Forward-Look 2-D Sonar Image Formation and 3-D Reconstruction

M. D. Aykin, S. Negahdaripour
ECE Department
University of Miami
Coral Gables, FL 33146
m.aykin@umiami.edu, nshahriar@miami.edu

***Abstract*—Sonar imagery deals with transmitting acoustic signals and measuring the reflected sound from the scene surfaces. The recorded signal encodes information about the shape and material properties of these surfaces. The relationship of the scene geometry to the image irradiance can be established by modeling the physics of the sonar image formation process. The inherent ambiguities in the interpretation of the 3-D world based on visual cues in a 2-D forward-scan sonar image arise as a result of both operating as a ranging device and loss of elevation angle information due to the projection geometry. One aspect of this work is modeling the image formed by a new class of high-resolution 2-D forward-looking sonar systems which supports our complementary aim of recovering the unknown zenith angles from image brightness and thus reconstruct 3-D objects. Our method applies to a single forward-scan sonar image, assuming that the scene objects have smooth surfaces that vary monotonically in terms of distance from the sonar, and cast visible shadows on a flat background. We present the results of experiments with real data to demonstrate the performance of our 3-D reconstruction technique.**

I. Introduction

A new class of high-frequency (MHz) forward-scan (FS) sonar systems is becoming more and more popular and widely used in the operation of autonomous underwater vehicles (AUV) for improving navigation accuracy [23], [32], detecting and identifying scene objects [42], [43] building large-area maps from seamless mosaics [24], [26], [30], [31], [35], [40], tracking/reacquisition of targets [12], [16], and enabling automated ship hull inspection and change detection [37]–[39].

In this work we derive an FS sonar image model for objects with diffuse reflectance that follow the Lambert's law. However, it can be readily revised to describe the sonar images of objects with different reflectance characteristics. As a second objective, we investigate the application of this model for reconstructing 3-D objects within benthic environments, while exploiting the visual cues from cast shadows on the sea floor.

The Lambertian reflection model has been commonly used in optics to describe the image shading of objects with diffuse reflectance [17], [21], [28]. The validity of Lambert's law depends on the surface roughness characteristics, relative to the wavelength of the incident/reflected signal. It has been applied in describing the reflectance of the sea floor, composed of a variety of material types, e.g., soft mud, gravel and rocky bottoms [9], [34], [36]. These assumptions hold even more strongly when the acoustic device operates at high frequencies, including side-scan sonars [2] as well as a new generation of 2-D FS sonar systems operating at 1 MHz and higher frequencies [1], [3], [4]. For side-scan sonars, measured sound power of the reflected signal is obtained as proportional to the n'th power of the cosine of the angle between the incidence sound direction and surface normal where measurements on various bottom characteristics validate that n takes values in between one and two [9], [34].

In optical SFS edge boundaries of objects from stereo (or multiple image) matchings have been used to obtain accurate 3-D reconstructions [8], [10], [14], [15], [22]. In this work we utilize cast shadows to encode information about the object silhouette. The size of a shadow depends on the object position relative to the source, the background surface and its background orientation [27], [41]. Large cast shadows, behind tall 3-D objects resting on the sea-floor, are induced where the sonar is deployed at sharp grazing incidence to cover larger regions within the field of view. These shadows provide important visual cues in determining the size and edge boundaries of objects in forward-scan (FS) sonar imagery. Some earlier works describe methods for estimating the height of a 3-D object, using the knowledge of sonar geometry relative to the sea-floor, where the shadow is cast [6], [20].

For optical images recorded by a stationary viewer under stationary lighting, a homology describes the relationship between a moving planar object and its shadow, cast on a planar background [41]. On the other hand, an FS sonar serves both as the transmitter (source) and receiver (viewer), and is not stationary in most practical applications. Ignoring smaller region corresponding to the images of moving fish, an FS sonar image is typically composed of 3 major region types, namely, 3-D objects, shadows, and the background sea-floor on which these shadows are cast [5]. The proposed 3-D reconstruction method makes use of the association between neighboring bright object and corresponding dark shadow regions. Here, in establishing the object-shadow correspondences, we exploit strong constraints arising from the fact that the source and viewer are coincident. This task is typically much more complicated for optical images, recorded under arbitrary and unknown source position relative to the camera.

In particular, we can often determine the 3-D positions of object contour points based on the 3-D positions of their

978-0-933957-40-4 ©2013 MTS

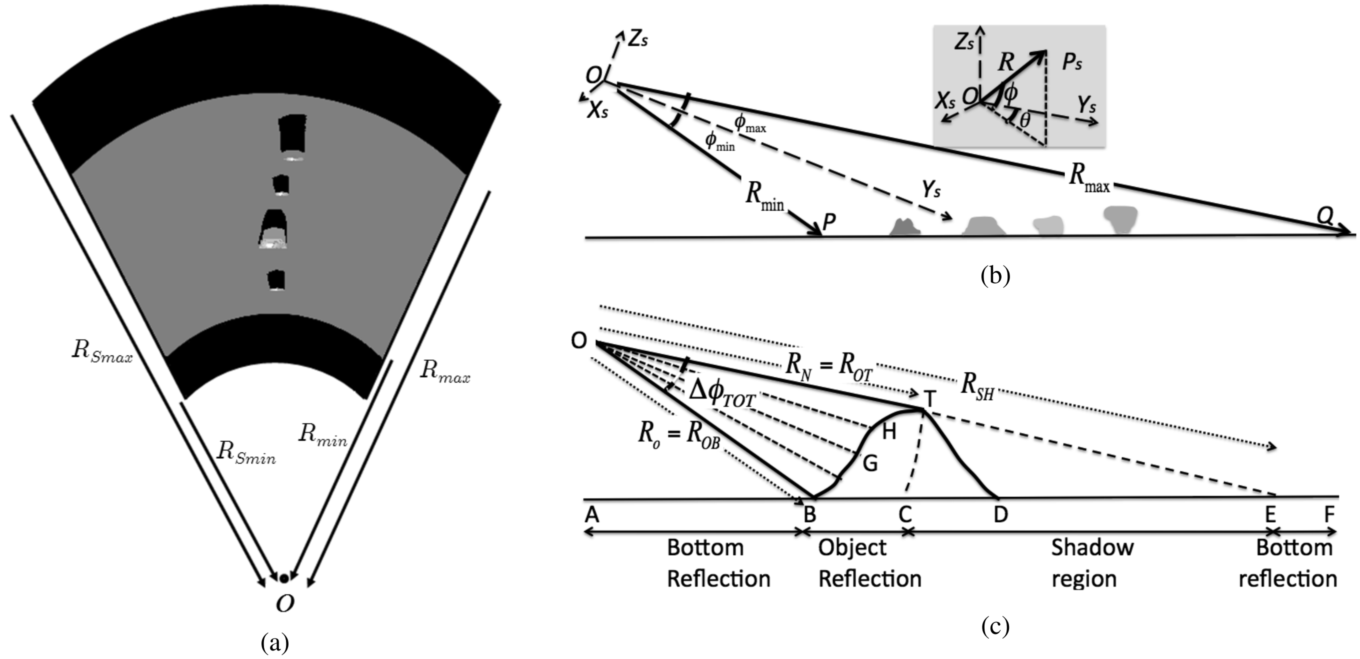

Fig. 1. (a) Sonar range window $\{R_{Smin}, R_{Smax}\}$ can be set by user to achieve optimal scene coverage. The actual minimum and maximum ranges $\{R_{min}, R_{max}\}$ of imaged scene depend on the zenith angle bounds $\{\phi_{min}, \phi_{max}\}$, sonar grazing angle α, distance and orientation of sea bottom encoded in the normal vector $\mathbf{n}$. (b) Field of view in the zenith direction. Directed sound pulse propagates along narrow arc bounded by $\{\phi_{min}, \phi_{max}\}$. (c) Three distinct region types in FS sonar images of a benthic world comprise 1) highlights formed by strong object reflectance (BC); 2) dark blobs corresponding to cast shadows (CE); 3) Intermediate gray levels due to sea-floor reflectance (AB and EF).

