A Measurement Model for Mobile Robot Localization using Underwater Acoustic Images

Antoni Burguera, Gabriel Oliver and Yolanda González
Dept. Matemàtiques i Informàtica - Universitat de les Illes Balears
Ctra. Valldemossa Km. 7,5 - 07122 Palma de Mallorca (SPAIN)
{antoni.burguera,goliver,y.gonzalez}@uib.es

Abstract

Likelihood fields (LF) have been used in the past to perform localization. These approaches infer the LF from range data. However, an underwater Mechanically Scanned Imaging Sonar (MSIS) does not provide distances to the closest obstacles but echo intensity profiles. In this case, obtaining ranges involves processing the acoustic data.

The proposal in this paper avoids the range extraction to build the LF. Instead of processing the acoustic images to obtain ranges and then using these ranges to infer a LF, this paper proposes the use of the acoustic image itself as a good approximation of the LF. The experimental results show the potential benefits of using this idea to define a measurement model to perform mobile robot localization.

1. Introduction

For a mobile robot to execute useful long-term missions, it has to keep track of its own position. This task is known as *mobile robot localization*. Thrun et al. [12] define mobile robot localization as the problem of determining the pose of a robot relative to a given map of the environment. However, the requirement of an *a priori* map limits the robot's autonomy and the range of scenarios where it can be deployed. Moreover, in a wide range of applications, especially outdoor and underwater, it is not possible to have an *a priori* map.

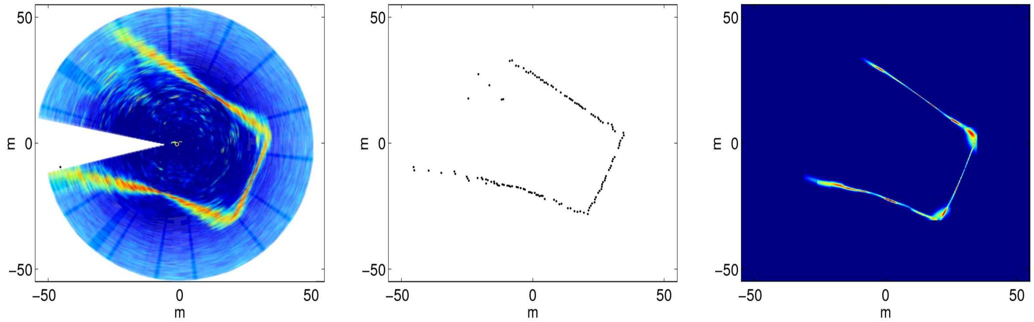
There are several studies in the literature dealing with the problem of determining the robot pose with no *a priori* map. Some strands of research [7, 8, 2] match the recent sensory input against a short history of previously gathered sensor data. These approaches have the advantage of reduced computational requirements, though they suffer from drift due to their incremental nature. Some other studies build a map of the environment simultaneously to the robot pose estimation and use the constructed map to improve the pose estimates [5]. These approaches are referred to as *Simultaneous Localization and Mapping* (SLAM). Although SLAM is known to have problems of computational complexity, different studies exist proposing solutions to this issue [9]. Moreover, SLAM has the important advantage of solving the drift problem when revisiting previously mapped areas.

Both incremental localization and SLAM approaches rely on their ability to match recent sensor data against previously gathered measurements. In some cases, this involves the detection of features, such as straight lines or corners, and then performing the matching at the feature level. Besides, some other studies avoid the use of features, providing a more general approach not depending on the type of environment. Examples of matching techniques not depending on features are those inspired on the *Iterative Closest Point* (ICP) algorithm [7] and those based on the use of *Likelihood Fields* (LF) [11], both designed to work with range measurements. In the context of localization, these matching processes constitute the so called *measurement model*.

A LF is defined as a function of (x, y) coordinates depicting the likelihood of obstacle detection. They are often built from *a priori* maps, though different studies show that local LF can be constructed from local sets of range readings [3]. For example, the *Normal Distributions Transform* (NDT) [1] builds the LF as a grid of Gaussians, each of them being computed from spatially disjunct subsets of range measurements. Also, the *LF with Sum of Gaussians* (LF/SoG) [3] models each reading in a given set as a Gaussian and then defines the local LF as the sum of these Gaussians. In this context, matching two sets of readings involves building a LF from one of the sets and then maximizing a likelihood function that computes how much the second set of readings fits onto the LF given a certain displacement and rotation between both sets.

