a larger rate of incorrect decisions is expected.

#### IV. TARGET CLASSIFICATION AND LOCALIZATION WITH NEURAL NETWORKS

The algorithm summarized in the previous section does not provide a distinctive rule to differentiate several types of edges and cylinders. One way of differentiating edges and cylinders is to use the radius of curvature estimation method in [4]. In this work, neural networks are employed to identify and resolve parameter relations embedded in the characteristics of experimentally obtained sonar returns from all target primitives in a robust manner.

Panasonic transducers are used with aperture radius  $a = 0.65$  cm, resonance frequency  $f_0 = 40$  kHz, and  $\theta_0 \cong 54^\circ$  [20] (Fig. 1). The center-to-center separation of the transducers used in the experiments is  $d = 25$  cm. The entire sensing unit is mounted on a small stepper motor with step size  $1.8^\circ$  whose motion is controlled through the parallel port of a PC 486. Data acquisition from the sonars is through a 12-bit 1-MHz A/D card. Echo signals are processed on a PC 486. Starting at the transmit time, 10 000 samples of each echo signal are collected and thresholded. The amplitude information is extracted by finding the maximum value of the signal after the threshold is exceeded.

The targets employed in this study are cylinders with radii 2.5, 5.0 and 7.5 cm, a planar target, a corner, an edge of  $\theta_e = 90^\circ$ , and an acute corner of  $\theta_c = 60^\circ$ . Amplitude and TOF data from these targets are collected with the sensing unit described above at 25 different locations ( $r, \theta$ ) for each target, from  $\theta = -20^\circ$  to  $\theta = 20^\circ$  in  $10^\circ$  increments, and from  $r = 35$  cm to  $r = 55$  cm in 5-cm increments (Fig. 5). The target primitive located at range  $r$  and azimuth  $\theta$  is scanned by the sensing unit for scan angle  $-52^\circ \leq \alpha \leq 52^\circ$  with  $1.8^\circ$  increments. With the given scan range and motor step size, 58 ( $= (2 \times 52/1.8)$ ) sets of amplitude and TOF data ( $A_{aa}, A_{ab}, A_{ba}, A_{bb}; t_{aa}, t_{ab}, t_{ba}, t_{bb}$ ) are acquired for each target location. The structure of these amplitude and TOF characteristics is provided in [2].

The network employed has one hidden, one input, and one output layer. Although there are many ways of choosing input signals to train the network, this study uses differentials in amplitude and TOF signals as input signals. This choice is motivated by the form of the terms in the differentiation algorithm outlined in Section III. Therefore, the inputs to the neural network are 58 samples each of the difference signals  $A_{aa}(\alpha) - A_{ab}(\alpha)$ ,  $A_{bb}(\alpha) - A_{ba}(\alpha)$ ,  $t_{aa}(\alpha) - t_{ab}(\alpha)$ , and  $t_{bb}(\alpha) - t_{ba}(\alpha)$ . This means that the input layer has 232 neurons.

The hidden layer comprises 100 neurons. This number was determined by a process known as *enlarging*, which starts with a relatively small number of neurons and increases the size of the hidden layer until learning occurs. The number of output layer neurons is 21. The first seven neurons encode the target type. The next seven represent the target range  $r$  which is binary coded with a resolution of 0.25 cm. The last seven neurons represent the azimuth  $\theta$  of the target with respect to the line-of-sight of the sensing unit, which is also binary coded with resolution  $0.5^\circ$ .

