

Correlation of Imaging Sonar Acoustic Shadows and Bathymetry for ROV Terrain-Relative Localization

Jose Padial*
jpadial@stanford.edu

Shandor Dektor*
sgd@stanford.edu

Stephen M. Rock*[†]
rock@stanford.edu

*Aerospace Robotics Laboratory, Stanford University
496 Lomita Mall, Rm 250, Stanford, CA 94305, USA

[†]Monterey Bay Aquarium Research Institute (MBARI)
7700 Sandholdt Road, Moss Landing, CA 95039, USA

Abstract—

This paper demonstrates the use of sonar imagery as an automated pilot aid for localization of a remotely-operated vehicle (ROV) with respect to an *a priori* bathymetric terrain map. Specifically, a method is presented to correlate acoustic shadows in sonar imagery with expected visibility images generated from a bathymetry map, where acoustic shadows are significant drops in the sonar image intensity. An expected visibility image is generated for a given ROV position estimate with respect to a stored terrain map by extracting a bathymetry profile along the sonar scan plane and then evaluating a metric termed “differential height” to quantify visibility probability in the pixel space of the sonar image. Results from Monterey Bay Aquarium Research Institute (MBARI) ROV field trials using a Kongsberg mechanically-scanned imaging sonar are presented that demonstrate localization using the proposed method.

I. INTRODUCTION

The ability to localize a remotely-operated vehicle (ROV) accurately with respect to the seafloor terrain enables safer operation and improved navigation to sites of interest. Typical ROV positioning is accomplished through acoustic triangulation of the ROV from the surface vessel, e.g. using an ultra-short baseline (USBL) system. While this method of ROV positioning may be suitable for some mission profiles, it does not provide a terrain-relative position estimate of the ROV, and as such is insufficient for missions requiring accurate knowledge of ROV position relative to the seafloor. A terrain-relative position estimate is particularly important when the terrain map is inaccurately geo-referenced, as a USBL latitude/longitude estimate does not account for geo-referencing error. This motivates the development of a terrain-relative position estimator leveraging sensors commonly found on ROVs. In this work, imaging sonars are used.

Imaging sonars measure the intensity versus range for sonar beams over a span of azimuth angles. At each azimuth angle, the sonar transducer emits a pulse of acoustic energy and listens for return pings, much like sidescan sonar systems. Imaging sonars can be broadly classified into multi-beam and mechanically-scanned types. Multibeam imaging sonars form separate beams with narrow azimuthal and wide elevation beam-widths using a phased array of transducers. Mechanically-scanned imaging sonars physically rotate the sonar transducer in azimuth.

Sonar imagery is currently used by ROV pilots to infer

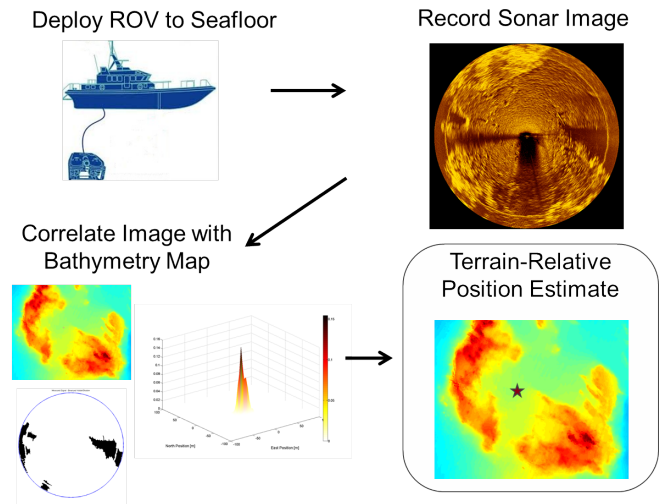


Figure 1: ROV operational motivation for the proposed sonar image-based localization.

map-relative position from features in the sonar image. While ROV pilots are skilled at this task, time and effort would be spared if it were automated. Furthermore, an automated process could explore a broader range of possible ROV position hypotheses to account for larger map geo-referencing errors. Figure 1 provides a schematic depiction of the operational motivation for this work.

Using sonar imagery as a quantitative navigation aid requires a measurement model to generate an expected sensor signal as a function of the vehicle pose and map, which can then be correlated with a measured sensor signal to estimate position. Forming a measurement model for sonar image intensity is difficult, as intensity returns are complex functions of grazing angle, terrain surface composition, and water properties [1]. Further, there are gains and filters applied to the raw signal internal to the sensor which are often, as is the case with this work, unknown to the operator. As such, predicting the intensity signal accurately requires more than a vehicle pose estimate and bathymetry map alone. *Acoustic shadows*, which are significant drops in sonar image intensity, however, are determined primarily by line-of-sight occlusion due to the geometry of ensonified terrain relative to the sonar transducer, and are subsequently well-suited to bathymetry-based correlation techniques.

This paper proposes a method to correlate acoustic shadows in sonar imagery with expected signals generated from a bathymetry map, and presents results from field trials that demonstrate position localization using the proposed approach. Section II briefly describes related prior work, particularly with respect to the use of sonar imagery for navigation. Section III provides technical detail of the approach proposed in this paper, specifically with regard to acoustic shadow labeling in the measured sonar image, expected visibility image generation, and position estimate weighting. Section IV presents results from field trials obtained in collaboration with the Monterey Bay Aquarium Research Institute (MBARI), where the proposed method was used to provide terrain-relative localization estimates for an ROV equipped with a Kongsberg mechanically-scanned imaging sonar.

II. RELATED WORK

Sonar imagery has been successfully utilized previously for underwater localization, however most methods differ from the approach presented in this paper in that they are based on correlating sonar image to sonar image as opposed to sonar image to bathymetry. These image-to-image methods typically correlate using computer-vision image features, and require a flat seafloor assumption. The flat seafloor assumption is employed due to the inherent ambiguity in extrapolating spatial information from sonar imagery; specifically, the azimuth and range of a given intensity return are known for sonar imagery, but the elevation angle is unknown.

In [2] a method is proposed to detect and match landmarks in sidescan imagery by level set evolution on Haralick feature maps, where the nature of the landmark registration is similar in nature to visual feature matching. The work of Fallon et al. [3] matched sidescan image features across multiple sidescan images and fused the spatial matching with acoustic ranging from a surface vessel for navigation drift mitigation. More recently, the work of King et al [4] explored the extension of computer vision feature technology to sidescan image registration, with results showing successful loop closure registration.

