Line Extraction from Mechanically Scanned Imaging Sonar

David Ribas[1], Pere Ridao[1], José Neira[2], and Juan Domingo Tardós[2]

[1] University of Girona
[2] University of Zaragoza

Abstract. The extraction of reliable features is a key issue for autonomous underwater vehicle navigation. Imaging sonars can produce acoustic images of the surroundings of the vehicle. Despite of the noise, the phantoms and reflections, we believe that they are a good source for features since they can work in turbid water where other sensors like vision fail. Moreover, they can cover wide areas incrementing the number of features visible within a scan. This work presents an algorithm to extract linear features from underwater structured environments including as major contributions a novel sonar model sensor and an adapted implementation of the Hough transform.

1 Introduction

Vehicle localization, map building and more recently, the simultaneous localization and mapping (SLAM) are fundamental problems to achieve true autonomous vehicles [1], [2]. One of the key issues on those techniques is to develop reliable systems to extract features from the environment in order to build maps or navigate thorough existing ones. Working in underwater environments is specially challenging because of the reduced sensorial possibilities. Acoustic devices are the most common choice while the use of cameras and laser sensors is limited to applications where the vehicle navigates very near to the seafloor. One of the issues on working on this kind of environments is the difficulty on finding reliable features. There are approaches using clusters of acoustic data as features [3], [4], some merge visual and acoustic information in order to improve the reliability [5], while others simply introduce artificial beacons to deal with complex environments [6]. Most of the previous work using mechanically scanned imaging sonars (MSIS) have focused on the use of point features assuming the robot remains static or moves sufficiently slow. In this work we propose an algorithm to take advantage of structured elements typically present in common underwater scenarios (drilling platforms, harbours, channels, dams,...) in order to extract line features. Moreover, our algorithm removes the "static" assumption. This paper briefly introduces MSISs, presents a novel sonar model which improves the sensor measurements characterization and depicts a modified version of the Hough transform algorithm for detecting line features (cross sections of planar structures present in the environment) in imaging sonar scans. Finally, the results and conclusions are presented.

2   Extracting Line Features from Acoustic Images

MSISs perform scans in a 2D plane by rotating a sonar beam through a series of small angle steps. For each emitted beam, distance vs. echo-amplitude data is returned. Thus, accumulating this information along a complete 360° sector, an acoustic image of the surroundings can be obtained (Fig. 1). Commonly, these devices have a slow scanning rate (f.i. a Tritech Miniking sonar head needs about 6 seconds to complete a 360° scan). For this reason, the vehicle movement along a complete scan usually induces important distortions in the acoustic image (Fig. 1b). Extracting features from this kind of images produces inaccuracies and yields to poor results. Therefore, the first step of the procedure consists on merging the raw sensor data together with the information from the vehicle's navigation system [7]. Incorporating the displacements and rotations of the sensor into the positional information of each sonar measurement leads to an undistorted acoustic image such the one represented in Fig. 1c.
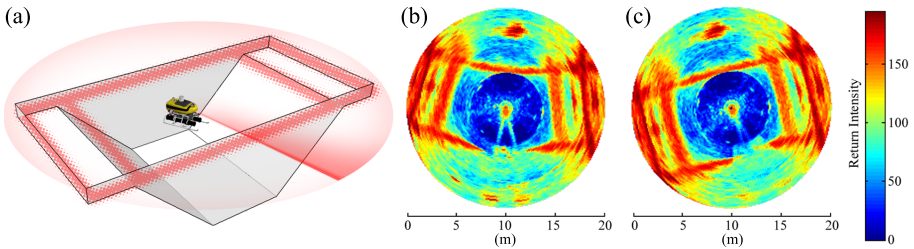

Fig. 1. (a) Schematic representation of the environment where the sonar data were gathered. The highlighted zones represent the expected sonar returns. Images generated from acoustic data, (b) distorted and (c) undistorted image through navigation integration.