The sensors providing range information, such as laser and ultrasonic range finders, are very common in terrestrial robotics. However, in underwater environments, which are the scope of this paper, it is more frequent the use of imaging sonars and profilers. Instead of providing ranges to the closest obstacles, these sensors provide acoustic profiles or acoustic images of the environment.

This paper proposes a measurement model based on LF. Our proposal uses the acoustic images provided by an



(a) (b) (c)

Figure 1. (a) Acoustic image. (b) Obtained ranges. (c) Inferred LF.

underwater *Mechanically Scanned Imaging Sonar* (MSIS) as the LF itself. The paper is structured as follows. Firstly, Section 2 motivates the approach. Then, the problem is stated in Section 3. Afterwards, Section 4 describes the acoustic image building process. The measurement model is presented in Section 5. Finally, Section 6 shows some experimental results and Section 7 concludes the paper.

2. Motivation

Figure 1-a shows an example of acoustic data provide by a MSIS. Some studies process these acoustic data in order to compute ranges [6]. Figure 1-b shows an example of ranges extracted from an acoustic image. This approach has proved to be useful when using measurement models based on ICP. However, when using measurement models based on LF, computing ranges may not be the best choice. In this case, computing ranges means discarding some information from the acoustic image data that has to be somehow inferred when building the LF. Figure 1-c shows a LF obtained from range data. It has been constructed using an interpolated NDT. By comparing this image to the original acoustic image, it is clear that valuable information has been lost during the process.

Our proposal is based on this idea. Instead of processing the acoustic images to obtain ranges and then using these ranges to infer a LF, this paper proposes the use of the acoustic image itself as a good approximation of the LF. This idea is supported by the experimental evidence in [10], where it is shown that the different echo intensity levels are related to the uncertainty in the detected object location. That is, uncertain detections spread the echo intensities around the detected object, while good detections provide clear echo intensity peaks nearby the detected object.

3. Problem statement

The experiments conducted in this paper have been performed using the sensor data gathered by the *Ictineu AUV*. This *Autonomous Underwater Vehicle* (AUV) was designed and developed at the University of Girona (see [10] for more details). Among other sensors, the AUV is endowed with a *Doppler Velocity Log* (DVL) which measures the velocities of the unit with respect to bottom and water, a *Motion Reference Unit* (MRU) that provides absolute attitude data by means of compass and inclinometers, and a *Mechanically Scanned Imaging Sonar* (MSIS).

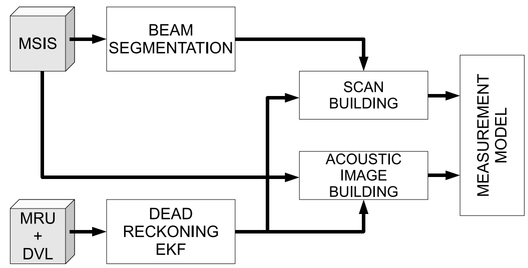


Figure 2. The proposed framework.

The MSIS obtains 360^o scans of the environment by rotating a sonar beam through 200 angular steps in about 13.8 seconds. At each angular position, a set of 500 values, named *bins*, is obtained representing a 50 m long echo intensity profile with a resolution of 10cm. Each of these sets of 500 bins will be referred to as *beam*. By accumulating this information, an *acoustic image* of the environment can be obtained.

Our proposal is summarized in Figure 2. First, DVL and MRU readings are fused by means of an *Extended Kalman Filter* (EKF) to obtain *dead reckoning* estimates. Also, the *beam segmentation* process [4] extracts range information from the MSIS beams. The raw MSIS beams, the obtained ranges and the dead reckoning estimates are stored in three buffers, called *beams history*, *ranges history* and *transformations history* respectively. When the MSIS has performed two 360^o rotations, the *scan building* combines the information in the transformations history and the readings history corresponding the the second MSIS rotation and builds the range scan S_k. Also, the transformations history and beams history obtained during the first MSIS rotation is used by the *acoustic image building* process to construct the acoustic image I_{k-1}. The transformations history data is needed to compensate the robot motion during the MSIS data gathering. Afterwards, the measurement model is in charge of computing the displacement and rotation between the acoustic image

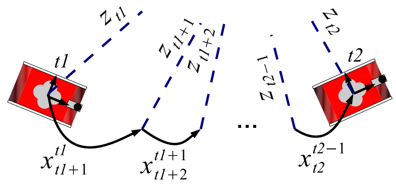


Figure 3. Robot motions and MSIS beams

and the range scan.

