PROCESSING OF SONAR SIGNALS USING NEURAL NETWORKS FOR ROBUST TARGET DIFFERENTIATION

Birsel Ayrulu and Billur Barshan

Department of Electrical Engineering
Bilkent University 06533 Bilkent, Ankara, Turkey

ABSTRACT

This study investigates the processing of sonar signals using neural networks for robust differentiation of commonly encountered features in indoor environments. Differentiation of such features is of interest for intelligent systems. Amplitude and time-of-flight (TOF) patterns of sonar signals acquired from a real system are processed using neural networks. Compared to earlier approaches, 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 TOF characteristics of these targets. High accuracies are achieved both for target differentiation and localization. An important observation follows from the robustness tests, which indicate that the amplitude information is more crucial than TOF for reliable operation.

1. INTRODUCTION

Neural networks have been employed efficiently as pattern classifiers in numerous applications [1, 2, 3]. These classifiers make weaker assumptions on the shape of the underlying distributions of input signals than traditional statistical classifiers and can prove more robust when the underlying statistics are unknown or the data are generated by a nonlinear system.

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 TOF characteristics of the echo 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.

Commonly employed ranging systems are based on *time-of-flight* (TOF) which is the time elapsed between transmission and reception of a pulse. Differential TOF models of targets have been used by several researchers [4, 5]. In earlier work, Barshan and Kuc introduce a method based on both amplitude and TOF information to differentiate planes and corners [6]. This algorithm is later extended to other target primitives in [7]. In the present paper, neural networks are employed to process amplitude and TOF information so as to reliably handle the target classification and localization problem.

2. BACKGROUND ON SONAR SENSING

In 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 τ at time t_o back at the receiver. Here, t_o is the TOF and c is the speed of sound in air.

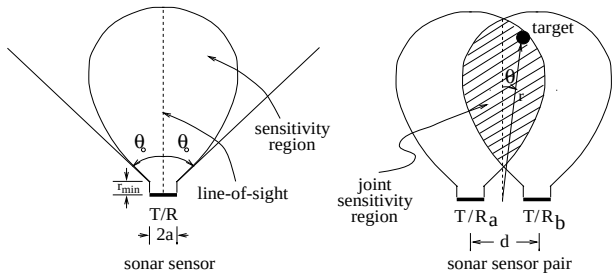

Figure 1: (a) Sensitivity region of an ultrasonic transducer. (b) The joint sensitivity region of a pair of ultrasonic transducers.

In general, it is observed that the echo amplitude decreases with increasing target range (r) and absolute value of the target azimuth ($|\theta|$). The echo amplitude falls below τ when $|\theta| > \theta_o$, which is related to the aperture radius a and the resonance frequency f_o of the transducer by $\theta_o = \sin^{-1}\left(\frac{0.61c}{af_o}\right)$ [8].

In our system, two identical ultrasonic transducers a and b with center-to-center separation d are employed to improve the angular resolution of a single transducer, limited by its beamwidth. Each transducer can operate both as transmitter and receiver and detect echo signals reflected from targets within its *sensitivity region* [Fig.1(a)]. Both transducers can detect targets located within the *joint sensitivity region*, which can be approximated by the overlap of the individual sensitivity regions [Fig.1(b)].

The target primitives modeled in this study are *plane*, *corner*, *acute corner*, *edge* and *cylinder* (Fig.2). Since the wavelength of operation ($\lambda \cong 8.6$ mm at $f_\circ = 40$ 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 [9]. Detailed physical reflection models of these target primitives with corresponding echo signal models are provided in [7]. In the following, 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 transmitting transducer, the second index denotes the receiver.

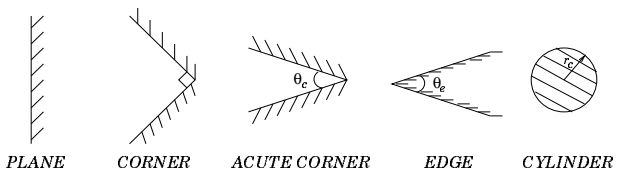

Figure 2: Horizontal cross sections of the target primitives differentiated in this study.

3. DIFFERENTIATION ALGORITHM

In this section, the target differentiation algorithm used in earlier work [7] is summarized. This will not only be useful in motivating the structure of the inputs to the neural network, but will also 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 [6] based on the idea of exploiting amplitude differentials in resolving target type. In [7], the algorithm is extended to include other target primitives using amplitude and TOF differentials. 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

Note that this algorithm does not provide a distinctive rule to differentiate edges and cylinders. In the above algorithm, $k_A(k_t)$ is the number of amplitude (TOF) noise standard deviations $\sigma_A(\sigma_t)$ and is employed as a safety margin to achieve robustness in the differentiation process. Differentiation is achievable only in those cases where the difference in amplitudes (TOFs) exceeds $k_A\sigma_A(k_t\sigma_t)$. If this is not the case, a decision cannot be made and the target type remains unknown.

4. DIFFERENTIATION WITH NEURAL NETWORKS

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 considered in a robust manner. Panasonic transducers are used with aperture radius $a = 0.65$ cm, resonance frequency $f_\circ = 40$ kHz, and $\theta_\circ \cong 54°$ [10] (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° whose motion is controlled through the parallel port of a PC with the aid of a microswitch. Data acquisition from the sonars is through a 12-bit 1 MHz PC A/D card. 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 $\theta_e = 90°$ and an acute corner of $\theta_c = 60°$. Amplitude and TOF data from these targets is collected with the sensing unit described above at 25 different locations (r, θ) for each target, from $\theta = -20°$ to $\theta = 20°$ in 10° increments, and from $r = 35$ cm to $r = 55$ cm in 5 cm increments (Fig.3). The target at range r and azimuth θ is scanned by the sensing unit for scan angle $-52° \leq \alpha \leq 52°$ with 1.8° increments.

