

A METHOD FOR EXTRACTING LINES AND THEIR UNCERTAINTY FROM ACOUSTIC UNDERWATER IMAGES FOR SLAM

David Ribas* Pere Ridao* José Neira**
Juan Domingo Tardós**

* *Universitat de Girona, Spain*

** *Universidad de Zaragoza, Spain*

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 phantom echoes 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, increasing 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 sonar sensor model, an adapted implementation of the Hough transform and a novel procedure to estimate feature uncertainty.

Keywords: Robot navigation, sensors, feature extraction, 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 (Durrant-Whyte and Bailey, 2006), (Bailey and Durrant-Whyte, 2006). 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 (Leonard *et al.*, 2001), (Tena *et al.*, 2001), or merging visual and acoustic information in order to improve the reliability (Williams and Mahon, 2004), while

other strategies simply introduce artificial beacons to deal with complex environments (Newman and Leonard, 2003). 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 (cross sections of planar structures present in the environment). Moreover, our algorithm removes the “static” assumption. This paper briefly introduces MSISs, a novel sonar model which improves the sensor measurements characterization, an algorithm for line feature extraction from imaging sonar scans based on a modified version of the Hough transform and a methodology for feature uncertainty estimation. Finally, the results and conclusions are presented.

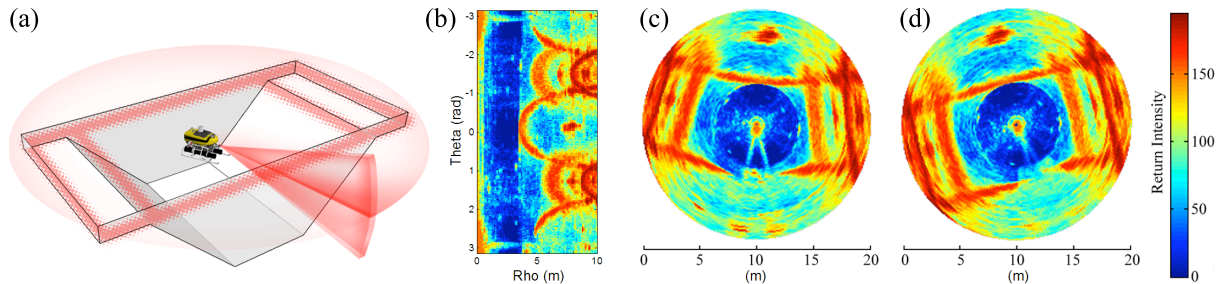


Fig. 1. (a) Schematic representation of the environment where the sonar data were gathered. (b) Raw data represented in polar coordinates. Each row corresponds to a single beam at a given orientation and each column to a bin in a particular position along the beam. (c) The same data represented in cartesian coordinates. Notice the distortion produced by the movement of the vehicle. (d) Undistorted image after integration with vehicle displacement.

2. WORKING WITH ACOUSTIC UNDERWATER IMAGES

MSISs perform scans in a 2D plane by rotating a fan-shaped sonar beam through a series of small angle steps. For each emitted beam, a set of bins (distance vs. echo-amplitude values) are obtained (Fig. 1b). Accumulating this information along a complete 360° sector produces an acoustic image of the surroundings (Fig. 1c). The beam typically has a large vertical beamwidth which makes possible the detection of obstacles at different heights. In the other hand, a narrow horizontal beamwidth increments the resolution of the device and improves the sharpness of the acoustic images.

2.1 Beam segmentation

Objects present in the environment appear as high echo-amplitude returns in acoustic images. Thus, only part of the information stored in each beam is useful for feature extraction. Therefore, a segmentation process can be done in order to obtain the more significant information. This process consists in two steps. First, only those bins with an intensity value over a threshold are selected and stored. This procedure separates the acoustic imprint left by an object in the image, from the noisy background data (Fig. 2b). The resulting imprint is used to estimate the feature uncertainty as explained in Section 4. The second step is to select among the thresholded data those bins which are local maxima and satisfy a “minimum distance between them” criterion (Fig. 2c). These local high intensity bins are the ones that most likely correspond to objects present in the scene. Thus, they are specially well suited as input to the feature extraction algorithm (Section 3) while, at the same time, the computational efficiency is improved since a small number of bins are involved.