In a similar vein, sonar imagery has been used in AUV Simultaneous Localization and Mapping (SLAM) solutions in the context of landmark detection and matching for improved navigation. In [5], distinctive features are identified in sidescan sonar imagery, and matched across images to estimate and track landmark positions. These landmark correspondences are used to improve the navigation estimate in order to project a more smooth and consistent sidescan map.

The work presented in this paper differs from these past works in that the presented method produces a navigation estimate with respect to an *a priori* terrain map, rather than relative to previous sonar images. Further, in this work there is no need for a flat bottom assumption in order to disambiguate the sonar intensity returns; expected visibility images are generated from a bathymetry map and projected into the sonar image range-azimuth domain, *which is an unambiguous mapping*.

This approach was first described by the authors in [6]. Two methods were presented: (1) XOR-based correlation using an expected shadow signal based strictly on line-of-sight occlusion, and (2) correlation using expected visibility images estimated from the relative heights of terrain with respect to

line-of-sight from the sonar transducer, termed “differential height”. The differential height method was shown to yield superior navigation performance, and is the basis for the work presented in this paper.

III. METHOD

The estimation method presented in this paper correlates acoustic shadows in measured sonar imagery with expected visibility images. Acoustic shadows are large drops in return intensity in the sonar image. While sonar intensity returns from low-reflectance terrain surfaces, e.g. featureless mud, can be low, geometric acoustic shadows have been observed to generate significantly lower return intensities. As such, shadow-based correlation is well suited to bathymetry-based localization.

Measured shadows are assigned hard labels according to sonar image intensity. This is accomplished via intensity thresholding, and/or a region-growing shadow labeling method. Further morphological operations may be performed on the binary shadow signal in order to clean it. Section III-A provides the details on shadow labeling.

The expected visibility image is probabilistic, and is generated by calculation of a metric termed “differential height”. For a given position estimate, a bathymetry profile of terrain points for each azimuth scan plane of the sonar image is extracted from the stored terrain map. A differential height value is calculated for each terrain point, which is a measure of how visible/occluded the terrain point is. A visibility probability for each terrain point is calculated from its differential height value and then projected into the range-azimuth space of the sonar image. Finally, each pixel in the expected visibility image is assigned a visibility probability through linear interpolation. The details of the differential height correlation method are presented in Section III-B.

ROV position hypotheses are populated in a non-parametric Point Mass Filter (PMF), and an expected visibility image for each hypothesis is correlated with the measured shadow image in order to yield a terrain-relative position estimate. Details of the filtering approach are presented in Section III-C.

A. Acoustic Shadow Labeling in the Measured Sonar Image

Acoustic shadows in the measured sonar image are assigned a hard label as follows:

$$M[u, v] = \begin{cases} 0 & \text{measured shadow} \\ 1 & \text{otherwise} \end{cases} \quad (1)$$

where M is the measured binary image, and (u, v) is the pixel index in the image.

The primary method of labeling shadows is the application of binary thresholding on sonar image intensity. Image pixels below the threshold intensity values are labeled as shadow.

If the ROV operator is not satisfied by the threshold shadow labels, a region-growing shadow labeling capability has been developed to assist labeling. The recursive region-growing algorithm takes as input a seed shadow location specified by the user, i.e. ROV operator. A mean intensity for the shadow region is initialized to the seed pixel intensity. The algorithm

recursively steps outward to the four neighboring pixels and checks two conditions: (1) if the intensity is within a threshold τ_1 of the seed intensity, and (2) if the intensity is within a threshold τ_2 of the mean shadow region intensity. If a pixel passes these two conditions, the mean intensity for the region is updated, and the algorithm recursively steps to the untested neighbors of this new shadow pixel. Pseudocode for the region-growing algorithm is given below:

Algorithm 1 Region-growing acoustic shadow labeling

Input: Sonar intensity image I , seed pixel (u_s, v_s)

Output: Binary shadow image M

Initialize:

$M[u, v] \leftarrow 1 \quad \forall (u, v) \text{ pixels}$

$\alpha \leftarrow I[u_s, v_s]$ where $\alpha \equiv$ mean shadow intensity

$\beta \leftarrow 1$ where $\beta \equiv$ count of shadow pixels in region

Call:

$M \leftarrow \text{regionGrower}(I, M, u_s, v_s, \alpha, \alpha, \beta)$

function regionGrower($I, M, u, v, \gamma, \alpha, \beta$)

if $|I[u, v] - \gamma| \leq \tau_1 \wedge |I[u, v] - \alpha| \leq \tau_2$ **then**

$\beta \leftarrow \beta + 1$

$\alpha \leftarrow ((\beta - 1)\alpha + I[u, v]) / \beta$

$M[u, v] \leftarrow 0$

for all $(u_N, v_N) \in \text{neighbor set of pixel}(u, v)$ **do**

$M \leftarrow \text{regionGrower}(I, M, u_N, v_N, I[u, v], \alpha, \beta)$

end for

end if

return M

In order to provide the capability to further clean the shadow detection process, morphological image processing is employed. The overall effect of the morphological process is to eliminate speckle noise in shadow regions. The image is first eroded, which consists of sliding a structuring element (disk or square) over the image and performing a logical AND of the pixel values over the structuring element. This has the positive effect of closing shadow holes, with the negative effect of growing shadow region boundaries. The image is then dilated to compensate for shadow boundary growth, which is the same process as erosion but with a logical OR instead of AND. Figure 2 presents a measured shadow image before and after morphological image processing.

B. Expected Visibility Image from Differential Height

An expected visibility image quantifies shadow/visibility confidence of the ensonified bathymetry for a given position estimate. This expected image is generated using the “differential height” metric. The key idea behind differential height is that some expected shadow assignments are more likely than others. For example, an expected acoustic shadow due to grazing should be trusted less than an acoustic shadow expected behind a large mound. As such, this measure accounts for measurement uncertainty due to variations in terrain, and implicitly accounts for uncertainty due to the noise of the sensor model.

The differential height value for terrain point m in the transducer ψ -azimuth scan plane of pose estimate i , $\delta z_{\psi, m}^i$,

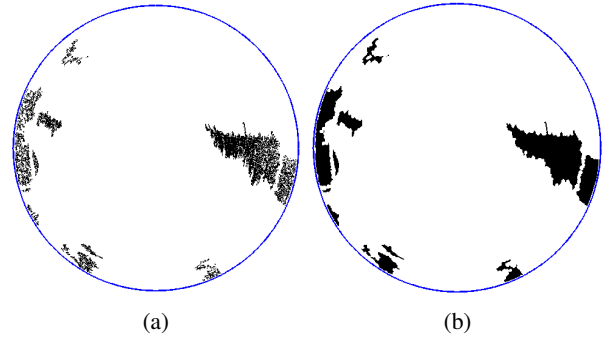


Figure 2: Measured shadow image before and after morphological image processing. (a) Raw shadow image. (b) Shadow image after morphological image processing.

reflects how occluded or visible the terrain point is. Specifically, the differential height for terrain point m is the distance that m is above or below its occluding line-of-sight, where the occluding line-of-sight is from the transducer to the terrain point j that most occludes m in the scan plane. This is illustrated in Figure 3.