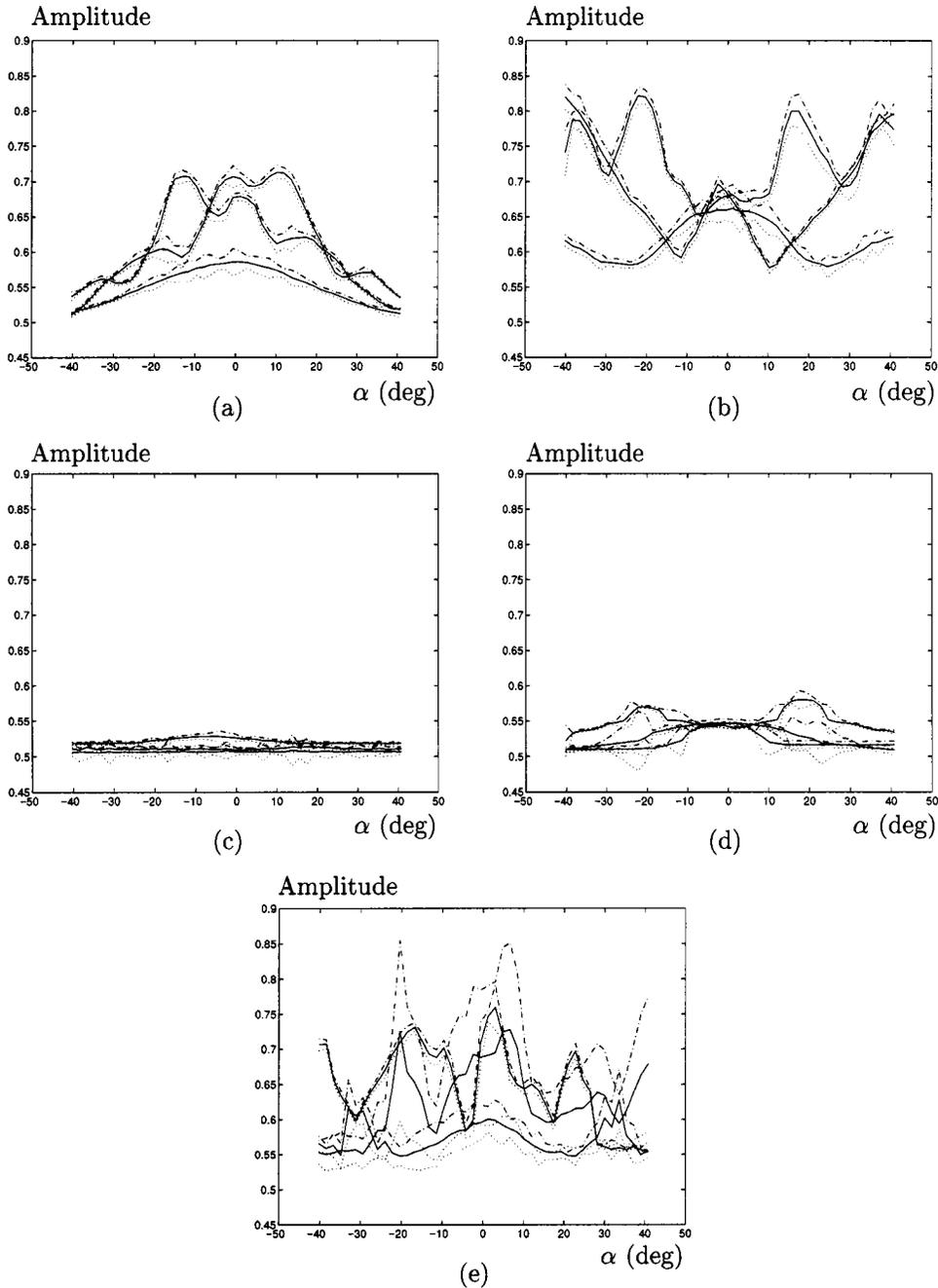


Fig. 3. Amplitude characteristics which incorporate the amplitude noise ( $\pm 3\sigma_A$ ) for the targets: (a) plane (b) corner (c) edge with  $\theta_e = 90^\circ$  (d) cylinder with  $r_c = 5$  cm (e) acute corner with  $\theta_c = 60^\circ$ . Here solid, dashed, and dotted lines correspond to the average over eight data sets, average  $+3\sigma_A$  and average  $-3\sigma_A$ , respectively.

Four sets of data are collected for each target location for each target primitive, resulting in 700 (=4 data sets  $\times$  25 locations  $\times$  7 target types) sets of waveform data for training. The network is trained with these 700 sets of data, using the back-propagation algorithm [26] with a learning constant equal to 0.01 and momentum constant equal to 0.9, and a sigmoid-type nonlinearity. With the software tool PlaNet [18], the weights are found in about 1 h.

The network is tested as follows. Each target primitive is placed in turn in each of the 25 locations shown in Fig. 5. Four sets of measurements are collected for each combination of target type and location, again resulting in 700 sets of experimentally obtained waveform data. The neural network estimates the target type, range, and azimuth from this data.

Table I gives the resulting percentages of correct target-type classification, correct range, and correct azimuth estimation. A range or azimuth estimate is considered correct if it is within an error tolerance of  $\epsilon_r$  or  $\epsilon_\theta$  of the actual range or azimuth, respectively. The average percentages over all target types are also given in the last row of the table. The percentage of correct target type classification is high at 95%. The percentage of correct range estimation lies in the range 74%–93%, and that for correct azimuth estimation lies in the range 89%–97%, depending on the error tolerance level ( $\epsilon_r$  or  $\epsilon_\theta$ ). For comparison, the average correct target type classification obtained using the differentiation algorithm given in Section III on the same data set is 61% and the average correct range and azimuth estimation percentages are 72% and 59%, respectively, for  $|\epsilon_r| = 1$  cm and  $|\epsilon_\theta| = 2^\circ$ .