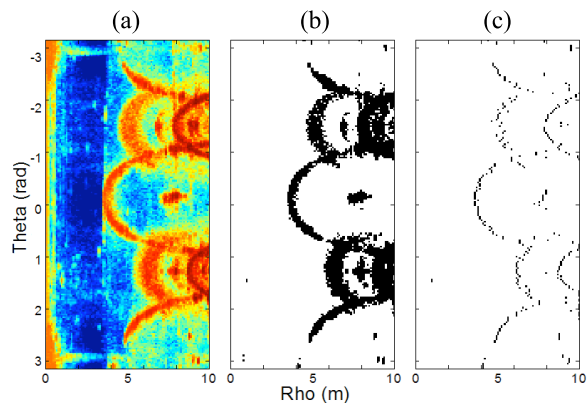


Fig. 2. (a) Raw sensor data. (b) Segmented data. (c) Selection of local maxima bins.

2.2 Undistorting the acoustic images

Commonly, MSISs 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. 1c). Extracting features from this kind of images produces inaccuracies and yields 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 (Ribas *et al.*, 2006). 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. 1d.

2.3 Dealing with a stream of beams

In order to deal with the stream of measurements produced by the continuous arrival of beams, we set a data buffer storing the beams contained within the most recent 180° scan sector. Whenever new beams corresponding to an unexplored zone arrive, old beams that fall outside the scan sector are discarded. The choice of a 180° sector is not arbitrary since this is the maximum zone that

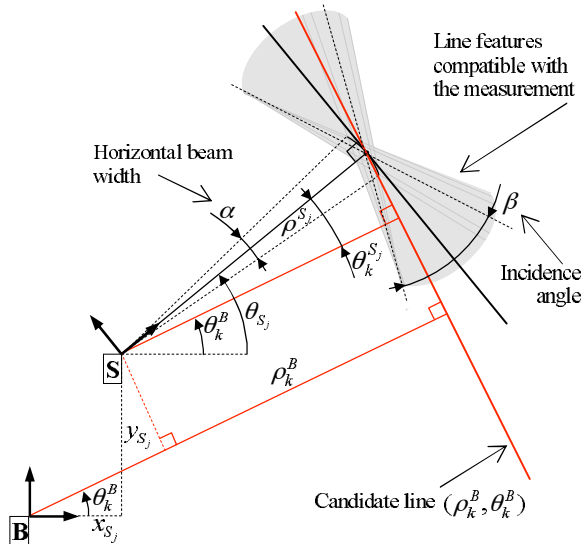


Fig. 3. Model of the sonar sensor for line features. Where B is the base reference frame and S is a reference frame attached to a beam.

a single line can cover within a sonar scan. Since calculations searching features are performed with every new beam (Section 3), the buffer should contain the bins which are local maxima (Fig. 2c) for the line detection process, the segmented beams (Fig. 2b) for uncertainty estimation and all its associated positions in the world coordinate system to deal with the movement-induced distortions.

3. DETECTION OF LINE FEATURES

As said before, walls and other planar structures produce line-shaped features in acoustic images. The Hough transform (Illingworth and Kittler, 1988) is a feature extraction technique which is specially well suited for this kind of situations. This algorithm 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 that most likely correspond to a real object in the environment. In our application, line features are described by two parameters, ρ^B and θ^B (distance and orientation with respect to a base frame B). Hence, the resulting Hough space (HS) is a two-dimensional space where the voting process and the search for maxima can be done efficiently. The base reference frame B can be set arbitrarily. However, our choice for B is the position of the sensor head when the last beam arrives. So, any detected line feature will be represented directly in the sensor coordinate frame. It is worth noting that B is not a fix coordinate frame. As the parametrization in the HS is performed in polar coordinates, setting

the reference in a fix position would produce a resolution loss with the increment of range. Hence, we need to resituate B near to the sensor as the vehicle moves. Unfortunately, this makes necessary to recalculate the HS with each change in the position of B. Another key issue is the quantization of the HS. In our case, we have observed that selecting the quantization equal to the angular and lineal resolutions of our sensor (typically, 1.8° and 0.1 m) works fine.