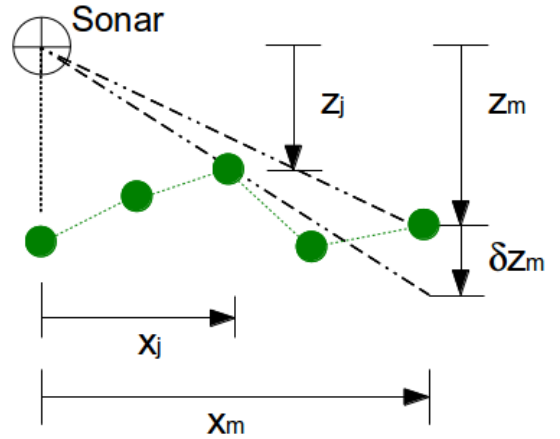


Figure 3: Differential height diagram. Shown is the differential height δz_m of a visible terrain point m above the occluding line-of-sight from terrain point j . Terrain points are shown in green.

If terrain point m is occluded according to line-of-sight from the sonar transducer, $\delta z_{\psi, m}^i$ is negative. Similarly, if the point is visible according to line-of-sight, $\delta z_{\psi, m}^i$ is positive. Equation 2 provides the definition of differential height.

$$\delta z_{\psi, m}^i = \min_j (x_m \frac{z_j}{x_j} - z_m), \forall j = 1, 2, \dots, m-1 \quad (2)$$

Differential heights are translated into probabilities by passing them through a sigmoid function, given by the following relation:

$$p_{\psi, m}^i = 0.5 + \lambda \frac{\delta z_{\psi, m}^i - \mu}{\sqrt{\gamma^2 + (\delta z_{\psi, m}^i - \mu)^2}} \quad (3)$$

where $p_{\psi,m}^i$ is the visibility probability for terrain point m , μ shifts the mean differential height, and γ and λ adjust the sigmoid shape. For the results presented in this work, $\lambda = 0.4$, $\mu = -0.2m$, $\gamma = 0.3m$. These parameters were estimated by fitting the sigmoid function to measured visibility probabilities in sonar imagery. Shifting the mean accounts for the observed behavior that expected shadows are overly predicted by pure line-of-sight when compared to measured shadows. Further, having λ less than the full possible value of 0.5 accounts for measurement uncertainty as a means to prevent filter overconfidence, as detailed in [7]. Figure 4 provides a plot of the sigmoid given by Equation 3.

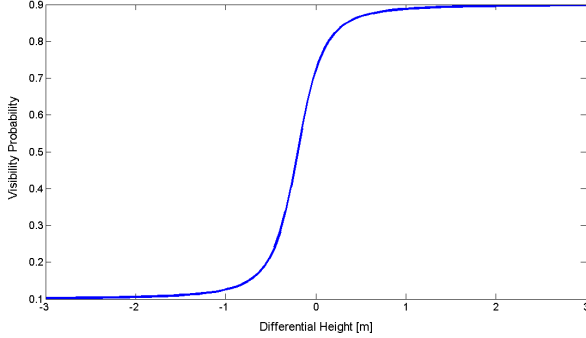


Figure 4: Sigmoid function given by Equation 3 with parameters $\lambda = 0.4$, $\mu = -0.2m$, $\gamma = 0.3m$

Once visibility probabilities have been calculated for terrain points in the scan planes of position estimate i , these values are projected into the range-azimuth space of the sonar image, and linearly interpolated in order to yield the expected visibility image $P[u,v]^i$ for each sonar image pixel (u,v) .

C. Multi-resolution Non-parametric Filtering

This work is intended for use with non-parametric position estimation filters. Examples of such filters include particle filters and point mass filters (PMFs), both of which are extensively used in underwater navigation solutions, and detailed in [8]. This work focuses on the output of a single measurement update for a PMF.

The PMF estimated state is 2-D North and East vehicle position $\mathbf{x} = [x_N, x_E]^T$. The orientation of the vehicle (pitch, roll, heading) is estimated by the ROV inertial navigation system (INS), and altitude provided by either an altimeter or the doppler velocity logger (DVL).

A multi-resolution filtering approach is employed to efficiently estimate ROV position. Initially, a coarse grid of discrete state hypotheses is instantiated about the USBL-estimated ROV position at 4m resolution. Each discrete position hypothesis in the PMF is assigned a normalized measurement weight $w^{(i,j)}$ according to Equations 4,5.

$$w_0^{(i,j)} = \prod_{(u,v) \in C} \underbrace{M[u,v]P[u,v]^{(i,j)}}_{\text{measured visible}} + \underbrace{(1 - M[u,v])(1 - P[u,v]^{(i,j)})}_{\text{measured shadow}} \quad (4)$$

$$w^{(i,j)} = \frac{w_0^{(i,j)}}{\sum_i \sum_j w_0^{(i,j)}} \quad (5)$$

where C is the “correlation region” of the sonar image. The correlation region contains the pixels that will be used in the weighting calculation, and is chosen *a priori* to exclude regions of the image that contain false shadows due to ROV geometry. Additionally, pixels near the radius equivalent to the ROV depth are automatically excluded (using the vehicle depth measurement), as the water surface generates intensity returns not based on the bathymetry (see Dataset 2 in Section IV).

Higher resolution grid cells are instantiated about PMF cells with weight above a threshold. The population of fine-resolution grid cells is done through a 4-neighbor connected scheme, shown pictorially in Figure 5. The red circles are the fine-resolution grid cells and the blue crosses the coarse-resolution cells.

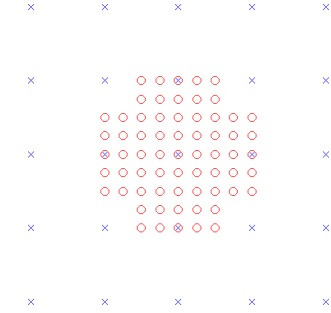


Figure 5: Connected 4-neighbor fine-resolution PMF grid cell population scheme. Blue crosses are coarse-resolution grid cells. Red circles are fine-resolution grid cells

For this work, the refined grid has a spatial resolution of 1m. This value was chosen as a compromise between accuracy and computational burden. Figure 6 presents the trend for number of grid cells to evaluate in the connected 4-neighbor population scheme for one cell of interest as a function of the ratio of fine-grid resolution to coarse-grid resolution.