shadows on a flat sea-floor [5], [29]. Once the 3-D outline is determined we make the assumption that the distances to object surface points vary monotonically, so on a scan line no two points of the scene surfaces are at the same distance from the sonar. Finally, starting on the boundary we apply our sonar shading model to determine the 3-D shape towards the interior (i.e. across the object).

We present the results for real pool data acquired with a DIDSON video camera, using cylindrical and spherical objects. The concave cylinder is directly reconstructed with the measured intensity. For the convex cylinder and sphere measured intensity is pre-processed before reconstruction since they do not obey our assumption of monotonically varying range and display a strong secondary spike along their intensity profiles due to the ground echo. Finally, the reconstruction of 3 distinct fish from the same lake habitat demonstrate the applicability of our approach for free swimming objects within the water column. We discuss some of the key issues in reconstruction accuracy, including two simultaneous reflectors, ground echo, and the validity of our assumptions.

II. Technical Background

A. Notation

The 3-D coordinate of a point $\mathbf{P}$ in sonar coordinate system is denoted $\mathbf{P}_s = (X_s, Y_s, Z_s)^T$. Utilizing spherical coordinates $(R, \theta, \phi)^T$, as illustrated in Fig. 1(b), the transformations between the Cartesian and the spherical coordinates are given by

$$\mathbf{P}_s = \begin{bmatrix} X_s \\ Y_s \\ Z_s \end{bmatrix} = R \begin{bmatrix} cos\phi \; sin\theta \\ cos\phi \; cos\theta \\ sin\phi \end{bmatrix} \tag{1}$$

and

$$\begin{bmatrix} R \\ \theta \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{X_s^2 + Y_s^2 + Z_s^2} \\ tan^{-1}(X_s/Y_s) \\ tan^{-1}(Z_s/(X_s^2 + Y_s^2)) \end{bmatrix} \tag{2}$$

where the azimuth and elevation angles are measured with respect to y-axis.

B. FS Sonar Imaging

Referring to Fig. 1, sound pressure waves induced by transmitted pings of an FS sonar propagate through the medium, and are reflected by the scene surfaces. The sonar records these reflected echoes. A high-frequency multi-beam FS sonar transmits beams of fixed vertical width bounded by $\{\phi_{min}, \phi_{max}\}$ along various azimuth directions θ. The ping emission and echo recording are scheduled optimally in order to image objects and the scene within a desired range window $\{R_{Smin}, R_{Smax}\}$ at a relatively high data rate. For each beam at azimuth direction θ, the range values are obtained by time-of-flight principle by recording the returning signal as a function of time t:

$$R = \frac{c \; t}{2} \tag{3}$$

where c is the sound velocity in the medium. Based on the range setting $[R_{Smin}, R_{Smax}]$, only the portion of the signal received within the time interval $[2\frac{R_{Smin}}{c}, 2\frac{R_{Smax}}{c}]$

is recorded. This signal is divided into equal discrete time intervals of

$$\Delta t = 2\frac{R_{Smax} - R_{Smin}}{c\ N_t} = t_{i+1} - t_i, \quad i = 0, 1, \ \ldots, N_t - 1 \tag{4}$$

where N_t, the desired number of intervals (i.e. $N_t = 512$ for DIDSON 1.8 [MHz] [4]) fixes the range resolution, and i represents the bin number. Respective fixed ranges R_i are related to t_i's through (3).

The measured average power in each interval determines the intensity value of a pixel in the constructed image. Through the sonar beam forming process, a so-called bin-beam 2-D acoustic image is produced (corresponding to intensity $\mathrm{I}(R, \theta)$ as a function of range and azimuth angles). Further processing by interpolation enables the construction of a polar xy image $\mathrm{I}(x, y)$, where $[x, y] = R[\sin\theta, \cos\theta]$.

III. FS Sonar Image Formation

A. Modeling Intensity Under Lambertian Assumption

The term "intensity" in acoustic literature refers to power per unit area [11], [33], [36]. The sound pressure amplitude decreases by the traveled distance R, giving rise to spherical spreading loss by R^2 for intensity:

$$I_I = k\ \frac{I_0}{R^2} \tag{5}$$

where I_I is the pressure intensity at the surface located at distance R from the source, I_0 represents the reference pressure intensity at 1 m from the source, and constant k accounts for unit conversion.

The scene/target surface characteristics dictate the relationship between the incident acoustic power and the portion scattered in various directions. Underwater objects, i.e. sea-bottom, rocks, reefs, usually have rough surfaces, relative to sonar wavelength. Thus their reflection characteristics may be modeled by diffuse reflection that obey the Lambertian model [9], [28], [34]. Additionally, the FS sonar is arranged at grazing incidence to maximize the overall area being viewed. Thus, even for objects with smooth surfaces, the direction of mirror-like (specular) reflection is generally far from the direction towards the sonar. The exceptions include small isolated local object patches where the normal surface is (nearly) coincident with the propagation direction of the sound wave.

As explained in Section II, the FS sonar acts both as the transmitter and receiver, and thus the angles of incident (transmitted) and reflected beams measured by the sonar are equal. Referring to Fig. 2 and following [36], the diffuse bottom backscatter for surface area dA satisfying the Lambert's reflection law is given by

$$I_S = \mu\ I_I\ cos^2\alpha\ dA \tag{6}$$

where I_I, I_S are respectively the incident and reflected sound intensities, the constant μ embodies unit conversion and the reflectivity of the surface, and α is the angle between the surface normal and incoming sound direction.