There are two points to be emphasized. On the one hand, the raw MSIS beams used to construct S_k can be used later to build I_k and, thus, the measurement model is able to operate every scan rotation. On the other hand, it is straightforward to see that the measurement model does not require that the acoustic image and the range scan have been consecutively gathered. That is, the measurement model can be used to estimate the rotation and displacement between S_i and I_j, for all $1 \leq i \leq k, 1 \leq j \leq k, i \neq j$. However, without loss of generality, the rest of the paper will assume a measurement model operating on S_k and I_{k-1}, to ease notation.

The rest of the paper is devoted to describe the acoustic image building and the measurement model. A description of the beam segmentation and the scan building is available in [4]. The dead reckoning EKF is explained in [10].

4. Building the acoustic image

Let us assume that the data corresponding to I_{k-1} has been gathered between the time step $t1$ and the time step $t2$. The relative robot motions during this time interval are stored in the transformations history and are as follows:

$$M = \{x_{t1+1}^{t1}, x_{t1+2}^{t1+1}, ..., x_{t2}^{t2-1}\} \tag{1}$$

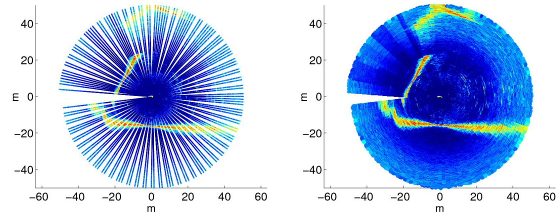
Let z_t denote the MSIS beam obtained at time step t. The set of beams involved in the construction of I_{k-1} is defined as

$$Z = \{z_{t1}, z_{t1+1}, ..., z_{t2}\} \tag{2}$$

This notation is summarized in Figure 3. As a beam is composed of 500 bins, each z_t is a vector containing 500 echo intensity values corresponding to distances up to 50 m with a resolution of 0.1m. Thus $z_t(i)$ contains the echo received for obstacles in an interval ranging from $(i-1)/10$m to $i/10$m. Let $x_t^R = [x, y, \theta]^T$ denote the relative pose a coordinate frame located at the MSIS and aligned with the beam orientation at time t with respect to the robot's coordinate frame. The endpoints of the aforementioned intervals with respect to the robot are as follows:

$$z_t^R(i) = \begin{bmatrix} x + (i/10)\cos\theta \\ y + (i/10)\sin\theta \end{bmatrix} \tag{3}$$

By using the information in M, z_t^R can be represented with respect to the robot pose at time j, $t1 \leq j \leq t2$ while taking into account the robot motion as follows:



(a) (b)

Figure 4. (a) First approximation. (b) Including interpolated values.

$$z_t^j = \begin{cases} z_t^R & t = j \\ x_{j+1}^j \oplus x_{j+2}^{j+1} \oplus ... \oplus x_t^{t-1} \oplus z_t^R & t > j \\ x_{j-1}^j \oplus x_{j-2}^{j-1} \oplus ... \oplus x_t^{t+1} \oplus z_t^R & t < j \end{cases} \tag{4}$$

where $x_{k-1}^k \equiv \ominus x_k^{k-1}$, and $\oplus$ and $\ominus$ denote the compounding and inversion operators, commonly used in stochastic mapping.

By means of this equation, all the end points of the mentioned intervals can be represented with respect to a single coordinate frame while correcting the motion induced distortions. Let us define this coordinate frame at the central position of the trajectory stored in M, although other positions could be used. In this way, a first approximation of the acoustic image can be built by assigning to each end point the echo intensity corresponding to the whole interval. Figure 4-a shows the resulting image. The effects of the range resolution cannot be appreciated due to the figure resolution. However, the effects of the discrete angular sampling of the MSIS are clearly visible. It is important to emphasize that the angular separation between consecutive beams is not identical due to the robot motion.

In order to alleviate the problems of the discrete MSIS sampling, both in angle and range, it is necessary to interpolate the echo intensity values for those points where no samples are available. For example, the echo intensity for each point could be selected as the one corresponding to the closest sample. Figure 4-b shows the result of this simple interpolation. Also, the interpolated value could be the result of combining the closest samples. The acoustic image shown in Figure 1-a performs the interpolation by combining the sample values within an angular window. Let $I_{k-1}(x)$ denote the function providing the interpolated echo intensity corresponding to a given coordinate x. Let us call this function the acoustic image.