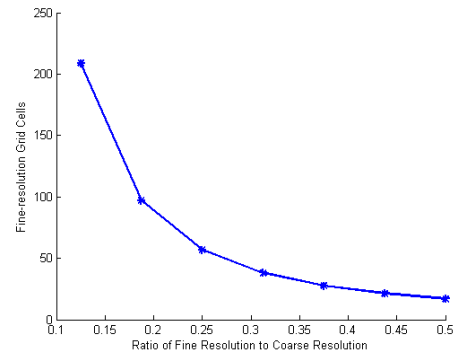


Figure 6: Connected 4-neighbor computation trend for one coarse-grid cell subdivision as a function of the ratio of fine-grid resolution to coarse-grid resolution.

A filter mean and covariance matrix are estimated by Equations 6,7. While the primary data output of the estimation filter is the MLE position, the mean and covariance are helpful for the evaluation of filter confidence.

$$\hat{\mathbf{x}} = \sum_{i=1}^{N_N} \sum_{j=1}^{N_E} \begin{bmatrix} x_N(i) \\ x_E(j) \end{bmatrix} w^{(i,j)} \quad (6)$$

$$\Sigma_x = \sum_{i=1}^{N_N} \sum_{j=1}^{N_E} \left(\begin{bmatrix} x_N(i) \\ x_E(j) \end{bmatrix} - \hat{\mathbf{x}} \right) \left(\begin{bmatrix} x_N(i) \\ x_E(j) \end{bmatrix} - \hat{\mathbf{x}} \right)^T w^{(i,j)} \quad (7)$$

IV. EXPERIMENTAL RESULTS

Results were obtained using field data from Monterey Bay Aquarium Research Institute (MBARI) ROV runs that demonstrate the feasibility of sonar imagery as a quantitative terrain-relative navigation sensor. Data from two ROVs were collected, with each ROV outfitted with a Kongsberg mechanically-scanned imaging sonar located on the vehicle top, as shown in Figure 7. The sonar images obtained have a maximum range of 100m. A 1m-resolution digital elevation map (DEM) of the terrain was used for each dataset.

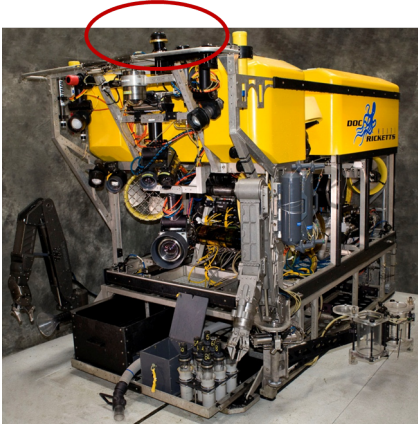


Figure 7: MBARI DocRicketts ROV with a Kongsberg mechanically-scanned imaging sonar circled in red. Image courtesy of mbari.org.

Three ROV datasets were collected in the Monterey Bay for testing the proposed shadow-based correlation method. For each dataset, a scanning sonar image was collected while the ROV remained motionless. Each dataset was tested with a different means of shadow labeling. The shadows in Dataset 1 were manually selected and grown with the region-growing algorithm from Section III-A to simulate the possible use with an ROV pilot interface. The shadows in Dataset 2 were labeled through intensity thresholding with some region-growing augmentation. The shadows in Dataset 3 were labeled solely by intensity thresholding.

The expected visibility images were generated with identical sigmoid parameters for all data sets as specified in Section III-B. The coarse-resolution grid was 96m by 96m with 4m spacing. The coarse-resolution filter weight threshold for evaluation at higher resolution (1m) was set to 0.05.

Table I presents statistics on the three dataset estimators. The estimate mean and covariance estimates were calculated using the fine-resolution grid weights exclusively. This was reasonable for these data, as the cells subdivided to yield the fine-resolution grid cells contained the following percentages

of the coarse-resolution grid position estimate weights for the datasets: (Dataset 1) 99.97%, (Dataset 2) 97.15%, and (Dataset 3) 97.76%. Note the meter-level uncertainties of the filter estimates. Though there is a lack of truth data, the alignment of the MLE expected visibility images with the measured shadow images presented in the remainder of this section strongly indicate the accuracy of the MLE position estimation. As presented in Table I, the MLE positions for the three datasets are offset from the USBL-estimated positions on the order of 30-40m for the three datasets.

Set	\mathbf{x}_{mle}	r_{mle}	\mathbf{x}_{mean}	r_{mean}	σ_1, σ_2
1	4, 35	35.2	4.0, 34.8	35.0	0.5, 0.4
2	2, 39	39.0	1.4, 39.4	39.4	1.5, 0.7
3	11, 26	28.2	10.7, 27.4	29.4	2.6, 0.5

Table I: Dataset estimator statistics. All values are in meters. For each set, vector quantities \mathbf{x}_{mle} , \mathbf{x}_{mean} and scalar ranges r_{mle} , r_{mean} are measured from the USBL-estimated ROV position for that set. Vector quantities \mathbf{x}_{mle} , \mathbf{x}_{mean} are in local East, North coordinates. The 1σ covariance ellipse semi-major and semi-minor lengths are given by σ_1, σ_2 .

Dataset 1 captures the imaging of a large boulder at a deep site (roughly 2900m depth). The measured shadow image was labeled by the region-growing algorithm presented in Section III-A, as there are a small number of clearly identifiable shadows, requiring only a few user clicks to identify the shadows. Figure 8 presents the correlation results. The values given are normalized measurement weights as calculated by Equations 4.5. The strong correlation spike in both the coarse- and fine-resolution evaluations is a solid validation of the proposed method, though this result is in many ways a “control” test, as the position estimation should be greatly aided by the large boulder feature surrounded by relatively flat terrain.

Figure 9 shows the expected visibility image for the Dataset 1 MLE position estimate, the measured shadows, and the alignment image for the measured and expected images. In the expected visibility image, blue indicate stronger confidence in shadow, with red indicating higher confidence in visibility. A key result from Figure 9 is that the MLE expected visibility image is well-aligned qualitatively with the measured sonar image, as shown in the alignment plot. The alignment plot was generated by thresholding the expected visibility image at 0.5 to form a binary expected image; red in the alignment plot indicates visibility agreement between measured and expected images, green indicates shadow agreement, and blue indicates disagreement. The grey hatched region is the non-correlation region, i.e. the complement to the correlation region discussed in Section III-C.