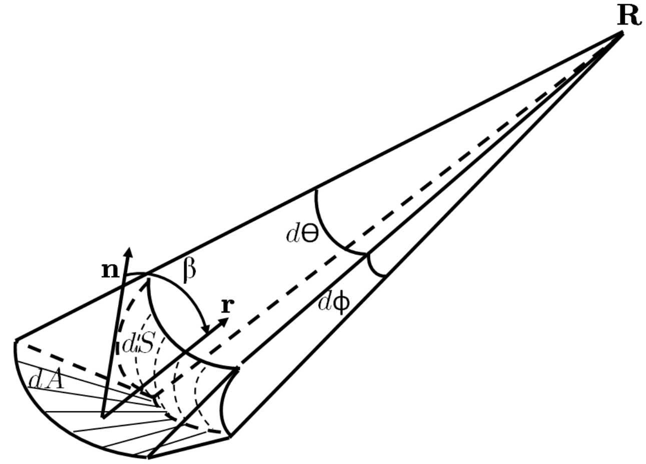

Fig. 2. Infinitesimal areas dA and dS, representing patches on the target surface and solid angle on the sphere centered at the receiver (sonar) in the direction of sound propagation. **n** and **r** respectively represent the local surface normal and propagation direction of the emitted sound pulse towards the patch.

Accounting for the spherical loss of the reflected wave according to (5), the measured intensity at the receiver is

$$I_M = k\ \frac{I_S}{R^2} \tag{7}$$

$$= K\ \frac{I_0}{R^4}\ cos^2\alpha\ dA \tag{8}$$

where $K = \mu\ k^2$ accounts for the constant terms.

The average intensity for a surface patch, with infinitesimal ΔA area as seen from the sonar, is given by

$$I^{avg} = \iint_{\Delta A} K\ \frac{I_0}{R^4}\ cos^2\alpha\ dA \tag{9}$$

Excluding multi-path returns, ΔA as seen from the sonar represents the area within distances R to $R+\Delta R$ and azimuth angles θ to $\theta + \Delta\theta$, where, ΔR and $\Delta\theta$ denote infinitesimal changes in range and azimuth angles.) As depicted in Fig. 2, the true surface patch dS corresponding to the apparent area $dA = dS/\cos\alpha$ is given by $dS = R^2\ cos\phi\ d\theta\ d\phi$. Substitution into (9) yields

$$I_{ij}^{avg} = K\ I_0 \iint_{\Delta A} \frac{1}{R^2}\ cos\alpha\ cos\phi\ d\theta d\phi \tag{10}$$

as the model for the formation of an FS sonar image, which encodes information about the geometry of local patches on the surface of a scene with diffuse reflection. We can utilize (10) to simulate the intensity profile of surfaces with known shapes. Note that the interference effects for multiple scatterers placed at the same range is not accounted by (10). Also, the reconstruction of the scene in 3-D from a given 2-D image is usually under-determined since ΔA can generally be a combination of multiple separate patches (i.e. spanned in between $\{R, R + \Delta R\}$ and $\{\theta, \theta + \Delta\theta\}$).

Under a simplifying assumption that each pixel is formed by a unique planar patch spanning θ_j to $\theta_j + \Delta\theta_j$ and ϕ_{ij} to

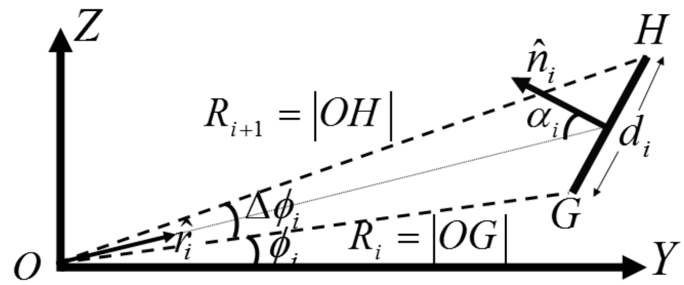

Fig. 3. Modeling object surface as piecewise planar patches, a sample patch and average propagation direction of sound pulse emitted towards the patch are respectively represented by $\hat{\mathbf{n}}_i$ and $\hat{\mathbf{r}}_i$. Triplet $\{\Delta\phi_i, R_i, R_{i+1}\}$ defines a unique triangle $\triangle^{GOH}$, from which orientation and size of i^{th} patch can be determined.

$\phi_{ij} + \Delta\phi_{ij}$ as in Fig. 3, equation (10) reduces to

$$I_{ij}^{avg} = K\ I_0 \int_{\phi=\phi_{ij}}^{\phi_{ij}+\Delta\phi_{ij}} \int_{\theta=\theta_j}^{\theta_j+\Delta\theta_j} \frac{1}{R^2}\ cos\alpha\ cos\phi\ d\theta d\phi \tag{11}$$

where (i, j) are the bin and beam numbers in the corresponding bin-beam image $I(R, \theta)$. This simplification leads to a practical method for 3-D estimation of the scene structure from a single FS sonar image as discussed in Section IV. More precisely, since the range R and azimuth angle θ of each patch is measured by the sonar, the image formation problem and subsequently the 3-D surface reconstruction can be posed in terms of the unknown zenith angles ϕ_{ij}. Therefore, surfaces with monotonically varying ranges (in sonar coordinates) can be reconstructed by the use of equation (11) together with boundary constraints (i.e. edges) as explained in Section IV.

IV. Surface Reconstruction

In this work 3-D reconstruction with FS sonars is performed by the use of (11), which is obtained through the assumption that intensity of a pixel is a function of a single 3-D patch. This assumption is satisfied by all the surfaces where range from the sonar is a monotonic function of the azimuth and elevation. In other words let

$$R = f(\theta, \phi). \tag{12}$$

and $f()$ be monotone. Then intensity of a pixel is determined by only a single 3-D patch and reconstruction is performed as explained in Sections IV-A and IV-B. Also for a group of objects that have two simultaneous reflectors along their profiles, a pre-processing method is discussed in Section IV-C.

A. Constraints for 3-D Contour

Following the methodology described by [5], initially the scene is assumed to be dominantly comprised of a relatively flat background plane (i.e. the sea-floor). Resting on the sea-floor are various 3-D objects that are to be reconstructed in 3-D. In our method, we require knowledge of the background plane equation, namely, the normal vector $\mathbf{n}(R, \theta)$ in the sonar coordinates. This is determined by applying the method described by [29]. An initial elevation map of the bottom

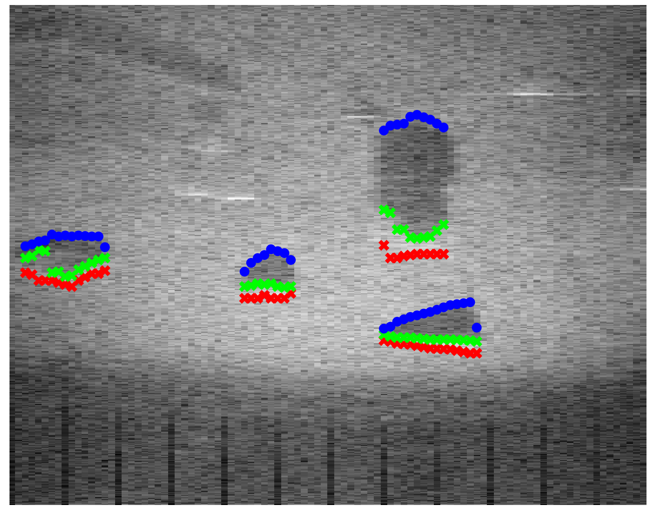

Fig. 4. Object boundaries and shadow edges are determined on bin-beam image. The bottom and occluding contours are respectively shown by red and green colors while the shadow edge is plotted by blue.

can be constructed from the normal vector. To be precise, for each point $\mathbf{P}$ on the surface with range R and azimuth θ, we calculate the elevation angle from $\mathbf{P} \cdot \mathbf{n} = -1$. This initial elevation map (or ϕ-map) is used to compute the zenith angles at the contours of all 3-D objects. For this purpose, object regions and their corresponding shadows are marked utilizing the method in [5] as follows:

- Compensating for reduced sensitivity towards the edges of the image;
- Computing the normal of the sea-floor to obtain an elevation-map of the sea-floor;
- Extracting highlights (objects);
- Pairing up object regions with corresponding cast shadows;
- Eliminating false positives, and adding new features, where necessary (manual correction);
- Computing the total elevation coverage $\Delta\phi_j^{TOT}$ (see Fig. 1(c)) of each object across the occupied beams.