5. The Measurement Model

This section proposes a method to compute the displacement and rotation between the coordinate frame of the range scan S_k and the one of the acoustic image I_{k-1}. It is straightforward to derive a method to compute displacement and rotation from I_{k-1} and I_k, avoiding the

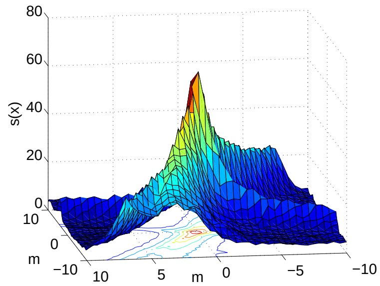


Figure 5. Evaluation of the score function

range extraction. However, this would result in much higher computational requirements.

Let $S_k = \{p_1, p_2, ..., p_N\}$, where the p_i are the points belonging to the range scan represented with respect to the scan coordinate frame. Let the score function $s(x)$ be defined as follows:

$$s(x) = \sum_{i=1}^{N} I_{k-1}(x \oplus p_i) \quad (5)$$

where x represents a displacement and rotation from I_{k-1} to S_k. This function provides higher values for those x that project the scan points onto large echo intensity values. As the scan points have been extracted from large echo intensities, the closer the value of x to the right displacement and rotation, the larger the score value. Accordingly, the motion estimate from the acoustic image to the scan is as follows:

$$x_S^I = \arg\max s(x) \quad (6)$$

Due to the sample based nature of I_{k-1}, maximizing such function requires some numerical optimization method.

6. Experimental Results

This section shows some preliminary, yet promising, experimental results. The tests have been performed building the range scan and the acoustic image from the same MSIS data. In this way, the displacement and rotation between them is perfectly known to be $[0; 0; 0]^T$, constituting the ground truth.

The score function $s(x)$ has been evaluated for different couples of scans and acoustic images, showing clear maximums around the ground truth. Figure shows the results of one of these experiments, showing the values of $s(x)$ for different displacements while keeping the rotation constant. The used acoustic image and range scan are those shown in Figure 1. A clear maximum appears close to the ground truth $[0, 0, 0]^T$.

7. Conclusion

This paper presents a method to match two sets of acoustic profiles provided by a MSIS. The matching method can be used as a measurement model to perform scan matching or scan based SLAM using underwater sonar data. Also, the proposed score function could be directly used to evaluate particles in Monte Carlo Localization. The experimental results, although preliminary, show the potential benefits of the approach.

Acknowledgements

This work is partially supported by DPI 2008-06548-C03-02, FEDER funding and FP7-TRIDENT Project SCP8-GA-2010-248497. The authors are grateful to the members of the Underwater Robotics Lab, in Universitat de Girona, for the data sets.

References

[1] P. Biber and W. Straβer. The normal distribution transform: a new approach to laser scan matching. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 2743 – 2748, October 2003.

[2] A. Burguera, Y. González, and G. Oliver. A probabilistic framework for sonar scan matching localization. *Advanced Robotics*, 22(11):1223–1241, 2008.

[3] A. Burguera, Y. González, and G. Oliver. On the use of likelihood fields to perform sonar scan matching localization. *Springer Autonomous Robots*, 26(4):203–222, May 2009.

[4] A. Burguera, Y. González, and G. Oliver. Underwater scan matching using a mechanical scanned imaging sonar. In *Proceedings of Intelligent Autonomous Vehicles*, 2010.

[5] H. Durrant-Whyte and T. Bailey. Simultaneous localization and mapping (SLAM): part I. *IEEE Robotics and Automation Magazine*, 13(2):99–110, June 2006.

[6] E. Hernández, P. Ridao, D. Ribas, and J. Batlle. MSISpIC: A probabilistic scan matching algorithm using a mechanical scanned imaging sonar. *Journal of Physical Agents*, 3(1):3–11, 2009.

[7] F. Lu and E. Milios. Robot pose estimation in unknown environments by matching 2D range scans. *Journal of Intelligent and Robotic Systems*, 18(3):249–275, March 1997.

[8] J. Minguez, L. Montesano, and F. Lamiraux. Metric-based iterative closest point scan matching for sensor displacement estimation. *IEEE Transactions on Robotics*, 22(5):1047–1054, October 2006.

[9] P. Piniés and J. Tardós. Large scale SLAM building conditionally independent local maps: Application to monocular vision. *IEEE Transactions on Robotics*, 24(5), October 2008.

[10] D. Ribas, P. Ridao, J. Tardós, and J. Neira. Underwater SLAM in man made structured environments. *Journal of Field Robotics*, 11-12(25):898–921, 2008.

[11] S. Thrun. A probabilistic online mapping algorithm for teams of mobile robots. *International Journal of Robotics Research*, 20(5):335–363, 2001.

[12] S. Thrun, W. Burgard, and D. Fox. *Probabilistic Robotics*. The MIT Press, 2005.