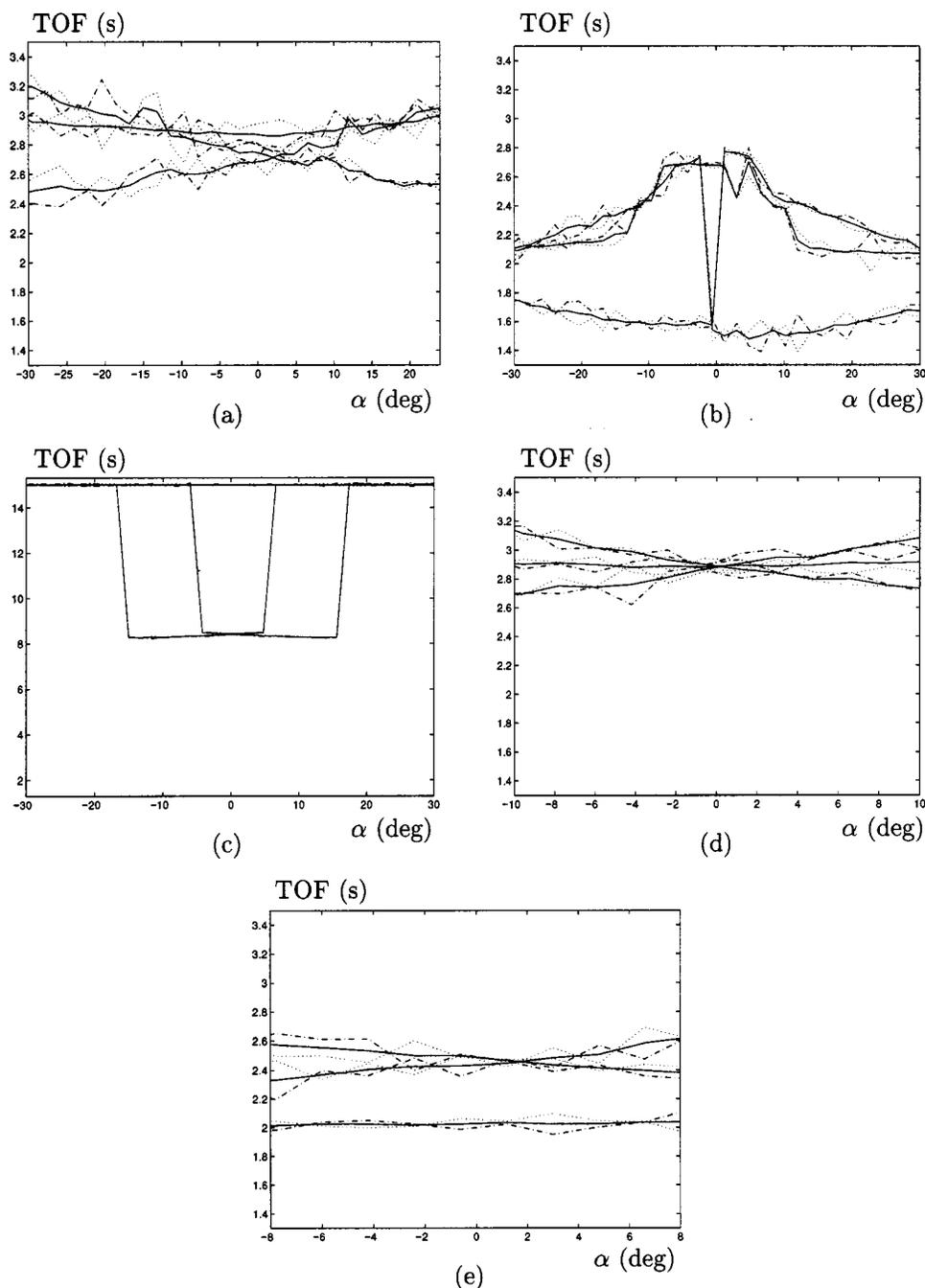


Fig. 4. TOF characteristics which incorporate the TOF noise ( $\pm 3\sigma_t$ ) for the targets: (a) plane (b) corner (c) edge with  $\theta_e = 90^\circ$  (d) cylinder with  $r_c = 5$  cm (e) acute corner with  $\theta_c = 60^\circ$ . Here solid, dashed, and dotted lines correspond to the average over eight data sets, average  $+3\sigma_t$ , and average  $-3\sigma_t$ , respectively.

The network is also tested for targets situated arbitrarily in the continuous estimation space and not necessarily confined to the 25 locations of Fig. 5. The results are given in Table I in parentheses. As expected, the percentages in this case are lower than those for the training positions; the network gives the best results when a target is situated exactly at one of the training sites. Noting that the network was trained only at 25 locations and at grid spacings of 5 cm and  $10^\circ$ , it can be concluded from the percentage of correct range and azimuth estimates obtained at error tolerances of  $|\epsilon_r| = 0.125$  and 1 cm and  $|\epsilon_\theta| = 0.25^\circ$  and  $2^\circ$ , that the network demonstrates the ability to interpolate between the training grid locations. Thus, the neural network maintains a certain spatial continuity between its input and output and does not haphazardly map positions which are not drawn from the 25 loca-

tions of Fig. 5. The correct target type percentages in Table I are quite high and the accuracy of the range/azimuth estimates would be acceptable in many applications. If better estimates are required, this can be achieved by reducing the training grid spacing in Fig. 5.