A sample image with the determined object and shadow edges are demonstrated in Fig. 4. Since these objects sit on the sea-floor their shadows occur right behind them. The bottom contour of an object may be assumed to have the same elevation as the sea-floor. Therefore, using the surface normal of the sea-floor, $\mathbf{n}$, ϕ-angles along each object's bottom edge may be obtained [29]. On the other hand, the occluding contour is shown to have the same elevation as its shadow as in Fig. 1(c). Since the shadows are assumed to appear on the planar sea-floor, we may compute the elevation values through the shadow edges (by $\mathbf{P} \cdot \mathbf{n} = -1$), which are equivalent to the elevation values along the corresponding occluding object contours.

A flying object on the water column may also be identified similarly by selecting the shadow edge corresponding to the object bottom as well as its top. Since no distinct echoes can be observed from an acoustic shadow zone [25], nothing is

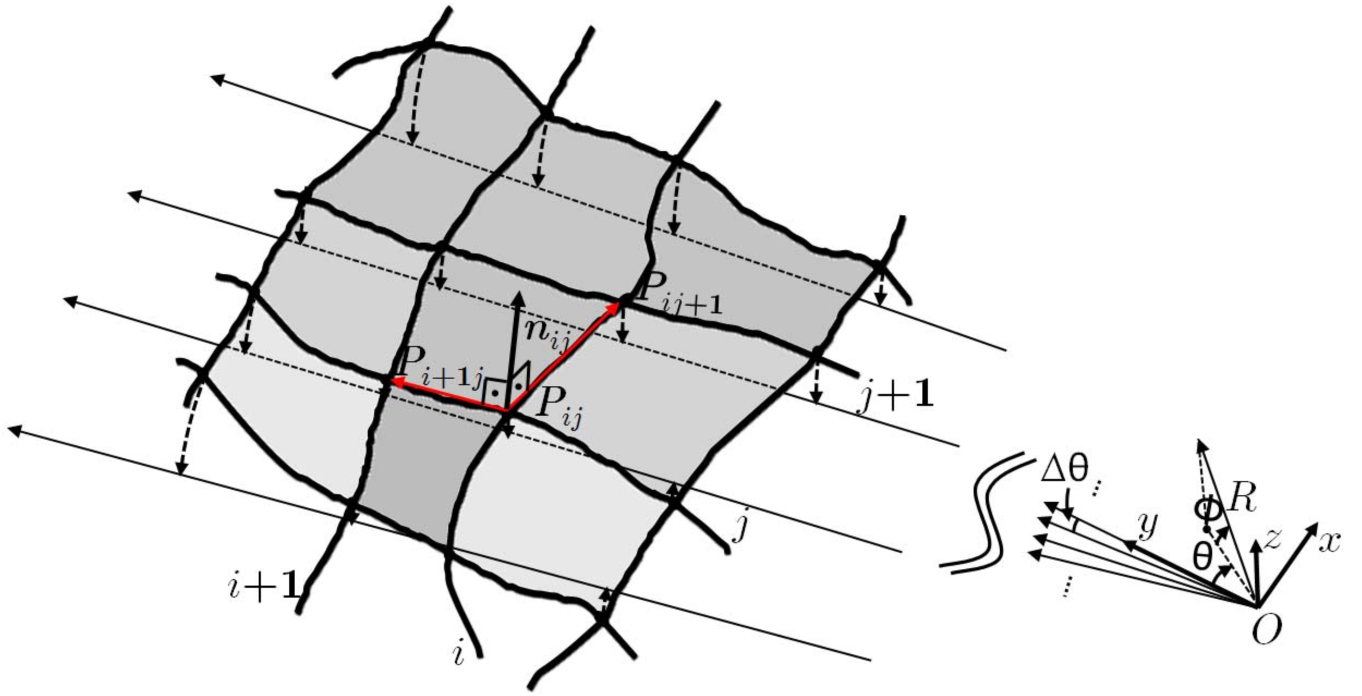

Fig. 5. Reconstruction is performed along constant range rows i and azimuth angles j. Starting on the bottom boundary, interior shading determines the complete 3-D structure. Updating the reflection coefficient, computation is repeated until the reconstructed occluding boundary coincides with the expected one determined by shadow constraints.

known about the completely shadowed pixels [13], hence these regions may not be reconstructed [7]. For example $|CE|$ in Fig. 1(c) may take various arbitrary shapes without having any effect on the measured signal. In our implementation these shadow regions are approximated to be the same level as the sea-floor.

B. Reconstructing the Interior

Using the scheme depicted in Fig. 5 an iterative solution which sequentially estimates consecutive discrete points ($P_{ij} = [x, y, z]_{ij}^T$) is proposed. Referring to *classical* SFS works [8], [18], [19] this kind of procedure may be defined as the sequential estimation of *characteristic strips* that start at the boundary and proceed towards the interior region. Another work [7] performs 3-D reconstruction of a side-scan sonar image by sequentially estimating depth for consecutive discrete points on a planar grid.

In Fig. 5 let ϕ_{ij}'s be given along the i'th row so that the 3-D coordinates of the object $\{\mathbf{P}_{ij}$, for $i = b_j, b_j{+}1, \dots o_j$, and $j = l, l+1, \dots, r\}$ are known. Here l and r represent the left/right boundaries of an object (in beam-bin image) while b_j and o_j represent the bottom/top boundaries. Then in order to estimate the 3-D coordinate $\mathbf{P}_{i+1j}$, solving for the unknown ϕ_{i+1j} is sufficient. Equivalently, computing the differential elevation coverage of the ij'th patch is enough since $\phi_{i+1j} = \phi_{ij} + \Delta\phi_{ij}$. At each point $\mathbf{P}_{ij}$ along the surface, four unknowns, $\Delta\phi_{ij}$ and $\mathbf{n}_{ij} = [n_x n_y n_z]_{ij}^T$, have to be determined. Referring to Fig. 5 the surface normal must satisfy

$$\mathbf{n}_{ij} \cdot \mathbf{d}_{ij} = 0 \tag{13}$$

$$\mathbf{n}_{ij} \cdot \mathbf{e}_{ij} = 0 \tag{14}$$

where $\mathbf{d}_{ij} = \mathbf{P}_{ij+1} - \mathbf{P}_{ij}$ and $\mathbf{e}_{ij} = \mathbf{P}_{i+1j} - \mathbf{P}_{ij}$. Furthermore

$$||\mathbf{n}_{ij}|| = 1 \ \forall i, j \tag{15}$$

Now let us rewrite (11) in terms of the surface normal

$$I_{ij}^{avg} = \frac{K\ I_0}{R^2} \int_{\phi=\phi_{ij}}^{\phi_{ij}+\Delta\phi_{ij}} \int_{\theta=\theta_j}^{\theta_j+\Delta\theta_j} \mathbf{n} \cdot \mathbf{r}\ cos\phi\ d\theta d\phi \tag{16}$$

