

Fast Target Classification Using Sonar

Andrew Heale and Lindsay Kleeman

Intelligent Robotics Research Centre
Department of Electrical and Computer Systems Engineering
Monash University, Victoria 3800, Australia

Andrew.Heale@eng.monash.edu.au Lindsay.Kleeman@eng.monash.edu.au

Abstract

This paper describes a new sonar system that can perform target localisation in two dimensions and classification into planes, concave corners and convex edges with no extra time overhead. That is, the sensor transmits on two transmitters a short time apart, thereby collecting echoes in virtually the same time as a single transmitter system. Moreover, the time separation of the transmitted pulses acts to identify the particular sonar system so that interference from other systems can be rejected. The sensor combines two previous sonar research efforts on double pulse coding [2] and classification [1] in a real time DSP based sensing module that is also smaller than previous sensors. Since the classification is performed with such a short delay between transmitter firings, the sensor could be deployed on moving platforms to achieve on-the-fly mapping. This paper describes the sonar hardware, the Maximum Likelihood Estimation (MLE) classification approach and experimental results.

1 Introduction

How much information can sonar provide? Accurate range and bearing measurements of multiple targets have been achieved [1,2,4,5], interference can be rejected [2,3,4], and targets can be classified [1,5,6]. Properly implemented, sonar accuracy may exceed that of laser rangefinders, while still being cheap and reliable.

Knowing the shape, or class, of a target assists in robot localisation. It enables prediction of how the target will appear from different sensor positions, and it simplifies matching sensed environmental features to a map.

It is now common to use the three classes of target: *plane*, 90° concave *corner* and convex *edge*. The minimum requirements to classify targets into these categories are two transmitter positions and two receivers [1].

Existing methods require multiple readings of the target, to obtain coordinates of virtual images

of a transmitter in two different positions. This means moving a single transmitter, or incorporating two transmitters into the sensor. Either way, in the interval between taking the two readings there can be significant air movement, which contributes to errors in the measurement, and reduces the reliability of the eventual classification.

If the interval can be reduced, reliability should be improved. Additionally, the latency before the target can be classified will be reduced.

We present a method of reducing this latency, and two methods of classifying the target as a *corner*, *plane*, *edge*, or unknown.

The whole process can be compressed into a single processing cycle by firing two transmitters nearly simultaneously. Indeed, the precise separation used can then identify the sensor, and aid in eliminating crosstalk and interference – a technique known as *double pulse coding* [2]. This paper presents new work that integrates interference rejection with target classification.

Using DSP technology we have produced a sensor that provides high range and bearing accuracy, implements a proven interference rejection method, and classifies multiple targets. It does this all with a single sensing of the environment, and in a package smaller than previously achieved – [1] has a transmitter separation of 260 mm, this paper shrinks this to 40 mm. The repetition rate depends on the number of pulses processed (clutter in the environment), and is typically 15-27 Hz.

This paper is organised as follows. Section 2 introduces the sensor hardware and configuration while the association of echo arrival times is discussed in section 3. Geometrical models of the target types are presented in section 4, so that classification algorithms can be described in section 5 where two approaches are discussed: the *delta classifier* that is based on the bearing difference resulting from firing the two transmitters, and the MLE classifier that exploits all four times of flight directly. Section 6 presents experimental results from the sensor.

Table 1 Hardware	
T1	right-hand transmitter
T2	left-hand transmitter
R1	right-hand receiver
R2	left-hand receiver
T1'	virtual image of T1
T2'	virtual image of T2
d	separation between transceivers (m)
Delta Classifier	
R₁	θ_1 measured coordinates of T1'
R₂	θ_2 measured coordinates of T2'
delta	$\theta_1 - \theta_2$
ω	theoretical delta value
MLE Classifier	
rm_{m n}	measured range from T _m to R _n .
re_{m n}	estimated range from T _m to R _n .

2 Four TOF Sensor

An important feature of sonar is that time of flight errors are well correlated in time and space [1, 2]. By placing two receivers close together, we can still obtain a very accurate measurement of bearing despite the short baseline [1], because the bearing calculation depends critically on the difference between the two times-of-flight from a single transmitter to the two receivers. The error in this difference is typically much smaller than the straight time of flight errors since these are highly correlated.