In addition to the above-mentioned structure, multistage and modular network structures have also been implemented and tested. In the multistage network, training is performed in two or three stages by introducing the target type, range, and azimuth information in sequence. In the modular structure, three separate networks for target type, range, and azimuth have been employed. However, these structures did not result in substantial improvement in the differentiation and localization process.

The network is further tested to investigate its robustness in physically plausible failure or missing data situations. The same 700 sets of test

TABLE I  
PERCENTAGES OF CORRECT CLASSIFICATION, RANGE ( $r$ ), AND AZIMUTH ( $\theta$ ) ESTIMATION. NUMBERS GIVEN IN PARENTHESES ARE THE RESULTS WHEN 100 ARBITRARY LOCATIONS IN CONTINUOUS ESTIMATION SPACE ARE USED FOR TESTING, WHEREAS THE NUMBERS BEFORE THE PARENTHESES ARE FOR WHEN TRAINING POSITIONS ARE USED FOR TESTING

target type	% of correct classif.	% of correct $r$ estimation				% of correct $\theta$ estimation			
		error tolerance $\epsilon_r$				error tolerance $\epsilon_\theta$			
		$\pm 0.125$ cm	$\pm 1$ cm	$\pm 5$ cm	$\pm 10$ cm	$\pm 0.25^\circ$	$\pm 2^\circ$	$\pm 10^\circ$	$\pm 20^\circ$
plane	100(90)	62(47)	66(50)	78(69)	88(80)	91(66)	96(68)	98(93)	99(92)
corner	99(100)	90(72)	90(72)	92(86)	95(91)	88(81)	90(85)	92(90)	93(90)
edge ( $\theta_e = 90^\circ$ )	99(96)	51(49)	59(56)	82(78)	92(90)	72(60)	77(69)	89(81)	97(95)
acute corner ( $\theta_c = 60^\circ$ )	98(99)	81(71)	83(75)	91(83)	95(88)	85(79)	87(80)	93(90)	97(92)
cylinder ( $r_c = 2.5$ cm)	90(88)	77(55)	80(60)	89(78)	94(87)	97(79)	97(83)	97(95)	98(99)
cylinder ( $r_c = 5.0$ cm)	89(70)	75(61)	77(65)	82(78)	89(91)	98(73)	98(78)	99(98)	99(96)
cylinder ( $r_c = 7.5$ cm)	92(86)	82(55)	86(60)	92(76)	98(86)	95(54)	97(60)	99(94)	99(96)
average	95(90)	74(59)	77(63)	87(78)	93(88)	89(70)	92(75)	95(92)	97(94)

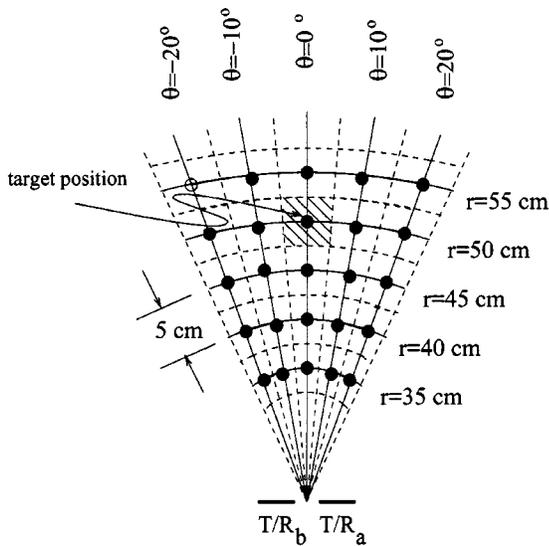


Fig. 5. Discrete network training locations.

data are used, but with some of the network inputs equated to zero. The results are tabulated in Table II for both the case where testing is done at the training locations, and the case when testing is done at arbitrary locations. Rows 1–4 correspond to the case when one of the differential input channels is made completely unavailable to the network. Rows 4–8 represent failure of one of the transducers. Rows 7 and 8 also correspond to the case when the target does not fall within the joint sensitivity region of the two transducers. Rows 9–12 correspond to the case when the amplitudes of the echo waveforms fall completely below the preset threshold level. This happens when the target is very far away from the sensor or too far off its line-of-sight. In this case, TOF information cannot be extracted although amplitude information is still available. Rows 13–16 correspond to the complementary case where TOF information is available but amplitude information is not. Finally, the effect of the absence of randomly selected samples of the input data is investigated and presented in row 17. Here, 25% of the input data is made unavailable to the network by randomly setting

some of the input samples to zero. Note that this percentage is the same as the percentage of samples excluded when one of the input channels is completely blocked. The result of these tests indicate that amplitude information is much less dispensable, despite the fact that TOF is the more commonly used parameter. It can also be concluded from the table that there is a 1%–16% decrease in performance when the object is at an arbitrary location as compared to when it is at a training location.