Under the assumption that the intensity of each pixel is solely contributed by a single planar patch among which $\mathbf{n}$ is constant as shown in Fig. 3, one can further simplify the above integral by noting $\mathbf{r}$ being approximately constant along the integral limits (i.e. over the planar patch). Thus $\mathbf{n} \cdot \mathbf{r}$ term is approximately constant as well and can be taken outside the integral while the rest simplifies to

$$I_{ij}^{avg} = \vartheta_j\ \mathbf{n}_{ij} \cdot \mathbf{r}_{ij}\ \Delta\phi_{ij} \tag{17}$$

where $\vartheta_j = K\ I_0\ \Delta\theta_j$ is a constant. Note that $\mathbf{P}_{i+1j}$ and $\mathbf{r}_{ij}$ can be written in terms of $\mathbf{n}_{ij}$ and $\Delta\phi_{ij}$ as follows:

$$\mathbf{P}_{i+1j} = \begin{bmatrix} (R_{i+1j})\ cos(\phi_{ij} + \Delta\phi_{ij})\ sin(\theta_{i+1j}) \\ (R_{i+1j})\ cos(\phi_{ij} + \Delta\phi_{ij})\ cos(\theta_{i+1j}) \\ (R_{i+1j})\ sin(\phi_{ij} + \Delta\phi_{ij}) \end{bmatrix} \tag{18}$$

$$\mathbf{r}_{ij} = \frac{\mathbf{P}_{ij}}{||\mathbf{P}_{ij}||} \tag{19}$$

Therefore, the unknowns $\mathbf{n}_{ij}$ and $\Delta\phi_{ij}$ can be computed by simultaneously solving (13), (14), (15) and (17) where P_{i+1j} and $\mathbf{r}_{ij}$ are appropriately substituted from (18) and (19).

Finally the 3-D reconstruction procedure is given as follows:

1) Find the 3-D coordinates along the bottom and occluding contour of the surface to be reconstructed (i.e. P_{ij} for $i = b_j$ on the bottom contour and $i = o_j$ on the occluding contour of the j'th beam) by using the planar background assumption and shadow cues as explained in Section IV-A.
2) Initialize $\vartheta_j = 2(\frac{\sum_i \sqrt{I_{ij}^{avg}}}{\Delta\phi_j^{TOT}})^2$.
3) Compute P_{ij} ($\forall\ i,\ j$) towards the interior region by simultaneosly solving (13), (14), (15) and (17).
4) If the computed P_{ij} along the occluding boundary (i.e. $i = o_j$) are close to their values obtained in step 2 stop. Else set

$$\vartheta_j = \vartheta_j \frac{\sum_{i=b_j}^{o_j} \Delta\phi_{ij}}{\Delta\phi_j^{TOT}}$$

and go to step 3.

Here steps 2 and 4 enable the reconstructed 3-D surface to extend across the boundary determined from the constraints in Section IV-A.

C. Pre-Processing the Data

For a group of objects that violate the monotonically increasing range assumption, 3-D reconstruction can be performed after reducing the distortions in the measured signal by i) resolving simultaneous reflections from two patches and ii) discarding the ground echo.

1) Handling Two Separate Reflectors at the Same Range: "Monotonically increasing range" assumption may not be obeyed by all objects for which the reconstruction results would not be accurate. Even then, a special group of surfaces can be 3-D reconstructed after pre-processing the measured intensity. For easier understanding the concept is explained in 2-D on a circular arc as given in Fig. 6(a). For such a shape, starting from the bottom, range from the sonar first decreases until the closest point is reached over the circular arc and then keeps increasing. Many geometric shapes (i.e. sphere, cylinder, rectangle) and thus human made materials (pipes, sea-mines, submarines, boxes, bricks) may fall into the category of objects that display *two* reflecting patches for some range. Obviously, this is a violation of the *monotonically increasing range* (as well as *single patch at a given range bin*) assumption.

Consider Fig. 6(a,b) where the intensity values in between AB and AB' are caused by simultaneous reflections from two separate patches over the circular arc. First, three characteristic points (A, B, C) on the intensity profile are marked. Then, under a symmetry assumption (i.e. assuming both patches are about the same size), we use *half* the intensities and insert additional range bins to distinguish BA from AB' and generate a more accurate intensity profile across the object as demonstrated in Fig. 6(c).

Note that if the ground echo is not strong enough to yield a noticeable secondary spike, we recommend directly using

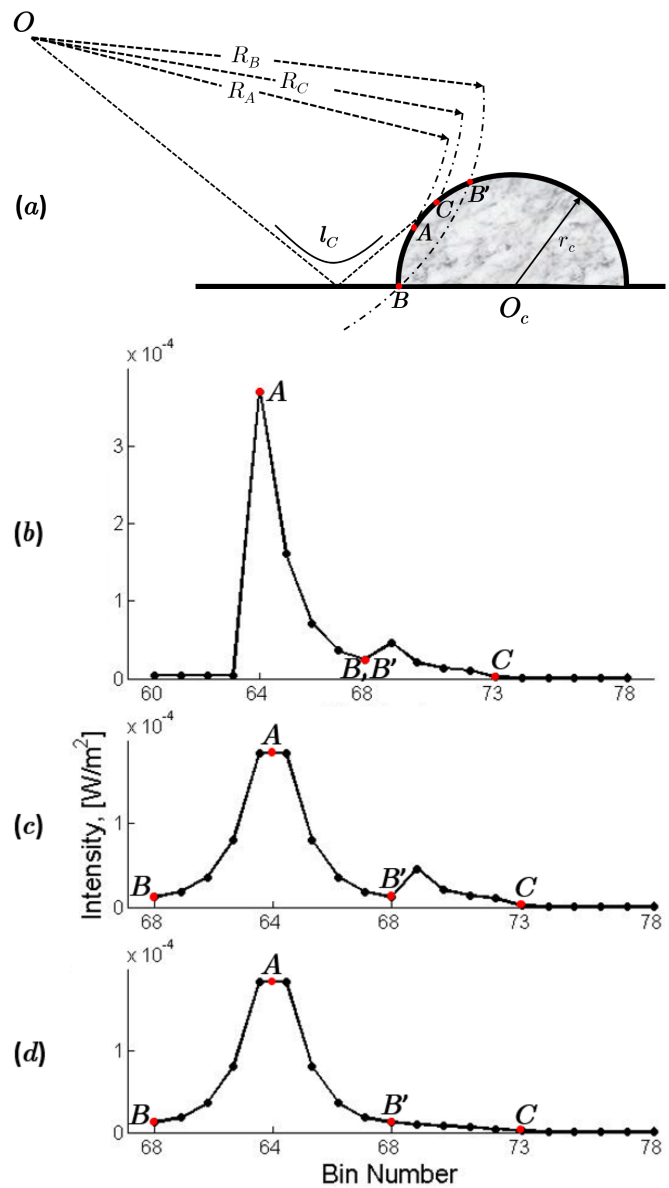

Fig. 6. (a) A circular object has two distinct simultaneous reflecting patches (i.e. in between BA and AB'). (b) The total measured intensity is processed so that (c) BA and AB' are separately represented in the intensity function by using half the values of the measured intensity. (d) As a last touch the ground-echo is removed and replaced by an interpolation of the remaining intensity function