Additionally placing the two transmitters close together produces highly correlated errors for all the four time-of-flight measurements of a target. This is important because the dominant factor in determining the class of a target is the difference between the two measured bearings.

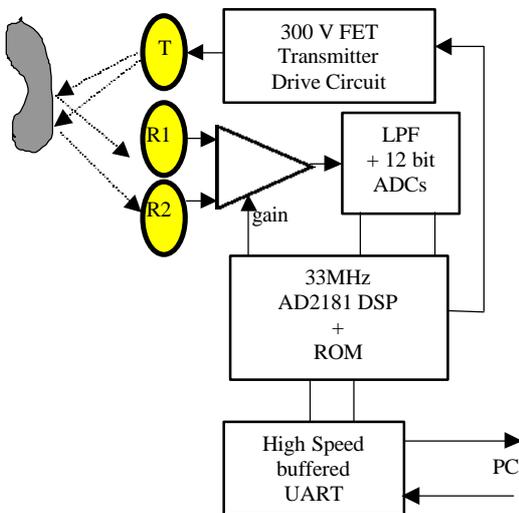


Figure 1 – Hardware Block Diagram

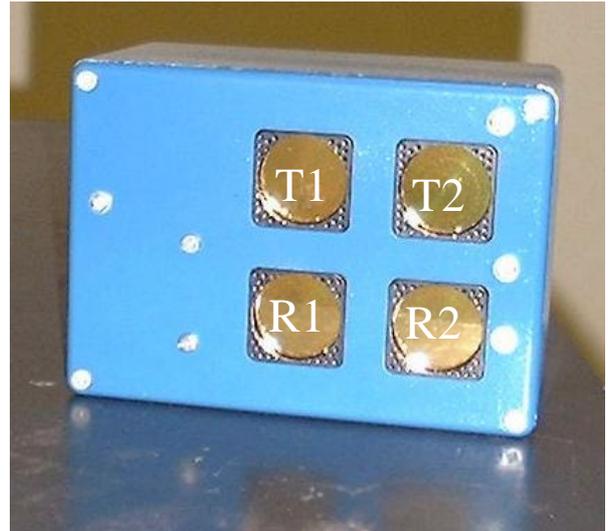


Figure 2 – the sensor unit, showing the layout of the transmitters T1, T2, and the receivers R1, R2.

Since both classifiers described in Section 5 depend primarily on differences in time-of-flight, this results in low error rates from the classifiers.

Our sensor package measures about 15 cm x 10 cm x 7 cm, it is powered by a single 5 V supply and communicates with its host by a high-speed serial link. Two transmitters are mounted above two receivers, forming a square with only 40 mm between centres. We use Polaroid 7000 series transducers with the front grille removed.

The received signals are amplified, low pass filtered and digitised at 1 Mz and 12 bit precision, then processed on an Analog Devices 33 MHz ADSP2181. The DSP also generates the transmit waveforms and communicates with the host via an external UART.

A pulse is fired from the right transmitter first, and rapidly followed from the left (200 μ s delay is typical). Echoes are digitised and processed on a DSP, yielding up to four arrival times for reflections from each target.

We find it clearer to think of these as arrival times, rather than times-of-flight, since at this point we do not know whether an echo is a reflection of the first or second pulse fired, so actual time-of-flight is ambiguous.

3 Forming Tuples

Matching up all the echoes from a single target is not trivial. Sometimes, not all four echoes will be detected. In a cluttered environment, they may be interleaved with echoes from another target.

Both transmitted pulses are the same shape, as this accelerates processing (fewer matched filters are required), but means that echoes of first and

second pulses are indistinguishable. However, geometry provides a means to differentiate them.

We rely on the fact that the time-of-flight from a transmitter at A to a receiver at B is the same as going from B to A; and knowledge of the precise interval between the two transmitted pulses. Since our hardware has separate transmitters and receivers, we use the times-of-flight from T1 to R2, and from T2 to R1. By symmetry, these times are always equal for vertical targets (plane, corner or edge), because they reflect sound from the point where they intersect the horizontal plane through the sensor.

The first phase is a search for a pair of received pulses bearing the same interval as the transmitted pulses. We call these 'double pulse pairs'. For example, if the transmitted pulses were separated by 200 μs , and we get a pulse on the R2, then 200 μs later ($\pm 1 \mu\text{s}$, or so) we get a pulse on R1, we can reasonably assume that the first pulse had been transmitted by T1, and the second by T2. The two received pulses should also be of similar amplitude.