The sensing unit has been mounted on a small mobile robot navigating in a test room consisting of eight planes, six corners, and two edges. Data are collected at five different positions by scanning the environment (Fig. 6). At each scan angle  $\alpha$ , a decision about the type of the target which is closest to the line-of-sight of the sensing unit is made by employing both the differentiation algorithm and the neural network classifier. The correct decision percentages averaged over the entire scan are shown in Fig. 6, indicating that the neural network outperforms the differentiation algorithm at all five positions. We emphasize that the neural network employed here is the same one trained with the discrete locations of Fig. 5; it was not trained specifically for this room. Yet, it had a success rate ranging between 71%–90%, except for position 1 where the proximity of the several features to each other, and the fact that the two edges are closer to the robot than the minimum training distance overwhelm the network. The correct decision percentages when decisions of five nodes are fused by Dempster–Shafer rule of combination and simple majority voting, described in the next section, are 68.8% and 52.5%, respectively. These percentages are still less than the correct decision percentage obtained by the neural network classifier at nodes 2, 3, 4, and 5.

The next example is related to the path planning of a mobile robot in an environment cluttered with cylinders. In addition, there is a corner behind the cylinders which causes higher order and multiple reflections to occur. A mobile robot equipped with the previously described sensing unit is aimed to get from its current position to the goal following the shortest path possible (Fig. 7). We again compare the differentiation algorithm with the neural network classifier. The decisions made when the sensor is looking toward each of the three cylinders are also indicated on the same figure. The differentiation algorithm detects two of the cylinders as planes and one as a corner, whereas the neural network classifier can correctly identify all cylinders. (In this example,

TABLE II  
PERCENTAGES OF CORRECT CLASSIFICATION, RANGE, AND AZIMUTH ESTIMATION WHEN THE TARGETS ARE TESTED IN THE TRAINING POSITIONS AND AT ARBITRARY POSITIONS (IN PARENTHESES)

testing condition	type %	range %	azimuth %
$A_{aa} - A_{ab} = 0$	58(50)	20(13)	34(19)
$A_{bb} - A_{ba} = 0$	59(45)	22(16)	35(19)
$t_{aa} - t_{ab} = 0$	94(83)	71(58)	89(73)
$t_{bb} - t_{ba} = 0$	95(84)	71(58)	89(73)
$A_{aa} = t_{aa} = 0$	25(22)	13(10)	16(11)
$A_{bb} = t_{bb} = 0$	22(19)	17(14)	15(10)
$A_{ab} = t_{ab} = 0$	16(15)	9(7)	16(9)
$A_{ba} = t_{ba} = 0$	25(24)	13(9)	13(7)
$t_{aa} = 0$	94(82)	69(56)	89(73)
$t_{ab} = 0$	95(83)	74(61)	90(74)
$t_{ba} = 0$	95(82)	74(60)	90(74)
$t_{bb} = 0$	94(81)	67(55)	87(72)
$A_{aa} = 0$	25(21)	13(10)	22(14)
$A_{ab} = 0$	16(15)	9(7)	16(10)
$A_{ba} = 0$	24(23)	9(7)	13(8)
$A_{bb} = 0$	23(19)	17(13)	15(9)
25% of inputs zeroed	81(73)	38(32)	65(53)

use of the differentiation algorithm would have resulted in doubling of the path length.)

## V. COMPARATIVE ANALYSIS AND DISCUSSION

In this part, the network is compared to the algorithm discussed in Section III (for the three target types which can be identified by that algorithm) and two previously proposed methods [2], [24]: Dempster-Shafer evidential reasoning [22] and majority voting [15], which are employed to resolve conflicts in the decisions of multiple sensors. Dempster-Shafer evidential reasoning is useful for representing and manipulating beliefs of decision makers about an event, especially in situations where their beliefs reflect uncertainty or ignorance. Voting is a simple tool for resolving conflicts when the emphasis is on the view of the majority. Both of these fusion methods and neural networks share the feature of being nonparametric (i.e., no underlying distribution of input data or noise is assumed).

For target differentiation based purely on raw data, the algorithm of Section III gives a correct differentiation percentage of 61%. In [24], based on this algorithm, sensors assign probability masses to plane, corner, and acute corner target types using Dempster-Shafer evidential reasoning. Combining the opinions of 15 sensing nodes using Dempster's rule of combination improves the correct decision percentage to 87%. When the sensors' beliefs about target types are counted as votes and the majority vote is taken as the outcome, the number rises to 88%.