the original data (without pre-processing) for reconstruction since in that case points B and/or C cannot be distinguished. Moreover, although for a group of objects (with two reflectors at the same range) the reconstruction accuracy improves by resolving simultaneous reflections from two patches, the problem becomes too complicated to be generally resolved for numerous simultaneous reflectors on an irregularly profiled surface.

2) Discarding Ground Echo: Proposed reconstruction method(s) assume that the intensity is consisted by only the direct reflection of the sound. In other words, all the higher order echoes due to multiple reflection of the sound wave

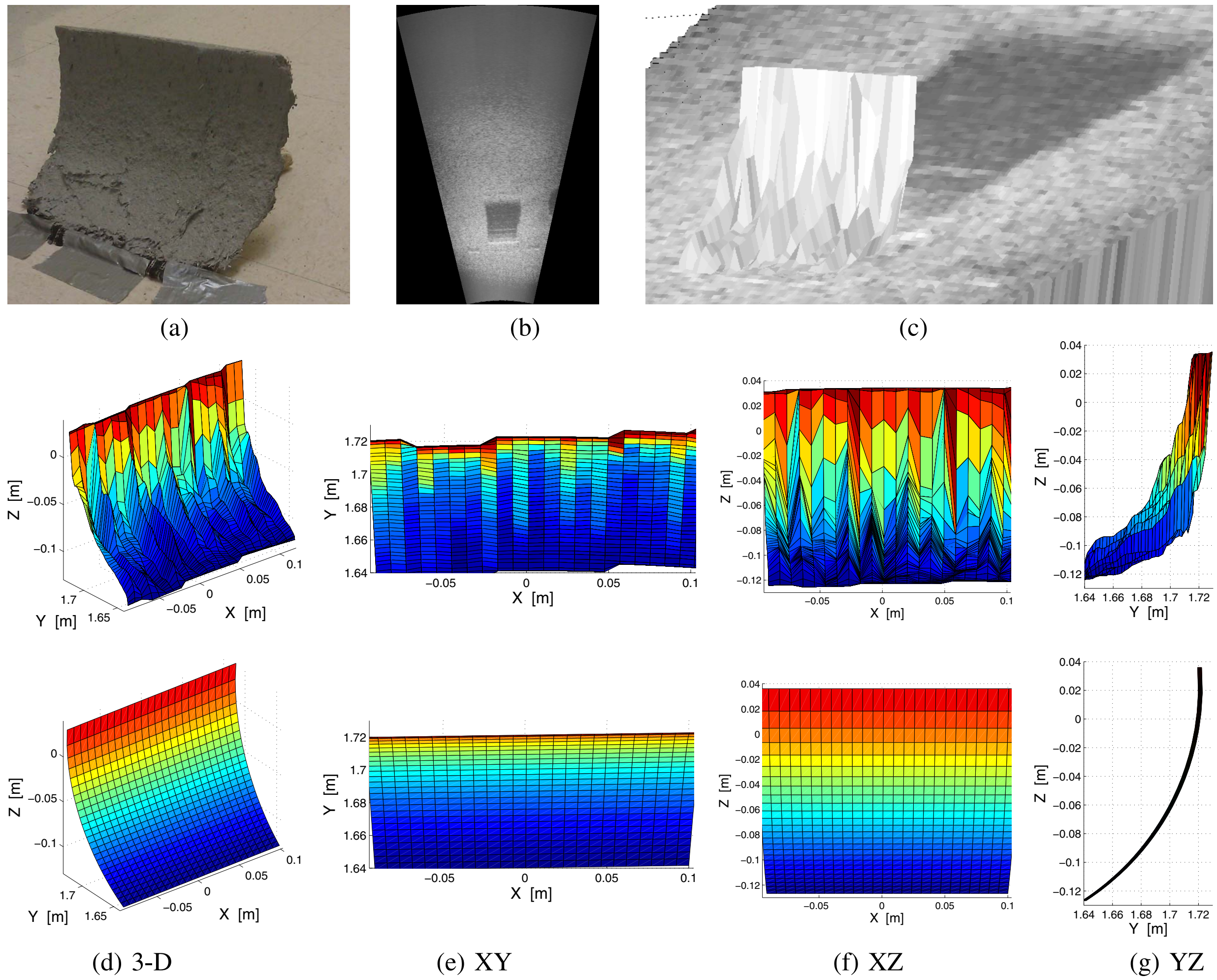

Fig. 7. Reconstruction of hemi-cylinder (concave). (a) Hemi-cylinder rock. (b) DIDSON image of hemi-cylinder. (c) Reconstructed hemi-cylinder. (d-g) Projections of 3-D reconstruction for cylindrical target (top) are compared to object model (bottom).

are assumed zero. However, practically the echoes formed by sound reflection through the ground to the object (or object to ground) may be large enough to cause strong secondary spikes in the intensity function as shown in Fig. 6(b). Such complex variations on intensity degrades the accuracy and causes the reconstructed surface to be distorted especially along the range bins corresponding to the secondary spike. Consider the circular arc in Fig. 6(a), where points A and B respectively represent the object's closest point to the sonar and object's bottom (i.e. intersection with the floor), while point C represents the end of ground-echo. Measured and processed intensities are given in (b) and (c) while the echo-discarded intensity is shown in (d). In this work, the secondary spike is simply manually marked (i.e. points B and C in Fig. 6(b)) and discarded while the remaining intensity function is interpolated in between BC.

V. Experimental Results

Images collected in a pool facility using well-known geometrical shapes, such as convex/concave hemi-cylinders and a hemi-sphere as target objects, provide data with ground truth; see Fig. 7, 8, 9(a). Their DIDSON views are given in (b). The 3-D reconstruction of these objects are shown in (c-g). The results for the concave hemi-cylinder (Fig.7) are obtained by directly using the measured intensities since it obeys the monotonically increasing range assumption and no secondary intensity spike is observable. On the other hand, partly along their profiles a convex hemi-cylinder and hemi-sphere have two simultaneous scattering patches for a given range. Furthermore we also observe a considerable amount of ground-echo content in the measured signals. Double reflections along these objects are accounted for and the ground-echo is discarded as explained in Section IV-C. Comparisons with the ideal object models in Fig. 8, 9(c-g) demonstrate that our method can provide decent profiles.