The converse does not apply – that is, a pulse at R1 followed 200 μs later by a pulse at R2 does not indicate a valid target. The timing of such pulses depends on the bearing to the target, so it is not useful for discriminating genuine signal from interference.

The second stage is a search for the same reflected pulse at both R1 and R2. These two times-of-flight allow us to determine the bearing of a target, so we call them a 'bearing pair'. The criteria for forming such a pair are that the amplitudes are similar (within a factor of 2) and the times-of-flight differ by less than 22 μs (corresponding to about 11° from normal incidence). Additionally, if the pairing of a given pulse is ambiguous it is ignored rather than producing possibly incorrect results.

4 Target Models

To determine the class of target indicated by a set of returns, or tuple, we must know what to expect in each case. That is, we want to know the relationships that hold between the four ranges, designated r_{11} , r_{12} , r_{21} and r_{22} ; or the relationship between the two measured bearings, θ_1 and θ_2 . These relationships are different for each of the three target classes we consider.

4.1 Plane

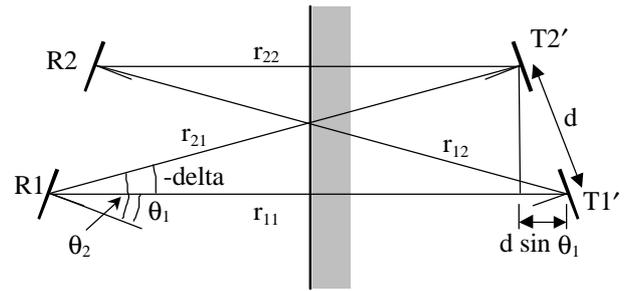


Figure 3 – Reflections from a plane.

From the geometry we have

$$r_{22} = r_{11} - 2d \sin q_1 \quad (1)$$

Applying the cosine rule to the triangle formed by R1, R2, T1':

$$\begin{aligned} r_{12}^2 &= d^2 + r_{11}^2 - 2dr_{11} \cos(90 - q_1) \\ &= d^2 + r_{11}(r_{11} - 2d \sin q_1). \end{aligned} \quad (2)$$

So by substitution

$$r_{12} = \sqrt{d^2 + r_{11}r_{22}}. \quad (3)$$

4.2 Corner

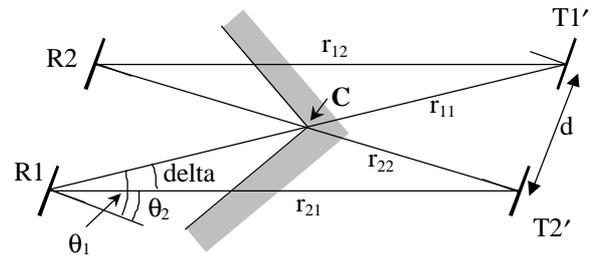


Figure 4 – Reflections from a corner.

The lines labelled r_{11} and r_{22} bisect each other at the corner, C.

We apply the cosine rule twice more, first to the triangle R1, R2, T1' and multiply the result by 2 then to triangle R1, R2, C and multiply the result by 4:

$$2r_{12} = 2d^2 + 2r_{11}^2 - 4dr_{11} \sin q_1 \quad (4)$$

$$r_{22}^2 = 4d^2 + r_{11}^2 - 4dr_{11} \sin q_1 \quad (5)$$

Take the difference and solve for r_{12} :

$$r_{12} = \sqrt{(r_{11}^2 + r_{22}^2)/2 + d^2} \quad (6)$$

4.3 Edge

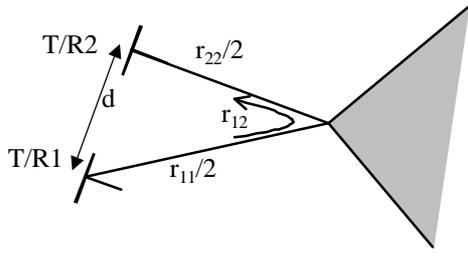


Figure 5 – Reflections from an edge.