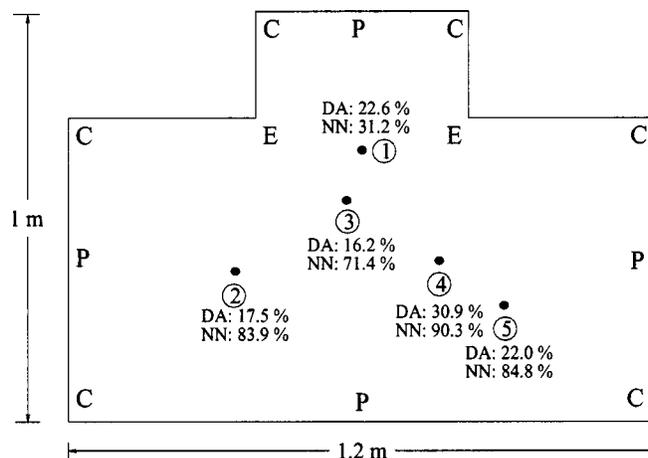


Fig. 6. The experimental test room. For each node, correct decision percentages are given for the differentiation algorithm (DA) and the neural network (NN).

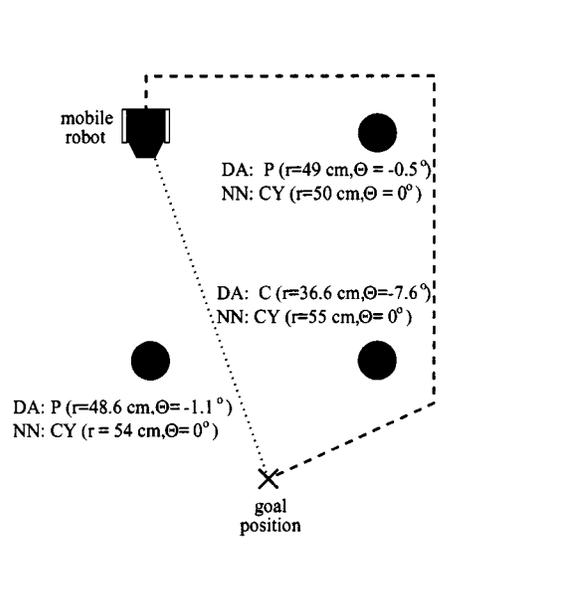


Fig. 7. Mobile robot path planning. The dotted and dashed lines correspond to the minimum-length paths designed by employing the neural network (NN) and the differentiation algorithm (DA), respectively.

Moreover, using various ordering strategies in the voting algorithm further increases this number to 90%. Using these two fusion methods, only planes, corners, and acute corners can be differentiated. Using the neural network described in this paper, seven different target types can be differentiated and localized employing only a single sensor node, with a higher correct decision percentage (95%) than with the previously used decision rules employing several sensors. The fact that the neural network is able to distinguish all target types indicates that it must be making more effective use of the available data than the methods used earlier. The neural network's performance shows that the original training data set does contain the information sufficient to differentiate the seven target types, but the other methods mentioned above are not able to resolve this identifying information. The neural network allows differentiation of more targets with increased accuracy by exploiting the hidden identifying features in the differential amplitude and TOF characteristics of the targets. A comparison of the different methods is presented in Table III. It should be noted that the percentage of correct estimates for the neural network can be further

TABLE III  
COMPARISON OF THE VARIOUS METHODS OF TARGET DIFFERENTIATION. TARGET TYPES ENCLOSED IN BRACES CAN BE RESOLVED ONLY AS A GROUP

method	target types discriminated	accuracy	sensor configuration	learning	a priori information	associativity of "decision rule"
differentiation algorithm without noise statistics	P,C,AC	61%	single scanning node	no	no	associative
differentiation algorithm with noise statistics [2, 3]	P,C,AC	≤61%	single scanning node	no	yes	associative
Dempster-Shafer evidential reasoning [24]	P,C,AC, {E,CY,U}	87%	15 scanning sensing nodes	no	no	associative
voting [24]	P,C,AC, {E,CY,U}	88%	15 scanning sensing nodes	no	no	associative
		90%	ordered 15 scanning sensing nodes	no	no	not necessarily
radius of curvature estimation [4]	P,CY,{C,E}	80%	single stationary node	no	no	not applicable
neural network	P,C,AC,E,CY	95%	single scanning node	yes	no	not applicable

improved by decreasing the grid spacings in Fig. 5 and thus using a more finely spaced training set.

## VI. CONCLUSION

In this study, neural networks are employed to process real sonar data after trained to learn identifying parameter relations for the target primitives. This system uses amplitude as well as TOF data, allowing for improved differentiation and localization. The robustness of the network to partial removal of amplitude and TOF information has been investigated, demonstrating that the network is robust to different failure modes. The results suggest that amplitude measurements should be more widely exploited, rather than limiting sonar systems to the more conveniently available TOF measurements.

A comparison with previously investigated approaches to decision fusion (evidential reasoning and voting) indicates improved performance. Whereas these methods rely on our selection of important features from the observed data and assignment of beliefs accordingly, neural networks are capable of automatic selection of the discriminating features. Neural networks are of interest compared to other learning methods because they offer the ability to discriminate a larger number of target types than these methods, often with greater accuracy. Furthermore, they accomplish this using only a single sensing node. Had the number of sensing nodes (15) been reduced in the other methods, their accuracy would have been even worse. In addition, the neural network is found to be robust to a variety of failure modes. Although trained on a discrete and relatively coarse grid, the network is able to interpolate between the grid locations and offers higher resolution (especially in azimuth) than that implied by the grid size. The correct estimation rates for target type, range and azimuth can be further increased by employing a finer grid for training.