The feature extraction procedure works as follows: First, with each beam arrival the Hough space is initialized, and the buffered bins are referenced to B so they can be used to vote in such space. Then, the votes corresponding to each bin are assigned to the candidate lines by means of a sonar model. Finally, a search for winning candidates is performed.

3.1 Sonar Modeling

Each bin represents the strength of the echo intensity return in a particular place within theinsonified area. Due to the uncertainty produced by the horizontal beamwidth, a measurement cannot be assigned to a single point in the space. A common approach (Leonard and Durrant-Whyte, 1992), (Tardós *et al.*, 2002), is to represent a bin as an arc whose aperture represents the beamwidth uncertainty. Moreover, as a high intensity return is typically produced when the acoustic wave hits perpendicularly a surface, we can infer that all the surfaces tangent to the arc can explain the high intensity return. 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 of such images 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, $\beta = 60^\circ$) 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, $\alpha = 3^\circ$). 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. 3). Hence, the acoustic intensity represented by the bin should correspond to one of those candidate planes.

3.2 Voting

The next step is to determine the candidate lines that will receive the votes for each buffered bin. As

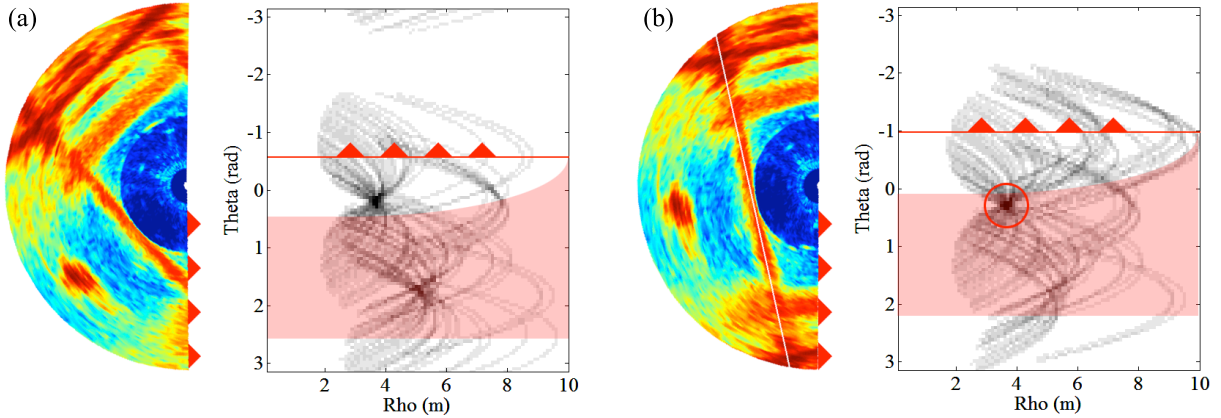


Fig. 4. Two examples of the voting process. The scan sector stored in the buffer is represented together with its corresponding generated voting space. The red triangular shapes mark the newest beam. The shaded zone represents where the candidates have received all the possible votes (a) Part of the target line is still outside the sector scan. Hence, it can receive more votes in the future. (b) The new beams can not add new votes to the target line. The line inside the shaded zone of the voting space have been fully observed and hence, it can be detected.

previously introduced in Section 3.1, the measurement is modeled as an arc in order to represent the uncertainty that appears due to the horizontal beam with α (Fig. 3). 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:

$$\theta_{S_j} - \frac{\beta}{2} \leq \theta_k^B \leq \theta_{S_j} + \frac{\beta}{2}. \quad (1)$$

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

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

In Fig. 4 it is shown how the set of voters looks like when assigned to the Hough space. Note that each selected cell of the Hough space only receives one single vote.