2.1   Beam Segmentation

At a given orientation of the transducer head, the sensor gathers an acoustic profile of the surroundings which is represented by a vector of echo amplitude values (called bins) each one corresponding to a particular distance along the emitted beam. Since objects present in the environment appear as high echo-amplitude returns, only part of the information stored in the vector is useful for feature extraction. Therefore, a segmentation process can be done in order to get the more significant information from the acoustic beam and, in addition, reduce the computational cost of the algorithm. In the context of this work several approaches have been tested (Fig. 2):

- Highest intensity: The bin with the maximum value over a threshold is selected.
- Peak values: The set of local maxima over the threshold are selected. Moreover, they must accomplish a "minimum distance between them" criterion.
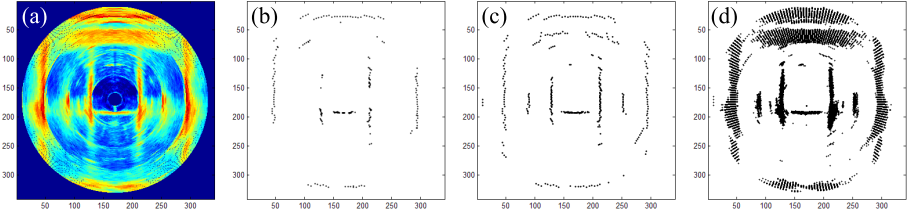
Fig. 2. (a) Acoustic image generated with the complete set of measurements, (b) highest intensity return selection, (c) peak values selection, (d) thresholding

.

- *Thresholding:* All the bins with values over a threshold are selected. The performance of these approaches is reported in section 3.

2.2 Sonar Modeling

Each single beam's bin represents the strength of the echo intensity return from a particular place within the insonified area (Fig. 3a). The acoustic beam projected along the central axis of the transducer has a 3° horizontal beamwidth and a 40° vertical beamwidth. As the sonar data is represented in the horizontal plane, the uncertainty due to the vertical aperture produces a blurred representation of the features along the beam (Fig. 3a). On the other hand, the uncertainty from the horizontal beamwidth is not implicitly represented in the resulting data. Typically, in order to represent this uncertainty, each bin can be described as an arc showing all the possible locations of tangent surfaces producing compatible sonar returns [8],[9]. While this simple model is well suited for air sonar ranging systems, it is not able to explain the acoustic images gathered with a MSIS. A careful analysis on such images (see Fig. 3b) reveals that their object detection capability is not limited to the arc-tangent surfaces. Even those beams which meet a surface with a considerable incidence angle (for the Miniking, $\beta = 60°$) produce a discernible high intensity profile. For this reason, we have adopted an extended model to describe the imaging sonar. Each bin represents a zone described by an arc which corresponds to the horizontal beamwidth α (in our sensor, $\alpha = 3°$). Given a resolution and the incidence angle β, for each point belonging the arc, its tangent surface as well as the intersecting planes with an incidence angle smaller than $\pm\beta/2$ are visible for the beam's bin (Fig. 4). Hence, the acoustic intensity represented by the bin should correspond to one of those candidate planes.

2.3 The Hough Voting Space

The Hough transform [10] accumulates the information from the sensor data into a voting table which is a parameterized representation of all the possible feature locations. Those features that receive a great number of votes are the ones with a relevant set of compatible sensor measurements and thus the ones
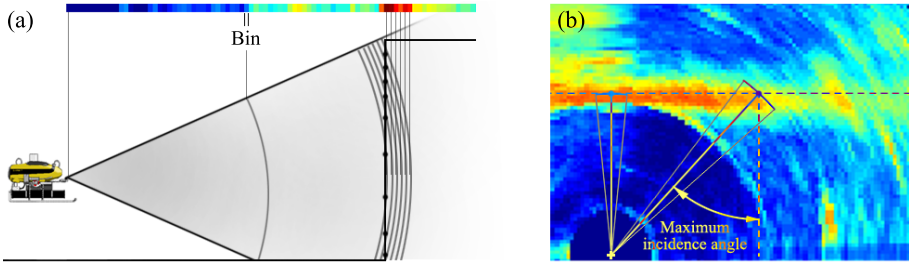
Fig. 3. (a) Dispersion of the high intensity returns provoked by the vertical aperture of the sonar beam, (b) representation of the horizontal beamwidth as an arc. Note that at the maximum incidence angle the line feature is still distinguishable although the arc is not tangent.

that most likely correspond to a real object in the environment. We have chosen line features to represent the planar objects existing in the scene. These line features are described by two parameters, ρ^B and θ^B (distance and orientation with respect to the base frame B). Hence, the resulting Hough space is a two-dimensional space where the voting process and the search for maxima can be done efficiently. Another key issue is the quantization of the Hough space. In our case, we have observed that tuning the quantization to match the angular and lineal resolutions of our sensor (typically, 0.1 m and 1.8°) works fine.