The results presented here suggest wider use of neural networks as robust pattern classifiers in sensor-based robotics. There is scope for further application of neural networks to sonar, based on the facts that sonar data is difficult to interpret, physical models can be complex even for simple TOF sonar, and expressions for sonar returns are very complicated even for the simplest target types. Acoustic propagation is also subject to distortion with changes in environmental conditions. Despite the fact that most surfaces encountered in typical indoor environments can be considered specularly reflecting, it would also be of interest to consider surfaces of varying roughness and reflectivity properties as the extension of this work. Although the amplitude and TOF characteristics of rough surfaces will exhibit variations from those of specular surfaces, the strength of the neural network approach is precisely that it can handle such situations by appropriate training without the need for explicit physical modeling. Future work will investigate scale- and shift-invariant features in order to reduce the number of training patterns needed. Unsupervised learning algorithms will also be considered to make the classification process more robust to changes in environmental conditions.

## REFERENCES

- [1] W. W. L. Au, "Comparison of sonar discrimination—dolphin and artificial neural network," *J. Acoust. Soc. Amer.*, pt. 1, vol. 95, no. 5, pp. 2728–2735, 1994.
- [2] B. Ayrulu and B. Barshan, "Identification of target primitives with multiple decision-making sonars using evidential reasoning," *Int. J. Robot. Res.*, vol. 17, no. 6, pp. 598–623, 1998.
- [3] B. Barshan and R. Kuc, "Differentiating sonar reflections from corners and planes by employing an intelligent sensor," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 6, pp. 560–569, 1990.
- [4] B. Barshan and A. Ş. Sekmen, "Radius of curvature estimation and localization of targets using multiple sonar sensors," *J. Acoust. Soc. Amer.*, vol. 105, no. 4, pp. 2318–2331, 1999.

- [5] M. Beckerman and E. Oblov, "Treatment of systematic errors in the processing of wide-angle sonar sensor data for robotic navigation," *IEEE Trans. Robot. Automat.*, vol. 6, pp. 137–145, Apr. 1990.
- [6] J. Borenstein and Y. Koren, "Error eliminating rapid ultrasonic firing for mobile robot obstacle avoidance," *IEEE Trans. Robot. Automat.*, vol. 11, pp. 132–138, Feb. 1995.
- [7] Ö. Bozma and R. Kuc, "Building a sonar map in a specular environment using a single mobile sensor," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 12, pp. 1260–1269, 1991.
- [8] —, "A physical model-based analysis of heterogenous environments using sonar—ENDURA method," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 497–506, May 1994.
- [9] I. E. Dror, M. Zagaeski, and C. F. Moss, "3-dimensional target recognition via sonar—a neural network model," *Neural Networks*, vol. 8, no. 1, pp. 149–160, 1995.
- [10] R. P. Gorman and T. J. Sejnowski, "Learned classification of sonar targets using a massively parallel network," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 7, pp. 1135–1140, 1998.
- [11] R. P. Lippman, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, pp. 4–22, Apr. 1987.
- [12] R. Kuc, "Three-dimensional tracking using qualitative bionic sonar," *Robot. Autom. Syst.*, vol. 11, no. 2, pp. 213–219, 1993.
- [13] —, "Biomimetic sonar recognizes objects using binaural information," *J. Acoust. Soc. Amer.*, vol. 102, no. 2, pp. 689–696, 1997.
- [14] R. Kuc and M. W. Siegel, "Physically-based simulation model for acoustic sensor robot navigation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 6, pp. 766–778, 1987.
- [15] L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition: An analysis of its behavior and performance," *IEEE Trans. Syst., Man, Cybern.*, vol. 27, no. 5, pp. 553–568, 1997.
- [16] J. J. Leonard and H. F. Durrant-Whyte, *Directed Sonar Navigation*. London, U.K.: Kluwer, 1992.
- [17] J. Manyika and H. F. Durrant-Whyte, *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*. New York: Ellis Horwood, 1994.
- [18] Y. Miyata, "A User's Guide to PlaNet Version 5.6, A Tool for Constructing, Running, and Looking into a PDP Network," Comput. Sci. Dept., Univ. Colorado, Boulder, CO, Jan. 1991.
- [19] T. Ogawa, K. Kameyama, R. Kuc, and Y. Kosugi, "Source localization with network inversion using an answer-in-weights scheme," *IEICE Trans. Inform. Syst.*, vol. E79-D, no. 5, pp. 608–619, 1996.
- [20] *Ultrasonic Ceramic Microphones*, Panasonic Corp., Burlington, MA, 1989.
- [21] H. L. Roitblat, W. W. L. Au, P. E. Nachtigall, R. Shizumura, and G. Moons, "Sonar recognition of targets embedded in sediment," *Neural Netw.*, vol. 8, no. 7/8, pp. 1263–1273, 1995.
- [22] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.
- [23] J. A. Simmons, P. A. Saillant, J. M. Wotton, T. Haresign, M. J. Ferragamo, and C. F. Moss, "Composition of biosonar images for target recognition by echolocating bats," *Neural Netw.*, vol. 8, no. 7/8, pp. 1239–1261, 1995.
- [24] S. W. Utete, B. Barshan, and B. Ayralu, "Voting as validation in robot programming," *Int. J. Robot. Res.*, vol. 18, no. 4, pp. 401–413, 1999.
- [25] S. Watanabe and M. Yoneyama, "An ultrasonic visual sensor for three-dimensional object recognition using neural networks," *IEEE Trans. Robot. Automat.*, vol. 8, pp. 240–249, Apr. 1992.
- [26] P. J. Werbos, "Backpropagation through time: What it does and how to do it," *Proc. IEEE*, vol. 78, pp. 1550–1560, Oct. 1990.
- [27] G. B. Willson, "Radar classification using a neural network," *Proc. SPIE, Optical Engineering and Photonics in Aerospace Sensing: Application of Neural Networks*, vol. 1294, pp. 200–210, 1990.
- [28] J. Zemanek, "Beam behavior within the nearfield of a vibrating piston," *J. Acoust. Soc. Amer.*, pt. 2, vol. 49, no. 1, pp. 181–191, 1971.