We also performed another experiment that involves the reconstruction of 3 fish in the water column and on the bottom. This confirms that our method equally applies for free swimming objects within the water column. The only requirement is that their cast shadows on the sea-floor are

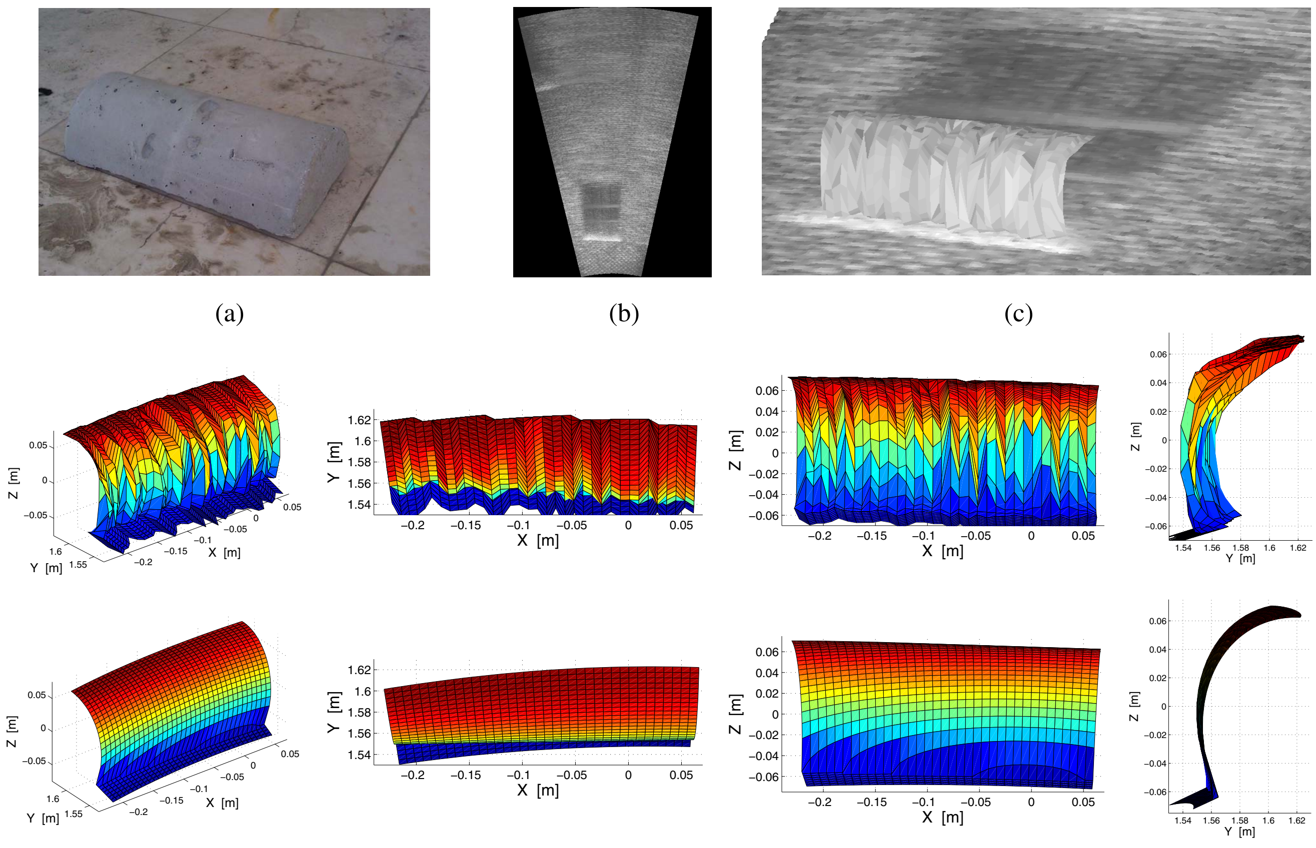

Fig. 8. Reconstruction of hemi-cylinder (convex). (a) Hemi-cylinder rock. (b) DIDSON image of hemi-cylinder. (c) Reconstructed hemi-cylinder. (d-g) Projections of 3-D reconstruction for cylindrical target (top) are compared to object model (bottom).

distinguishable. Referring to Fig. 10(a), the bigger fish swim highest within the water column, the second fish on the left is closer to the sea-floor, while the third fish seems to be right near the bottom. The side and front views of the reconstruction verify a sufficiently complete representation of the scene characteristics, in terms of object sizes, locations, shapes, and in fact some details of the larger fish (i.e. fin and tail). Not surprisingly, the 2-D FS image in 10(a)(a) is ambiguous in conveying which object is higher or lower with respect to the ground level, mainly because the knowledge of elevation angle ϕ is absent. The recovery of the unknown elevation angles enables to solve the puzzle as shown in (c), (d).

VI. Conclusion

In this work, we demonstrate that a 2-D FS sonar image contains more information than the eye may extract. We have developed a model to describe the formation of a 2-D forward-scan sonar image, and how one may recover the unknown third dimension, namely, the elevation angles from a single image, under some conditions. This is achieved by first determining the 3-D information at object boundaries based on cast shadow cues and then employing the proposed image formation model to obtain the remaining elevation values through each object. For objects with smooth, monotonic surfaces, whether resting on the sea-floor or moving within the water column, the reconstructions can be rather accurate. In addition to environmental characterization, mapping and modeling, such a reconstruction can play a key role in improving performance of image registration, mosaicing, landmark detection, object recognition, change detection, obstacle avoidance and precision navigation.

Acknowledgment

The first author is supported through University of Miami Graduate Fellowship Program. This work was supported in part by the Office of Naval Research under project N000141210076. Contents and opinions expressed in this paper represent solely the views of the authors, and do not necessarily represent official views of the ONR. Travel and conference expenses are supported in part by the Graduate Activity Fee Allocation Committee at University of Miami.

References

[1] *Blueview P900 Series Sonar*, http://www.blueview.com/P900-Series.html, [Online; accessed 31-July-2012].

[2] *EdgeTech - Ultra-high Frequency Side-scan Sonars 4125, 4200*, http://www.edgetech.com/edgetech/gallery/category/side-scan-sonar-systems, [Online; accessed 21-October-2012].