3.3 Line extraction

The algorithm looks for winning candidates each time a new beam arrives and a new voting space is generated. 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 180° scan sector and the newest sonar beam cannot provide more information). It is possible to determine the zone in the voting space where the candidates have received all the available votes (see Fig. 4). Basically, we can say that all the candidates beyond the boundary described by all the votes that an upcoming beam could

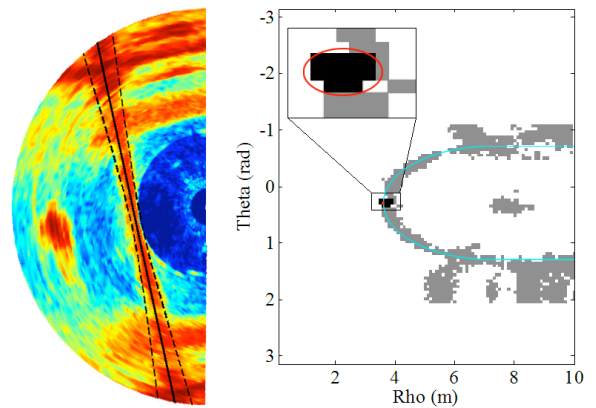


Fig. 5. Uncertainty estimation. The parameters of the winning line are represented as a small red dot in the segmented data. The line trace is represented in cyan. In the detail, all the neighboring compatible lines are represented in black. The red ellipse groups those lines and represents the feature uncertainty. The output line feature is represented in the scan sector.

generate, accomplish this condition. Performing the detection in this way, we can ensure that our algorithm detects the lines as soon as they are completely visible. After a line detection, all the bins involved in the election of the candidate are removed from the buffer so they do not interfere with the detection of further features.

4. UNCERTAINTY ESTIMATION

When a line has been detected among the candidates in the HS the next step is to estimate its uncertainty. As the winning line has obtained a considerable number of votes, it must correspond to a high intensity zone of the buffered beams (see

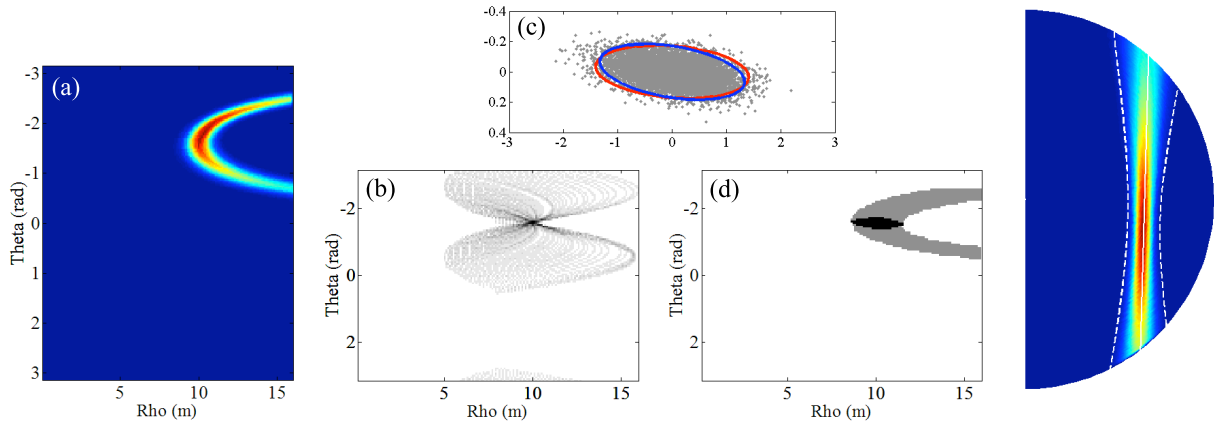


Fig. 6. Testing the algorithm with synthetic data. (a) Raw sensor data generated from ρ and θ given a normally distributed uncertainty. Some correlation affects the two variables to increase the difficulty of the test. (b) The voting space clearly identifies the line. (c) Points generated with the test distribution to produce the synthetic data. The blue ellipse represents the real multivariate normal distribution for a 95% confidence level while the red one is the output of our algorithm. (d) The same estimated uncertainty as obtained by the algorithm. (e) The estimated line feature fits almost perfectly the synthetic one.