## Choosing Good Distance Metrics and Local Planners for Probabilistic Roadmap Methods

Nancy M. Amato, O. Burchan Bayazit, Lucia K. Dale,  
Christopher Jones, and Daniel Vallejo

**Abstract**—This paper presents a comparative evaluation of different distance metrics and local planners within the context of probabilistic roadmap methods for planning the motion of rigid objects in three-dimensional workspaces. The study concentrates on cluttered three-dimensional workspaces typical of, for example, virtual prototyping applications such as maintainability studies in mechanical CAD designs. Our results include recommendations for selecting appropriate combinations of distance metrics and local planners for such applications. Our study of distance metrics shows that the importance of the translational distance increases relative to the rotational distance as the environment becomes more crowded. We find that each local planner makes some connections that none of the others do—indicating that better connected roadmaps will be constructed using multiple local planners. We propose a new local planning method we call *rotate-at-s* that often outperforms the common straight-line in C-space method in crowded environments.

**Index Terms**—Distance metrics, local planners, motion planning, probabilistic roadmaps.

### I. INTRODUCTION

Automatic motion planning has application in many areas such as robotics, virtual reality systems, and computer-aided design. Although many different motion planning methods have been proposed, most are not used in practice since they are computationally infeasible except for some restricted cases, e.g., when the robot has very few degrees of freedom [12], [16]. For this reason, attention has focused on randomized methods, such as randomized potential field methods, e.g., [5].

Recently, a class of randomized motion planning methods, called *probabilistic roadmap methods* (PRM's), has gained much attention (see, e.g., [1], [4], [11], [14], [22]). These methods use randomization during preprocessing to construct a graph in C-space (a *roadmap*). Queries are processed by connecting the initial and goal configurations to the roadmap, and then finding a path in the roadmap between these two connection points. PRM's have been shown to perform well in practice, answering difficult queries in fractions of seconds [4], [14].

Manuscript received May 19, 1998; revised June 28, 1999 and March 30, 2000. This paper was approved for publication by Associate Editor J. Laumond and Editor V. Lumelsky upon evaluation of the reviewers' comments. The work of O. B. Bayazit was supported in part by the Turkish Ministry of Education. The work of L. K. Dale was supported in part by a GE Foundation Graduate Fellowship and a Department of Education GAANN Fellowship. The work of D. Vallejo was supported in part by a Fulbright-CONACYT Scholarship. This work was supported in part by a National Science Foundation (NSF) CAREER Award CCR-9624315, by the NSF under Grant IRI-9619850, and by the Texas Higher Education Coordinating Board under Grant ARP-036327-017. This paper was presented in part at the IEEE International Conference on Robotics and Automation, 1998.

N. M. Amato, O. B. Bayazit, and L. K. Dale are with the Department of Computer Science, Texas A&M University, College Station, TX 77843-3112 USA (email: amato@cs.tamu.edu; burchanb@cs.tamu.edu; dalel@cs.tamu.edu).

C. Jones was with the Department of Computer Science, Texas A&M University, College Station, TX 77843-3112 USA. He is now with Sandia National Laboratories, Albuquerque, NM USA (email: cvjones@sandia.gov).

D. Vallejo is with the Department of Computer Science, Texas A&M University, College Station, TX 77843-3112 USA, on leave from Universidad de las Americas, Puebla, Mexico (email: dvallejo@cs.tamu.edu).

Publisher Item Identifier S 1042-296X(00)06955-X.