Dataset 2 was collected at shallower depth (86m) in Portuguese Ledge in the Monterey Bay. The measured shadow image was labeled mainly by intensity thresholding augmented by region-growing. Figure 10 presents the correlation results. Note the strong unimodal peaks in both the coarse- and fine-grid results, with the correlation peak well-behaved about the MLE estimate. The uncertainty ellipse is slightly larger than that of the Dataset 1 estimator, which is to be expected due to the ensonified terrain geometry.

Figure 11 shows the expected visibility image for the Dataset 2 MLE position estimate, the measured shadows, and the alignment image. Once again, there is strong agreement between the measured and expected shadow regions, specifically between the large shadow region on the right side of the image, and the smaller rock-created shadow region on the left. Note that the correlation region was specified differently for Dataset 2 as compared to Dataset 1. This is due to two factors: (1) a different ROV was used, so the false shadows caused by ROV geometry differ, and (2) the water surface creates false intensity returns and noise due to the shallow depth of the ROV in Dataset 2. The outer ring of non-correlation region pixels is due to the water surface returns.

Dataset 3 was collected in an area with ripples in the terrain. The measured shadow image was labeled by intensity thresholding alone. Figure 12 presents the correlation results. In the coarse-grid PMF estimation there is a unimodal peak as in Datasets 1 and 2. However, in the fine-resolution estimation there are two distinct peaks. Multimodality in navigation filters is not uncommon, and given the nonlinear measurement model should be expected in some cases, dependent on the terrain/shadow information content. The multimodality accounts for the larger 1σ confidence ellipse for Dataset 3 as opposed to Datasets 1 and 2. Envisioning a matured system, the expected visibility images for both peaks could be supplied to the end user (e.g. ROV pilot) in order to best inform him or her of the potential ROV positions. That the MLE peak has roughly twice the estimate weight as the secondary peak would additionally be useful to the end user.

Figure 13 shows the expected visibility image for the Dataset 3 MLE position estimate, measured shadows, and the alignment image. There is strong agreement in the large shadow region at the top of the image. This shadow is created by a geometric occlusion from the (slight) drop-off of a shelf. There is some disagreement in the smaller ripple shadows. This disagreement could stem from a number of factors, including changes in the terrain from when the bathymetry map was created (sand ripples can change). The key result from Figure 13 is that the MLE visibility image is well-aligned qualitatively with the measured sonar image, particularly in the large geometric shadow at the top. A strength of quantifying shadow confidence through the differential height method is that the minor shadow disagreements are weighted less than more significant shadow region agreements/disagreements, which are more stable to terrain variation and map inaccuracies.

The calculation of an MLE position for each dataset allows observation of the mapped terrain features that led to the sonar image features (shadows). Figure 14 shows terrain maps and the MLE expected visibility image for each dataset. The Dataset 1 boulder is clearly a large feature, but the large shadows in the Datasets 2 and 3 MLE visibility images are caused by relatively slight shelf-like terrain features and smaller rock/ripple features. This highlights the fact that at low ROV altitudes, even relatively minor terrain features can lead to prominent sonar shadows.

V. CONCLUSION

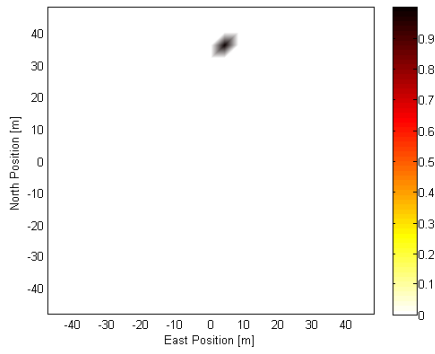
This paper described a method to correlate acoustic shadows in sonar imagery with expected visibility images generated from a bathymetry map for localization of a remotely-operated vehicle. The proposed approach could serve as a navigation augmentation to current ROV ship-based USBL acoustic positioning, and could specifically provide a terrain-relative estimate of ROV position that is inherently not afforded by USBL. Field trials obtained in collaboration with the Monterey Bay Aquarium Research Institute (MBARI) were presented that demonstrate localization using the proposed method.

ACKNOWLEDGMENT

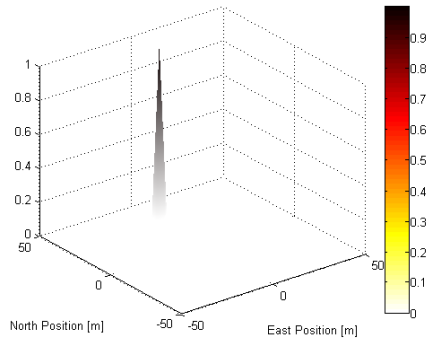
The authors thank the Monterey Bay Aquarium Research Institute (MBARI) for all data and the opportunity to test results at sea. Jose Padial is supported by NASA National Space Technology Research Fellowship NNX11AM92H. Shandor Dektor has received support from a National Defense Science and Engineering Graduate Fellowship and a Stanford Graduate Fellowship.

REFERENCES

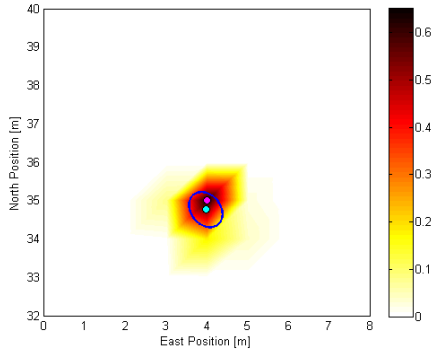
- [1] J. Bell, *A Model for the Simulation of Sidescan Sonar*. PhD thesis, Heriot-Watt University, 1995.
- [2] S. Stalder, H. Bleuler, and T. Ura, "Terrain-based navigation for underwater vehicles using side scan sonar images," in *Oceans 2008*, vol. 1-4, pp. 51–53, 2008.
- [3] M. F. Fallon, M. Kaess, H. Johannsson, and J. J. Leonard, "Efficient auv navigation fusing acoustic ranging and side-scan sonar," in *ICRA*, pp. 2398–2405, 2011.
- [4] P. King, B. Anstey, and A. Vardy, "Comparison of feature detection techniques for auv navigation along a trained route," in *MTS/IEEE OCEANS 2013*, pp. 1–8, 2013.
- [5] I. Ruiz, Y. Petillot, and D. Lane, "Improved AUV navigation using side-scan sonar," in *OCEANS 2003 MTS/IEEE: CELEBRATING THE PAST...TEAMING TOWARD THE FUTURE*, pp. 1261–1268, 2003.
- [6] J. Padial, S. G. Dektor, and S. M. Rock, "Correlation of sidescan sonar acoustic shadows and bathymetry for terrain-relative navigation," in *Unmanned Untethered Submersible Technology 2013*, (Portsmouth, NH), September 2013.
- [7] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. Intelligent Robotics and Autonomous Agents, MIT Press, 2005.
- [8] K. Anonsen and O. Hallingstad, "Terrain aided underwater navigation using point mass and particle filters," in *Position, Location, And Navigation Symposium, 2006 IEEE/ION*, pp. 1027–1035, 2006.