Fig. 2b). This zone (an arc) represents the imprint of the object (a line) in the polar representation of the acoustic image and its thickness is related to the line uncertainty. The points belonging to the arc are the polar representation of the sonar measurements corresponding to the uncertain line in Cartesian coordinates. If both, the scanner and the HS are represented in the same frame, the point with the minimum ρ (the maximum of the arc) would have the same rho-theta used to represent the line in the HS. If not, a coordinate system transformation can be used to relate both of them. It is worth noting that although HS is used to represent lines, it is actually a polar representation of the space and hence it can also represent points. Lines are in fact represented using their point with smallest distance to the origin. Therefore, it is possible to represent the buffered beams, for which we know its world coordinates, into the HS (referenced to B) where the detected line is represented (Fig. 5). The points belonging to the detected line describe an arc in the HS, and it is possible to compute their number of compatible measurements. Moreover, each realization of a stochastic line in Cartesian coordinates, accomplishes the following properties: (1) there exists a corresponding arc in the HS falling within the uncertain (thick) arc, (2) it has a number of compatible measurements similar to the already detected line, and (3) it is represented by a rho-theta point belonging to the neighbourhood of the already detected line. Therefore, it is possible to look for the points (representing line realizations) in the neighbourhood of the detected line satisfying the 3 properties. The region described by those points has an elliptical shape in HS and it is possible to estimate its parameters (red ellipse in Fig. 5). This ellipse can then be compared with the

ellipse described by a 2-dimensional multivariate normal distribution for a given confidence level. The mean of this normal distribution is taken as the output line feature, while the probability distribution itself is a representation of the feature uncertainty.

5. RESULTS & CONCLUSIONS

In order to validate the algorithm, several experiments with both real and synthetic data were carried out. Using synthetic data presents some advantages. First, it is possible to compare theoretical Gaussian lines (Fig. 6) with real ones (Figs. 7 and 8) and hence, evaluate how well real measurements are represented by a normally distributed (ρ, θ) variables. On the other hand, as we generate the set of synthetic data given particular values for the probability distribution, we can validate the output of our algorithm. In Fig. 6 the results for a synthetic experiment are presented. We generated a set of sonar measurements by injecting Gaussian noise at the ρ and θ values, but also by adding some correlation to the variables (Fig. 6a). The detection of the line feature is trivial as there are no noisy measurements or phantoms that can disturb the process (Fig. 6b). The uncertainty estimation is also very good as the zone representing que candidate lines (black shape in Fig. 6d) is elliptic, which reflects the Gaussianity of the original data. In Fig. 6c the estimated uncertainty (red ellipse) is compared with the probability distribution used to generate the data (blue ellipse). In general, if we can assume that for a particular segmentation threshold we could assure the existence of a real feature within

the imprint given a certain confidence level, we would be capable to estimate its uncertainty.

Results from real datasets obtained with a Tritech Miniking imaging sonar mounted in the ICTINEU^{AUV} are presented in Figs. 7 and 8. A Sontek Argonaut DVL sensor incorporating a magnetic compass was used to estimate the vehicle movement and to undistort the images. The first image was taken in the channels of a nautical club and, as it can be observed, the aspect of the real features matches the synthetic ones (Fig. 7). The second experiment was performed in the water tank of the Underwater Robotics Research Center at the University of Girona. As this environment is very confined and the material of its walls is highly reflective, the obtained data is much more noisy and some phantoms are present. However, the obtained results were still good (Fig. 8), demonstrating the robustness of the proposal.