A three-layer feed forward neural network is employed. The inputs to the neural network are samples of the amplitude and TOF differential signals $A_{aa}(\alpha) - A_{ab}(\alpha)$, $A_{bb}(\alpha) - A_{ba}(\alpha)$, $t_{aa}(\alpha) - t_{ab}(\alpha)$, and $t_{bb}(\alpha) - t_{ba}(\alpha)$. The target type, and its range and azimuth are produced at the output of the network.

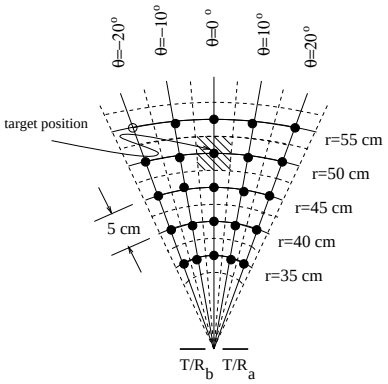

Figure 3: Discrete network training locations.

Four sets of data are collected for each target location for each target primitive, resulting in 700 (= 4 data sets × 25 locations × 7 target types) sets of training signals. The network is trained with these 700 sets of data, using the back-propagation algorithm.

The network is tested as follows: Each target primitive is placed in turn in each of the 25 locations shown in Fig.3. 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 r, and azimuth θ based on this data.

Table 1 presents the resulting percentages of correct target-type classification and correct r and θ estimation. A range or azimuth estimate is considered correct if it is within an error tolerance of ϵ_r or ϵ_θ of the actual r or θ 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 r estimation lies in the range 74–93%, and that for correct θ estimation lies in the range 89–97%, depending on the error tolerance level (ϵ_r or ϵ_θ). For comparison, the average correct target type classification obtained using the differentiation algorithm given in Sec.3 on the same data set is 61% and the average correct r and θ estimation percentages are 72% and 59% respectively for $|\epsilon_r| = 1$ cm and $|\epsilon_\theta| = 2°$.

The network is also tested for targets situated arbitrarily in the continuous estimation space and not necessarily confined to the 25 locations of Fig.3. The results are given in Table 5 in parentheses. As expected, the percentages in this case are lower than those for the training positions. Noting that the network was trained only at 25 locations and at grid spacings of 5 cm and 10°, it can be concluded from the percentage of correct r and θ estimates obtained at error tolerances of $|\epsilon_r| = 0.125$ cm and 1 cm and $|\epsilon_\theta| = 0.25°$ and 2°, 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 locations of Fig.3. The correct target type percentages in Table 1 are quite high and the localization (r, θ) accuracy would be acceptable in many applications. If better estimates are required, this can be achieved by reducing the training grid spacing in Fig.3.

The network is further tested to investigate its robustness in physically plausible failure or missing data situations. The same 700 sets of test data are used, but with some of the network inputs equated to zero. The results are tabulated in Table 2 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 echo amplitudes fall completely below τ. 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 the last row. 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 more commonly exploited. 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.

5. CONCLUSION

In this study, neural networks are employed for the processing of real sonar data for target differentiation and localization. The targets are discriminated using a practical pulse/echo sonar system. As a result of training, the network learns identifying parameter relations for the target primitives. This system relies on amplitude as well as TOF data, thus allowing for improved

target type	% of correct classif.	% of correct r estimation				% of correct θ estimation			
		error tolerance ϵ_r				error tolerance ϵ_θ			
		± 0.125 cm	± 1 cm	$\pm$ 5 cm	± 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)

Table 1: Percentages of correct classification, and r and θ estimation when the targets are tested in the training positions and at arbitrary positions (in parentheses). The resolutions of the r and θ encoders are 0.25 cm and 0.5°.

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 various failure modes. The results indicate that amplitude information should be more widely exploited.

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 θ) than that implied by the grid size. The correct estimation rates for target type, r and θ can be further increased by employing a finer grid for training.

In conclusion, the results presented here suggest wider use of neural networks as robust pattern classifiers for sonar signal processing. 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 systems, 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. Future work will investigate scale- and shift-invariant features of the targets and the employment of unsupervised learning algorithms.

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)

Table 2: The percentages of correct classification, r, and θ estimation when the targets are tested in the training positions and at arbitrary positions (in parentheses).

6. REFERENCES

[1] Dror I.E., Zagaeski M., and Moss C.F., *Neural Networks*, 8(1):149–160, 1995.

[2] Gorman R.P. and Sejnowski T.J., *IEEE T. Acoust. Speech, Signal Process.*, 36:1135–1140, 1998.

[3] Ogawa T., Kameyama K., Kuc R., and Kosugi Y., *IEICE T. Inf. Syst.*, E79-D(5):608–619, 1996.

[4] Bozma Ö. and Kuc R., *IEEE T. Pattern Anal. Mach. Intell.*, 13(12):1260–1269, 1991.

[5] Manyika J. and Durrant-Whyte H.F., *Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach*, Ellis Horwood, New York, 1994.

[6] Barshan B. and Kuc R., *IEEE T. Pattern Anal. Mach. Intell.*, 12(6):560–569, 1990.

[7] Ayrulu B. and Barshan B., *Int. J. Robot. Res.*, 17(6):598–623, 1998.

[8] Zemanek J., *J. Acoust. Soc. Am.*, 49(1 (Part 2)):181–191, 1971.

[9] Kuc R. and Siegel M.W., *IEEE T. Pattern Anal. Mach. Intell.*, PAMI-9(6):766–778, 1987.

[10] Panasonic Corp., Ultrasonic ceramic microphones, Burlington, MA, 1989.