2.4   Voting

Each time a beam is obtained, the segmented bins vote in the Hough space. The next step is to determine the candidate lines that will receive the votes. As previously introduced in section 2.2, the measurement is modeled as an arc in order to represent the uncertainty that appears due to the horizontal beamwith α (Fig. 4). Hence, θ_{S_j} will take values within an aperture of $\pm\alpha/2$ around the real angle of the transducer head. Then, for each θ_{S_j} value, a set of k candidate lines will be determined. As said before, not only the lines tangent to the arc are candidates, but also the ones inside the maximum incidence angle limits of $\pm\beta/2$. So, for each θ_{S_j} value we can define θ_k^B as:

$$\theta_{S_j} - \frac{\beta}{2} \leq \theta_k^B \leq \theta_{S_j} + \frac{\beta}{2}\,. \tag{1}$$

Finally, the ρ_k^B value that corresponds to each value of θ_k^B is calculated as:

$$\rho_k^B = x_{S_j} cos(\theta_k^B) + y_{S_j} sin(\theta_k^B) + \rho^{S_j} cos(\theta_k^{S_j})\,. \tag{2}$$

In Fig. 5a it is shown how the set of voters corresponding to a single sonar return looks like when assigned to the Hough space. Note that each selected cell of the Hough space only receives one single vote. However, we have tested two different
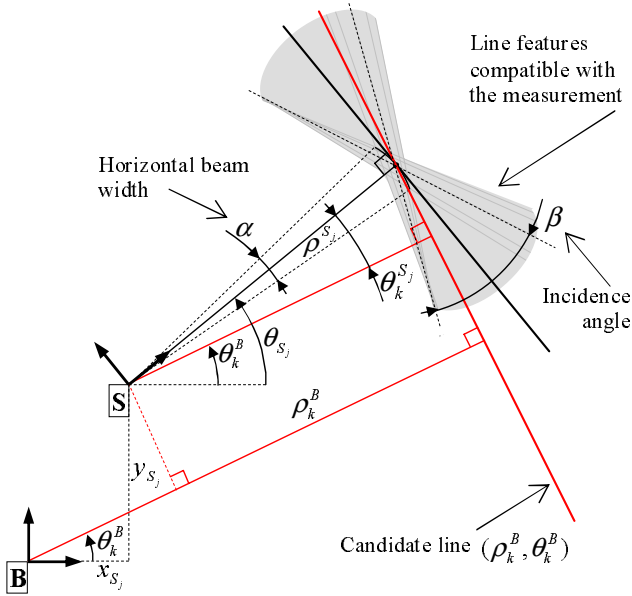
Fig. 4. Model of the sonar sensor for line features. Where B is the base reference frame and S is a reference frame attached to the current beam.

ways to assign values to the votes. The first method assigns to each vote a value of 1, while the second gives to the vote a value related with the echo intensity of its corresponding bin. In this way, we try to give more weight to those sonar returns that more strongly reflect the presence of objects in the environment.