For an edge, we can read r_{12} directly from the diagram:

$$r_{12} = (r_{11} + r_{22}) / 2. \quad (7)$$

4.4 Summary

$$\begin{aligned} r_{12} &= \sqrt{r_{11}r_{22} + d^2} && \text{– Plane} \\ r_{12} &= \sqrt{(r_{11}^2 + r_{22}^2) / 2 + d^2} && \text{– Corner} \\ r_{12} &= (r_{11} + r_{22}) / 2 && \text{– Edge} \\ r_{21} &= r_{12} && \text{– Always} \end{aligned} \quad (8)$$

5 Classification

Depending on the classification method used, we may calculate the range and bearing of the virtual image of each transmitter, and use the difference between bearings to classify the target. Alternatively, a maximum likelihood estimator is applied to the four range measurements to obtain a classification directly.

5.1 Delta classifier

When an echo of a pulse is detected at two receivers, we can triangulate to determine the bearing to the target. More precisely, if the target is a plane or corner, we can determine the bearing to the virtual image of the transmitter reflected in the target. Call this θ_1 . If the target is an edge then the same calculation is a good approximation of the target's position [1].

When two transmitters are used, the difference between the bearings thus obtained can identify the target type. We define

$$\text{delta} = \mathbf{q}_1 - \mathbf{q}_2. \quad (9)$$

Then, theory [1] predicts

$$\text{delta} = \begin{cases} 0, & \text{for an edge} \\ \mathbf{w}, & \text{for a corner} \\ -\mathbf{w}, & \text{for a plane} \end{cases} \quad (10)$$

$$\text{where } \mathbf{w} = \frac{1}{2} \tan^{-1} \left(\frac{\text{ReceiverSeparation}}{2 \times \text{RangeOfTarget}} \right).$$

Using thresholds at $\pm \frac{1}{2} \mathbf{w}$, a target can be classified as edge, plane or corner.

5.2 MLE classifier

When doing target classification using MLE, the measured data is compared with the model for each case, obtaining for each the most likely actual range values and also a least squares error that indicates which classes are plausible matches to the measured data.

Compared with the delta classifier, this has the advantage of fully utilising the available measurements, whereas the delta classifier only considers differences, and of producing an improved estimate of actual ranges in the cases where classification is successful.

5.2.1 Maximum Likelihood Estimation

Maximum likelihood estimation is a technique that uses noisy measurements of a system and determines the most likely actual state. We apply it to the four measured ranges, rm_{11} , rm_{12} , rm_{21} and rm_{22} , to obtain an estimate of the actual ranges, re_{11} and re_{22} , which are sufficient to fully describe the target position.

Let us begin by defining

- \mathbf{Y} – a vector of k noisy observations,
- \mathbf{X} – a state vector of i parameters,
- F – a function relating \mathbf{X} to \mathbf{Y} , and
- \mathbf{N} – the k noise components.

Then we can write the non-linear equation

$$\mathbf{Y} = F(\mathbf{X}) + \mathbf{N}. \quad (11)$$

If the system is linear with Gaussian conditional probability functions and Gaussian noise, then MLE can be applied. We linearise the non-linear model equations (8) obtained in section 4, about the measured values $\mathbf{X}_m = [rm_{11} \ rm_{22}]^T$. We define different variables to apply to the linear case:

- \mathbf{A} – a k dimensional observation vector
- \mathbf{B} – an i dimensional state vector
- \mathbf{J} – $k \times i$ Jacobian of F elaborated about \mathbf{X}_m
- \mathbf{N} – the k dimensional noise vector
- \mathbf{R} – the $k \times k$ noise covariance matrix

and write the linear equation

$$\mathbf{A} = \mathbf{JB} + \mathbf{N}. \quad (12)$$

Linearisation about \mathbf{X}_m gives

$$F(\mathbf{X}) - F(\mathbf{X}_m) \cong \mathbf{J} \cdot [\mathbf{X} - \mathbf{X}_m] \quad (13)$$

so now we rewrite the non-linear form from equation (11) as

$$\begin{aligned} \mathbf{Y} - F(\mathbf{X}_m) &= F(\mathbf{X}) - F(\mathbf{X}_m) + \mathbf{N} \\ &\cong \mathbf{J} \cdot [\mathbf{X} - \mathbf{X}_m] + \mathbf{N} \end{aligned} \quad (14)$$

which we recognise as equation (11) with

$$\begin{aligned} \mathbf{A} &= \mathbf{Y} - F(\mathbf{X}_m) \\ \mathbf{B} &= \mathbf{X} - \mathbf{X}_m. \end{aligned} \quad (15)$$