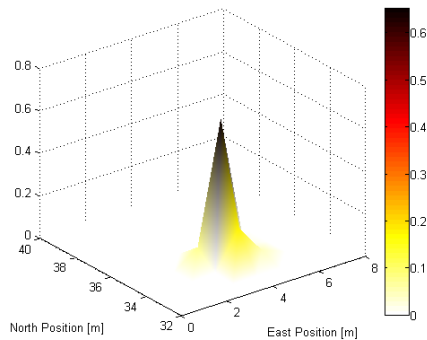
(a) Coarse-grid Top View



(b) Coarse-grid Perspective View

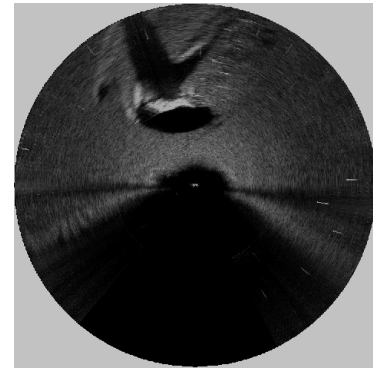


(c) Fine-grid Top View

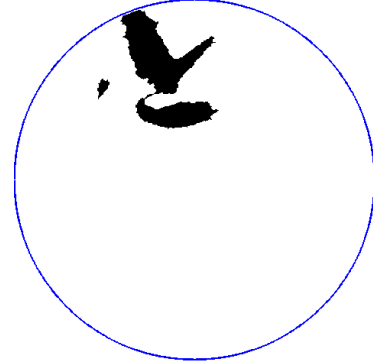


(d) Fine-grid Perspective View

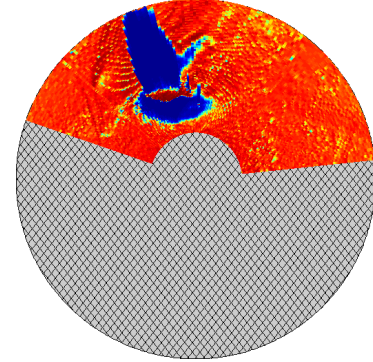
Figure 8: Coarse- and fine-grid PMF correlation results for Dataset 1 (boulder). Correlation results are normalized measurement weights for each position hypothesis. In (c) the magenta point is the MLE position, the cyan point is the mean position, and the blue ellipse is the 1σ confidence ellipse.



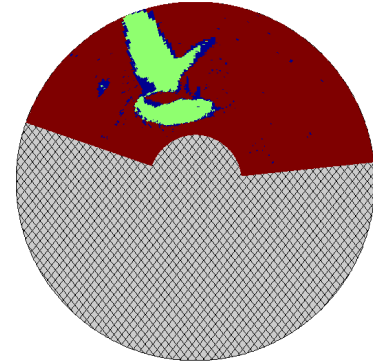
(a) Measured Sonar Image



(b) Measured Shadow Binary Image

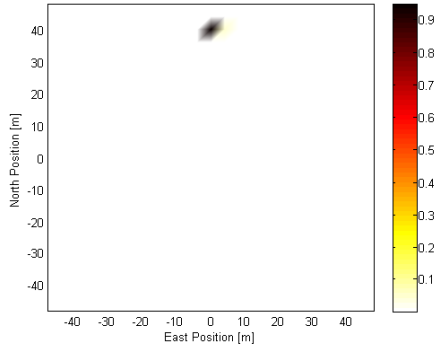


(c) MLE Visibility Probability Image

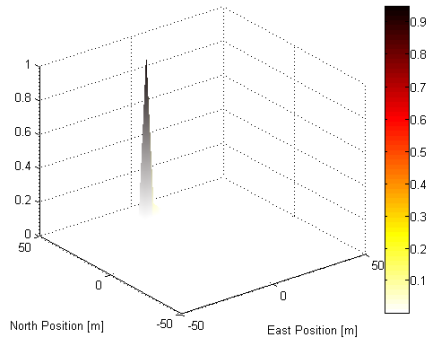


(d) MLE Alignment of Expected and Measured Images

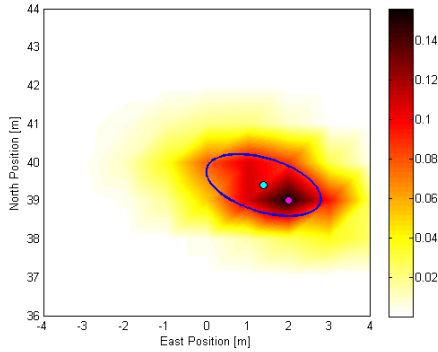
Figure 9: Dataset 1 measured sonar image, measured shadow image, PMF MLE expected visibility probability image, and alignment image. Note that in the MLE visibility image, blue indicates higher confidence in shadow, red indicates higher confidence in visibility. In the alignment plot, red indicates agreement between measured and expected visibility, green indicates agreement between measured and expected shadows, and blue indicates disagreement.