As the sensor constantly produces measurements, the number of votes assigned to the space grows without limit. Moreover, when a 360° scan around the sensor is completed, the new upcoming measurements would correspond to a zone already visited and hence, redundant votes could corrupt the candidate election. To avoid this, we tagged each bin with the angle of its corresponding beam so the oldest voters can easily be identified and discarded. The key issue of this idea is determining which measurements are identified as old voters. A rough option is to consider the voters tagged one 360° scan away from the actual beam orientation. A smarter approach is to determine the minimum scan sector necessary to detect a line feature. In the extreme case, a line can exist in a 180° sector (when the feature is very close to the sonar head so it is placed at the diameter of the scan). Therefore, each time a new measurement is incorporated, a voting process is performed for the measurements corresponding to the most recent 180° sector. Then, a search among the potential candidate lines is done. If a line obtains a sufficient number of votes (this changes depending on many factors such as the chosen beam segmentation procedure or the discretization of the voting space) it is considered as detected and thus, their voters can be eliminated (one bin can only correspond to a single line feature in our
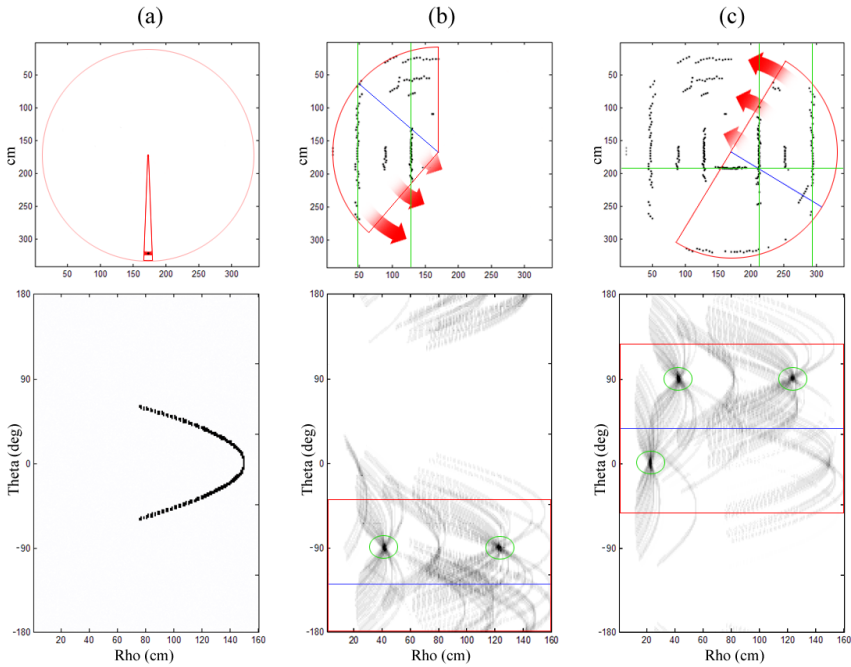
Fig. 5. (a) Votes assigned to the voting space for one single measurement. (b) Aspect of the voting space after a small scan sector. The actual voting scan sector is marked by a red line, the present features are in green, the blue line indicates the features detection zone. (c) Aspect of the voting space after almost one complete scan. Note that the old voters (the ones before the 180° scan) have been removed from the space.

application). If there is no line detected, the votes will be stored and used when new measurements arrive. This solves the problem of continuously dealing with new information provided by the sensor while, at the same time, keeps the number of votes low improving the overall efficiency of the algorithm (Fig. 5b and c).

2.5   Line Extraction

The algorithm looks for winning candidates each time a new beam arrives. However, we have to ensure that the algorithm detects the line when it has received all the possible votes (or what is the same, the line is totally inside the scan sector and the newest sonar beam cannot provide more information to it). Analyzing all line features existing in the sector scan it is possible to determine when they have received all the votes. A simple and robust strategy consist on setting the zone for the line detection 90° away from the last beam measurement, just in the bisector of the 180° scan sector (see the blue lines in Fig. 5b and c). Beyond this point, all the feature lines had been detected and hence, all the votes assigned to this lines had been removed so they cannot interfere with the detection of further features.
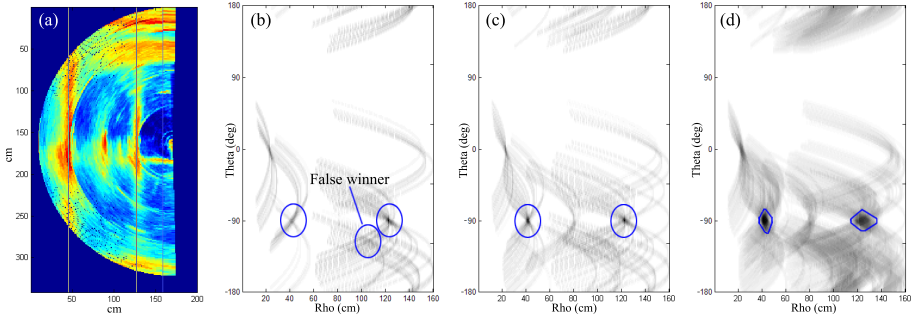
Fig. 6. (a) 180° scan sector used in the example, (b) aspect of the voting space when using only the highest value of each beam, (c) when using peak values, (c) when using a threshold for measurement selection. Note that the shapes of the peaks reflect the uncertainty of the lines in the scan.