The maximum likelihood estimate of \mathbf{B} , $\hat{\mathbf{B}}$, and the least squares error, S , are given by [7]

$$\begin{aligned} \hat{\mathbf{B}} &= [\mathbf{J}^T \mathbf{R}^{-1} \mathbf{J}]^{-1} \mathbf{J}^T \mathbf{R}^{-1} \mathbf{A}, \\ S &= [\mathbf{J} \hat{\mathbf{B}} - \mathbf{A}]^T \mathbf{R}^{-1} [\mathbf{J} \hat{\mathbf{B}} - \mathbf{A}]. \end{aligned} \quad (16)$$

Our input to the estimator is therefore

$$\mathbf{A} = \begin{bmatrix} 0 \\ 0 \\ rm_{12} - r_{12}(rm_{11}, rm_{22}) \\ rm_{21} - r_{21}(rm_{11}, rm_{22}) \end{bmatrix}, \quad (17)$$

and the Jacobian is

$$\mathbf{J} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial r_{12}}{\partial r_{11}} & \frac{\partial r_{12}}{\partial r_{22}} \\ \frac{\partial r_{21}}{\partial r_{11}} & \frac{\partial r_{21}}{\partial r_{22}} \end{bmatrix} \quad (18)$$

where the functions r_{ij} are from the appropriate form of equation (8), derived in section 4. The result vector is

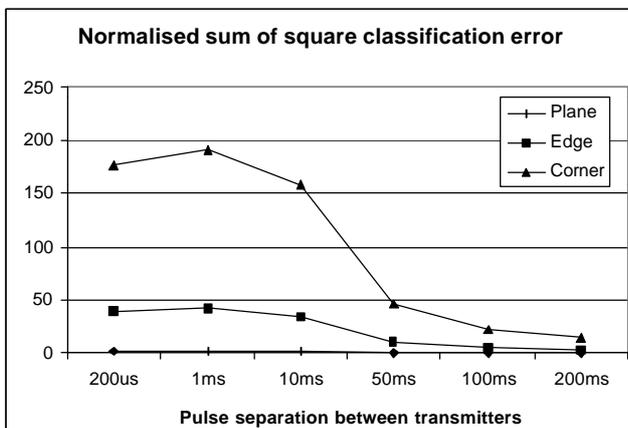


Figure 6 – Dependence of S on pulse separation.

$$\hat{\mathbf{B}} = \begin{bmatrix} re_{11} - rm_{11} \\ re_{22} - rm_{22} \end{bmatrix}, \quad (19)$$

thus yielding the estimated ranges, re_{11} and re_{22} .

5.2.2 Classification

The normalised least squares error, S , is minimised by MLE. The resulting value is a measure of how well the observed data fits the linear form in equation (11), and has a χ^2 distribution with $k - i$ degrees of freedom. We have $k = 4$, $i = 2$, therefore two degrees of freedom. A different value for S is obtained for each target class; the value corresponding to the correct target class will likely be much smaller than those for the incorrect classes.

We can set a threshold for S , c , specifying that only some small proportion, ϵ , of correctly classified measurements shall exceed c . That is,

$$P(S > c) = \epsilon. \quad (20)$$

Then the condition for accepting a given classification is that $S \leq c$. For the χ^2 distribution with two degrees of freedom, and a false rejection rate (ϵ) of 5%, $c = 5.99$.

If the target is found to belong to exactly one class, that class is accepted as the object's class. Otherwise its class remains unknown.

6 Results

The purpose of the first experiment described in this section is to show how the noise due to increasing pulse separation of the right and left transmitters, affects the MLE's ability to discriminate between target classes. Conversely, we aim to show the advantage of the high error correlation experienced when using short delays.

The experiment consists of sensing a wall at 3 m range using different pulse separations, from 200 μ s to 200 ms. These measurements were repeated 200 times to obtain experimentally four mean ranges, which may be taken as an ideal measurement of the wall, and the error covariance, \mathbf{R} of the measurements. Then we applied MLE for each target class and each pulse separation (i.e. differing covariance) to the mean, to obtain the least squares error, S .