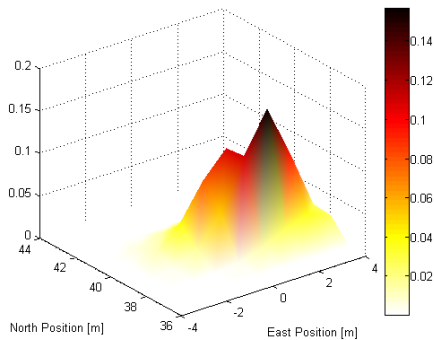
(a) Coarse-grid Top View



(b) Coarse-grid Perspective View

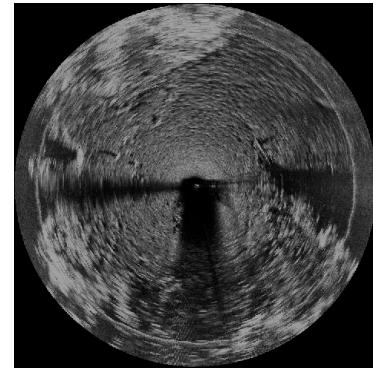


(c) Fine-grid Top View

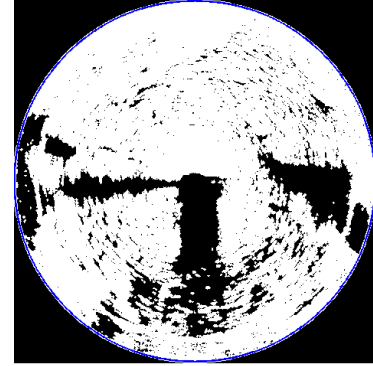


(d) Fine-grid Perspective View

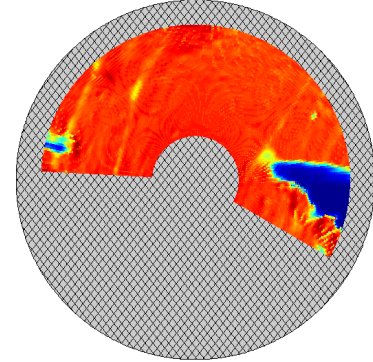
Figure 10: Coarse- and fine-grid PMF correlation results for Dataset 2 (Portuguese Ledge). Correlation results are normalized measurement weights for each position hypothesis. In (c) the magenta point is the MLE position, the cyan point is the mean position, and the blue ellipse is the 1σ confidence ellipse.



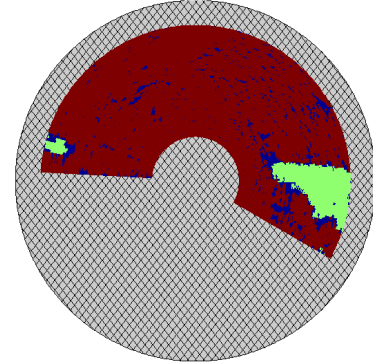
(a) Measured Sonar Image



(b) Measured Shadow Binary Image

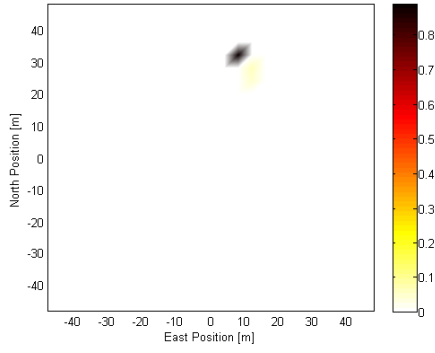


(c) MLE Visibility Probability Image

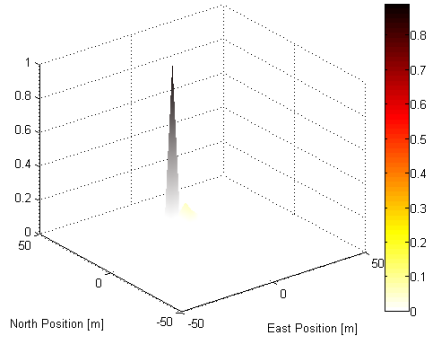


(d) MLE Alignment of Expected and Measured Images

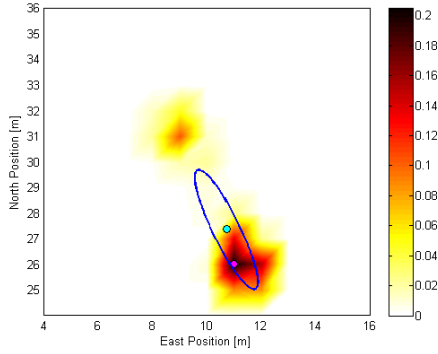
Figure 11: Dataset 2 measured sonar image, measured shadow image, PMF MLE expected visibility probability image, and alignment image. Note that in the MLE visibility image, blue indicates higher confidence in shadow, red indicates higher confidence in visibility. In the alignment plot, red indicates agreement between measured and expected visibility, green indicates agreement between measured and expected shadows, and blue indicates disagreement.



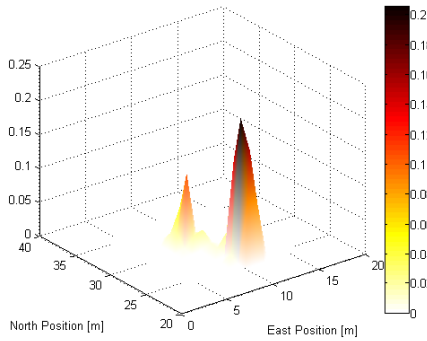
(a) Coarse-grid Top View



(b) Coarse-grid Perspective View

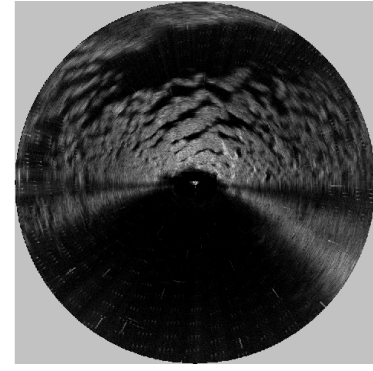


(c) Fine-grid Top View

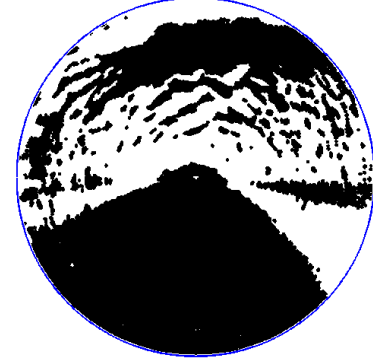


(d) Fine-grid Perspective View

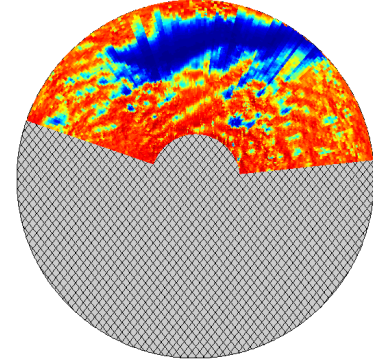
Figure 12: Coarse- and fine-grid PMF correlation results for Dataset 3 (ripple field). Correlation results are normalized measurement weights for each position hypothesis. In (c) the magenta point is the MLE position, the cyan point is the mean position, and the blue ellipse is the 1σ confidence ellipse.