3   Results and Conclusions

Several experiments were carried out gathering acoustic images in a water tank with a Tritech Miniking sonar mounted in an AUV. A Sontek Argonaut DVL sensor incorporating a magnetic compass was used to undistort the images. Selecting the parameters (threshold, incidence angle, minimum distance between local maxima, ...) according to the particular characteristics of our sensor and using the extended sonar model proposed in this paper, all the segmentation strategies were able to detect the line features. The proposed voting space (last 180° scan sector) together with the line extraction strategy allows to deal continuously with the sonar beams as provided by the sonar instead of waiting to collect a full scan before detecting the lines. With the proposed strategy, the maximum delay between the time instant in which the last bin of a line has been measured and the detection instant is 1.5 seconds (90° scan time) instead of the 6 seconds needed for mapping a full slam. in the future we will explore alternatives to reduce this delay to zero, looking for the Hough peaks not in linear region but in a nonlinear region.

Fig. 6 shows a representative experiment comparing the different beam segmentation strategies, using votes with a value of one. In our experiments, voting proportionally to the acoustic intensity didn't improve results appreciably. Hereafter we describe the conclusion related to each segmentation strategy:

– *Highest intensity:* This method is the one computationally more efficient because of the small number of votes. After the voting process, the winners appear in the Hough space as identifiable small peak zones. However, the low number of votes make easier the existence of "false winners" (a small but elongated shape can produce a similar number of votes that a line feature with lower intensity measurements, see Fig. 6b).
– *Peak values:* This method seems to be a good compromise between low computational cost and robustness. The higher number of votes gives a better

description of the elements present in the scene. The resulting Hough space is similar to the one obtained with the previous method. However, it seems to be more "contrasted" as winners present a much higher concentration of votes (Fig. 6c).

– *Thresholding:* This is by far the less efficient approach due to the large number of measurements involved in the voting. However, it presents a interesting property that the other methods had not shown: The thickness of the linear features present in the sonar scan seems to be related to the size and shape of the winner peaks in the Hough space (see Fig. 6d). As introduced in section 2.2, the vertical aperture of the fan shaped beam, among other causes, provokes a thickening of the elongated form corresponding to the line feature. This effect leads to uncertainty on the estimate of the real position of the feature and, for some applications, such as robot localization or SLAM, knowing this uncertainty is crucial. We think that, using this method, we can probably not only obtain the parameters of the existing features but also its related uncertainty. Further work will be done on this matter.

References

1. Durrant-Whyte, H., Bailey, T.: Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. Robotics and Automation Magazine (June 2006)
2. Durrant-Whyte, H., Bailey, T.: Simultaneous Localisation and Mapping (SLAM): Part II State of the Art. Robotics and Automation Magazine (September 2006)
3. Leonard, J.J., Carpenter, R.N., Feder, H.J.S.: Stochastic Mapping Using Forward Look Sonar. Robotica 19, 341 (2001)
4. Tena, I., Petillot, Y., Lane, D.M., Salson, C.: Feature Extraction and Data Association for AUV Concurrent Mapping and Localisation. In: Proc. of the IEEE International Conference on Robotics and Automation, Seoul, Korea (2001)
5. Williams, S., Mahon, I.: Simultaneous Localisation and Mapping on the Great Barrier Reef. In: Proc. of the IEEE International Conference on Robotics and Automation, New Orleans, USA (2004)
6. Newman, P.M., Leonard, J.: Pure range-only sub-sea SLAM. In: Proceedings of the IEEE International Conference on Robotics and Automation, Taipei, Taiwan (2003)
7. Ribas, D., Neira, J., Ridao, P., Tardós, J.D.: SLAM using an Imaging Sonar for Partially Structured Environments. In: Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China (2006)
8. Leonard, J.J., Durrant-Whyte, H.F.: Directed Sonar Sensing for Mobile Robot Navigation. Kluwer Academic Pub. London (1992)
9. Tardós, J.D., Neira, J., Newman, P., Leonard, J.: Robust Mapping and Localization in Indoor Environments using Sonar Data. International Journal of Robotics Research 21(4), 311–330 (2002)
10. Illingworth, J., Kittler, J.: A survey of the Hough transform. Computer Vision, Graphics, and Image Processing, Academic Press Professional 44(1), 87–116 (1988)