Figure 6 shows the results. The top trace may be viewed as "This is how much unlike a corner the wall appeared." The second trace is similar, but contemplating if the wall might be an edge. It is immediately obvious that although the target is accepted as a wall in all cases, for separations of 50 ms and over, the confidence that it could not be anything else (that is, definitely not a corner or

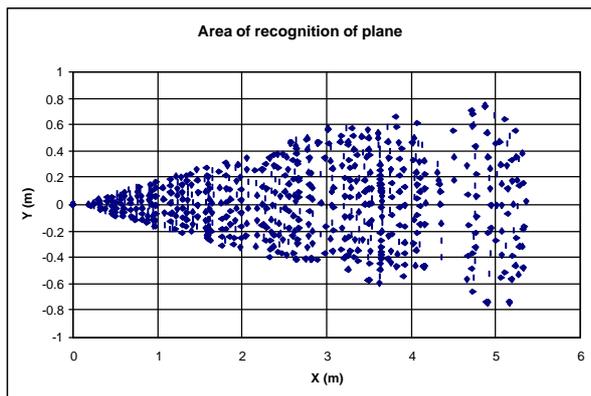


Figure 7 – Plane recognition region.

edge), is much diminished. In fact, for separations of 100 ms and 200 ms, the wall was accepted as a possible edge. The experiment was conducted under air conditions that were deliberately turbulent – a fan and an air conditioner were operating in the laboratory. We conclude that for durations up to 10 ms in these conditions, the air column may be assumed approximately static and errors are well correlated, but by 50 ms this assumption has broken down.

Another experiment was conducted to illustrate the region of recognition of planes that can be achieved. Figure 7 shows points corresponding to detected planes as the sensor is swept at various ranges. The shape of this region is constrained by the 22 μ s bearing association window described in section 3, and the 5.4 metre range limit.

7 Conclusions

This paper has described an approach to target localisation and classification that for the first time also integrates an interference rejection scheme proposed in [2]. The classification into corners, planes and edges, is performed with virtually no overhead, and very low latency – a significant improvement on previous systems.

Firing pulses from two closely spaced transmitters with a time separation in the order of 1 ms provides three important functions. Sufficient echoes are returned from a target to enable its classification. Accuracy is improved by utilising the high correlation of errors in the air column. Finally, the precise time separation serves to identify the sonar system and enable interference rejection.

The self-contained, real time DSP based sensing module is also smaller than previous sensors.

Further improvements in accuracy are expected by using a single transducer as both transmitter and receiver. Concurrent work on a sonar ring uses this design.

Since the classification is performed with such a short delay between transmitter firings, the sensor could be deployed on moving platforms to achieve on-the-fly mapping – this will be implemented in future work.

Acknowledgements

Steve Armstrong is gratefully acknowledged for technical support. The financial support of a large Australian Research Council Grant is acknowledged.

References

- [1] L. Kleeman and R. Kuc, "Mobile robot sonar for target localization and classification", *International Journal of Robotics Research*, Vol 14, No 4, August 1995, pp 295-318.
- [2] L. Kleeman, "Fast and accurate sonar trackers using double pulse coding", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Kyongju, Korea, October 1999, pp 1185-1190.
- [3] J Borenstein and Y Koren, "Error eliminating rapid ultrasonic firing for mobile robot obstacle avoidance", *IEEE Trans. Robotics and Automation*, Vol 11, No 1, pp 132-138, 1995.
- [4] K. Jorg and M. Berg, "Mobile robot sonar sensing with pseudo-random codes", *Proceedings 1998 IEEE Conference on Robotics & Automation*, Leuven, Belgium, May 1998, pp 2807-2812.
- [5] H. Peremans, K. Audenaert, and J. M. V. Campenhout, "A high-resolution sensor based on tri-aural perception," *IEEE Transactions on Robotics and Automation*, Vol 9, pp 36-48, 1993.
- [6] K S Chong and L. Kleeman, "Feature-based mapping in real, large scale environments using an ultrasonic array", *International Journal Robotics Research*, Vol 18, No 1, Jan 1999, pp 3-19.
- [7] A D Whalen, "Detection of Signals in Noise", Academic Press, New York, 1971a.
- [8] Hong Mun-Li and Lindsay Kleeman, "Ultrasonic classification and location of 3D room features using maximum likelihood estimation", *Robotica*, Vol 15, pp 483-491 and 645-652, 1997.