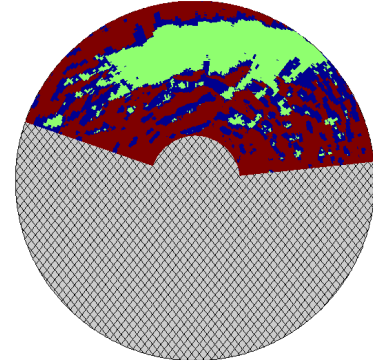
(a) Measured Sonar Image



(b) Measured Shadow Binary Image

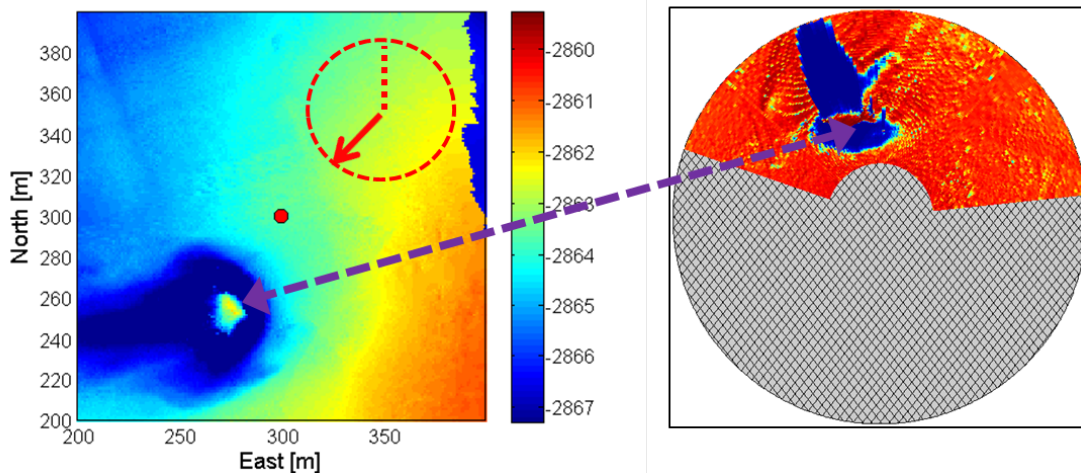


(c) MLE Visibility Probability Image

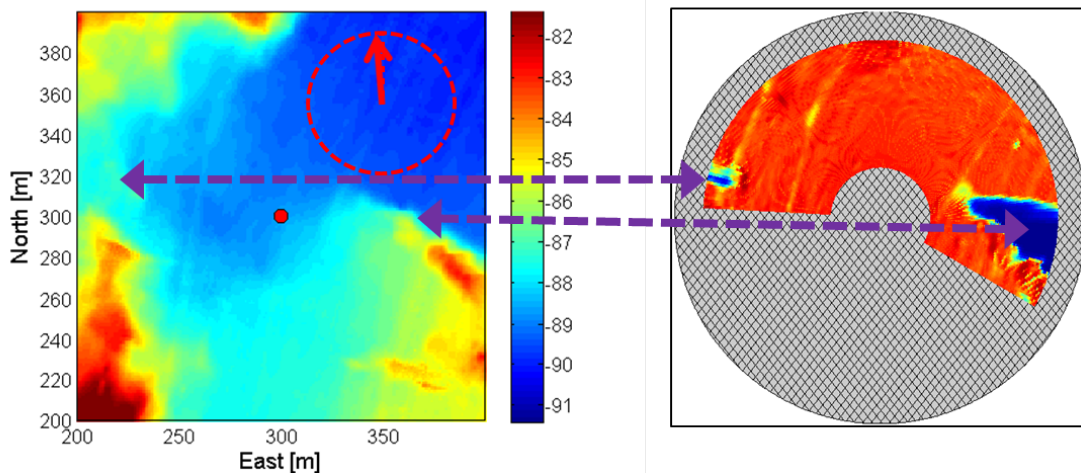


(d) MLE Alignment of Expected and Measured Images

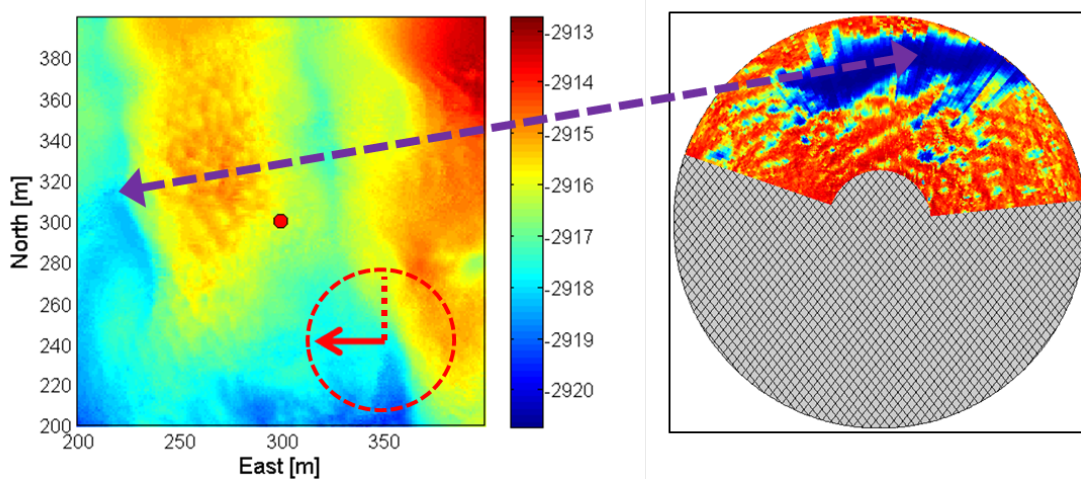
Figure 13: Dataset 3 measured sonar image, measured shadow image, PMF MLE expected visibility probability image, and alignment image. Note that in the MLE visibility image, blue indicates higher confidence in shadow, red indicates higher confidence in visibility. In the alignment plot, red indicates agreement between measured and expected visibility, green indicates agreement between measured and expected shadows, and blue indicates disagreement.



(a) Dataset 1



(b) Dataset 2



(c) Dataset 3

Figure 14: Identification of map feature correspondence to expected sonar shadow features based on estimated MLE positions for Datasets 1, 2, and 3. The red dot on each map is the MLE position. The red arrow on each map is the vehicle heading estimate obtained from the ROV INS. The purple arrow connects prominent features on the map to prominent features in the MLE expected visibility image.