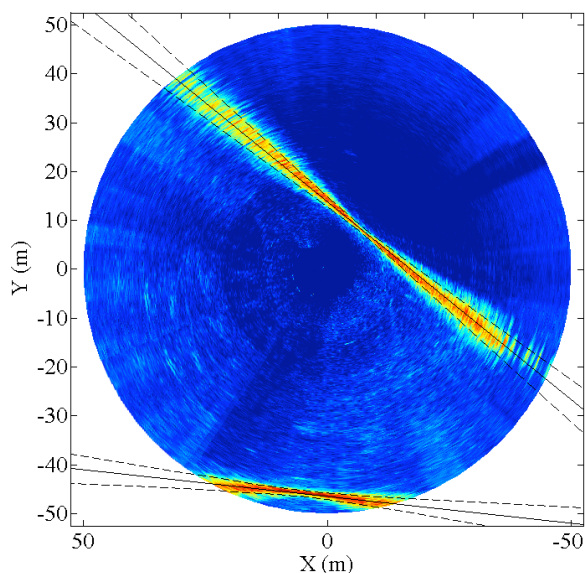


Fig. 7. Results obtained for a scan taken in a channel.

ACKNOWLEDGMENT

This research has been funded in part by the Dirección General de Investigación of Spain under projects DPI2006-13578 and DPI2005-09001-C03-01.

REFERENCES

- Bailey, T. and H. Durrant-Whyte (2006). Simultaneous localization and mapping (SLAM): Part II, state of the art. *Robotics and Automation* **13**(3), 108–117.
- Durrant-Whyte, H. and T. Bailey (2006). Simultaneous localisation and mapping (SLAM): Part I, the essential algorithms. *Robotics and Automation* **13**(2), 99–108.

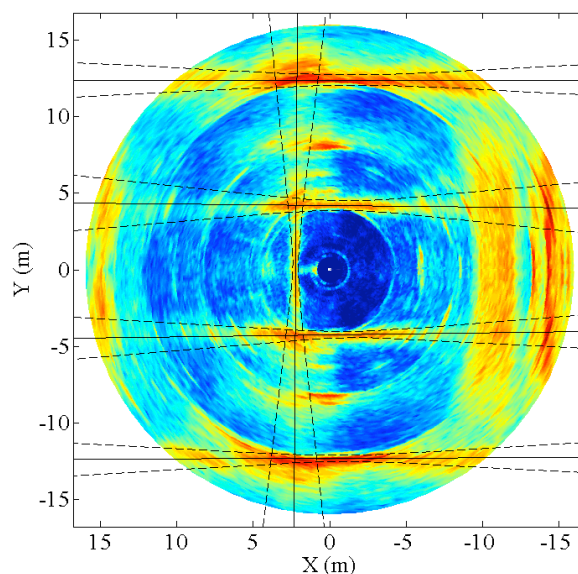


Fig. 8. Results obtained for a scan taken in a confined small water tank. The lines on the right side are not estimated as they are split between the start and the end of the scan.

- Illingworth, J. and J. Kittler (1988). A survey of the hough transform. *Computer Vision, Graphics, and Image Processing, Academic Press Professional* **44**(1), 87–116.
- Leonard, J.J. and H.F. Durrant-Whyte (1992). *Directed Sonar Sensing for Mobile Robot Navigation*. Kluwer Academic Pub.. London.
- Leonard, J.J., R.N. Carpenter and H.J.S. Feder (2001). Stochastic mapping using forward look sonar. *Robotica* **19**, 341.
- Newman, P.M. and J. Leonard (2003). Pure range-only sub-sea SLAM. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. Taipei, Taiwan. pp. 1921–1926.
- Ribas, D., J. Neira, P. Ridao and J.D. Tardós (2006). SLAM using an imaging sonar for partially structured environments. In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*.
- Tardós, J.D., J. Neira, P. Newman and J. Leonard (2002). Robust mapping and localization in indoor environments using sonar data. *International Journal of Robotics Research* **21**(4), 311–330.
- Tena, I., Y. Petillot, D.M. Lane and C. Salson (2001). Feature extraction and data association for AUV concurrent mapping and localisation. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. Seoul, Korea. pp. 2785–2790.
- Williams, S. and I. Mahon (2004). Simultaneous localisation and mapping on the great barrier reef. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. New Orleans, USA.