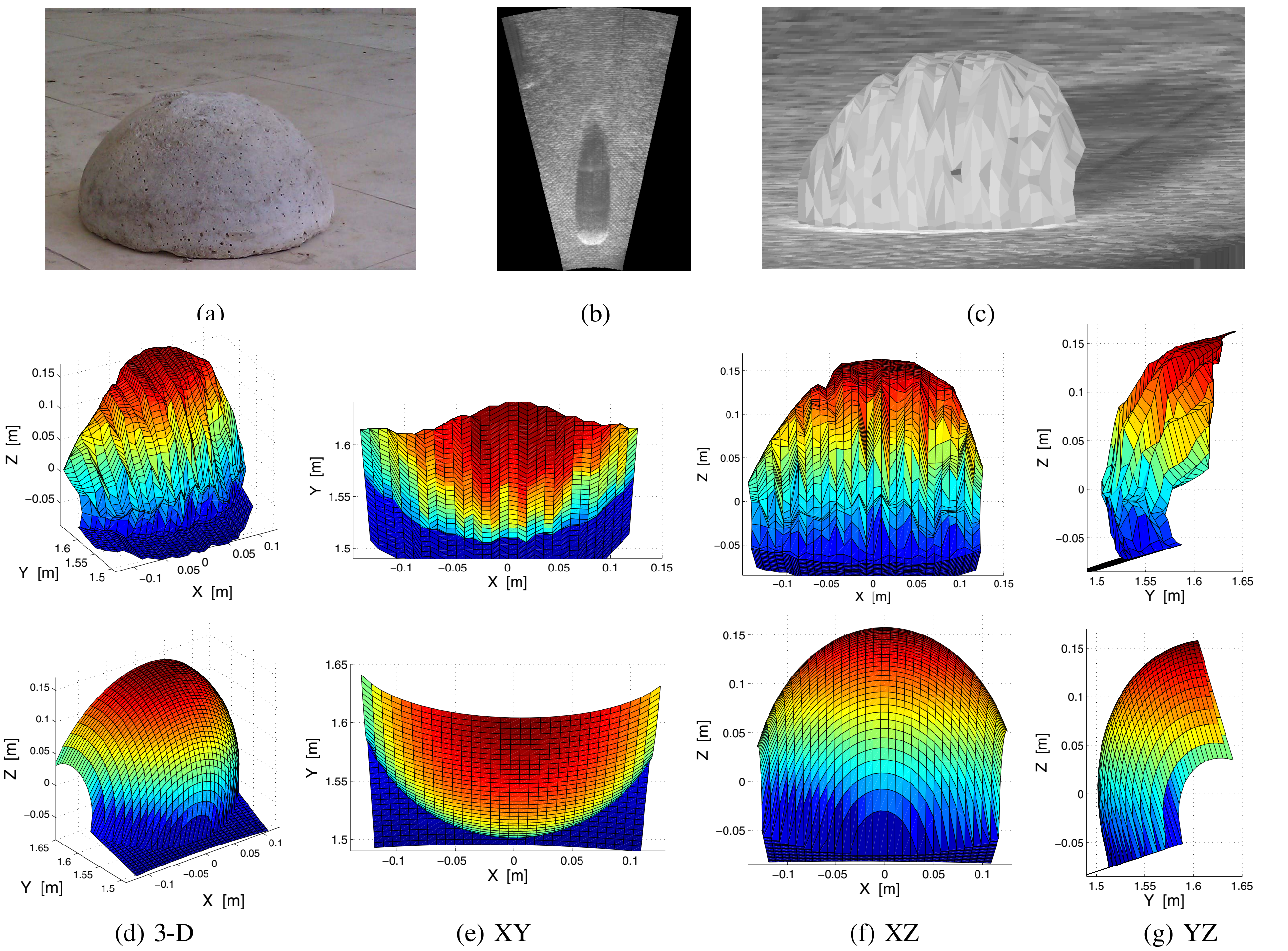

Fig. 9. Reconstruction of hemi-sphere. (a) Hemi-spherical rock. (b) DIDSON image of hemi-sphere. (c) Reconstructed hemi-sphere. (d-g) Projections of 3-D reconstruction for hemi-spherical target are compared to object model (bottom).

[3] *Gemini 720i - Multibeam Imaging Sonar*, http://www.tritech.co.uk/product/gemini-720i-300m-multibeam-imaging-sonar, [Online; accessed 31-July-2012].

[4] *Sound Metrics - DIDSON 300 m Sonar*, http://www.soundmetrics.com/products/imaging-sonars/didson-300, [Online; accessed 31-July-2012].

[5] M.D. Aykin and S. Negahdaripour, *On feature matching and image registration for 2-d forward-scan sonar imaging*, Journal Of Field Robotics, Accepted (2013).

[6] E.O. Belcher, W.L.J. Fox, and W.H. Hanot, *Dual-frequency acoustic camera: a candidate for an obstacle avoidance, gap-filler, and identification sensor for untethered underwater vehicles*, OCEANS'02 MTS/IEEE, vol. 4, IEEE, 2002, pp. 2124–2128.

[7] K. Bikonis, A. Stepnowski, and M. Moszynski, *Computer vision techniques applied for reconstruction of seafloor 3d images from side scan and synthetic aperture sonars data*, Journal of the Acoustical Society of America **123** (2008), no. 5, 3748.

[8] Andrew Blake, Andrew Zisserman, and Greg Knowles, *Surface descriptions from stereo and shading*, Image and Vision Computing **3** (1985), no. 4, 183–191.

[9] H. Boehme, NP Chotiros, LD Rolleigh, SP Pitt, AL Garcia, TG Goldsberry, and RA Lamb, *Acoustic backscattering at low grazing angles from the ocean bottom. part i. bottom backscattering strength*, The Journal of the Acoustical Society of America **77** (1985), 962.

[10] Heinrich H Bülthoff and Hanspeter A Mallot, *Integration of depth modules: stereo and shading*, JOSA A **5** (1988), no. 10, 1749–1758.

[11] W.S. Burdic, *Underwater acoustic system analysis*, vol. 113, Prentice-Hall Englewood Cliffs, New Jersey, 1984.

[12] D.E. Clark and J. Bell, *Bayesian multiple target tracking in forward scan sonar images using the phd filter*, Radar, Sonar and Navigation, IEE Proceedings-, vol. 152, IET, 2005, pp. 327–334.

[13] E. Coiras and J. Groen, *3d target shape from sas images based on a deformable mesh*, Proceedings of the 3rd International Conference on Underwater Acoustic Measurements (UAM), Nafplion, Greece, 2009, pp. 382–390.

[14] James Edwin Cryer, Ping-Sing Tsai, and Mubarak Shah, *Integration of shape from shading and stereo*, Pattern recognition **28** (1995), no. 7, 1033–1043.

[15] Pascal Fua and Yvan G Leclerc, *Object-centered surface reconstruction: Combining multi-image stereo and shading*, International Journal of Computer Vision **16** (1995), no. 1, 35–56.

[16] N.O. Handegard and K. Williams, *Automated tracking of fish in trawls using the didson (dual frequency identification sonar)*, ICES Journal of Marine Science: Journal du Conseil **65** (2008), no. 4, 636–644.

[17] B. Horn and M.J. Brooks, *Shape from shading*, MIT press, 1989.

[18] Berthold KP Horn, *Shape from shading: A method for obtaining the shape of a smooth opaque object from one view*, MIT Artificial Intelligence Laboratory (1970).

[19] ______, *Obtaining shape from shading information*, Shape from shading, MIT press, 1989, pp. 123–171.

[20] JB Hsieh, JI Olsonbaker, and WL Fox, *A screening application for image data collected by an acoustic lens sonar*, Tech. report, DTIC Document, 2005.

[21] K. Ikeuchi and B.K.P. Horn, *Numerical shape from shading and occluding boundaries*, Artificial Intelligence **17** (1981), no. 1, 141–184.

[22] Katsushi Ikeuchi, Memo No, and Katsushi Ikeuchi, *Constructing a depth*

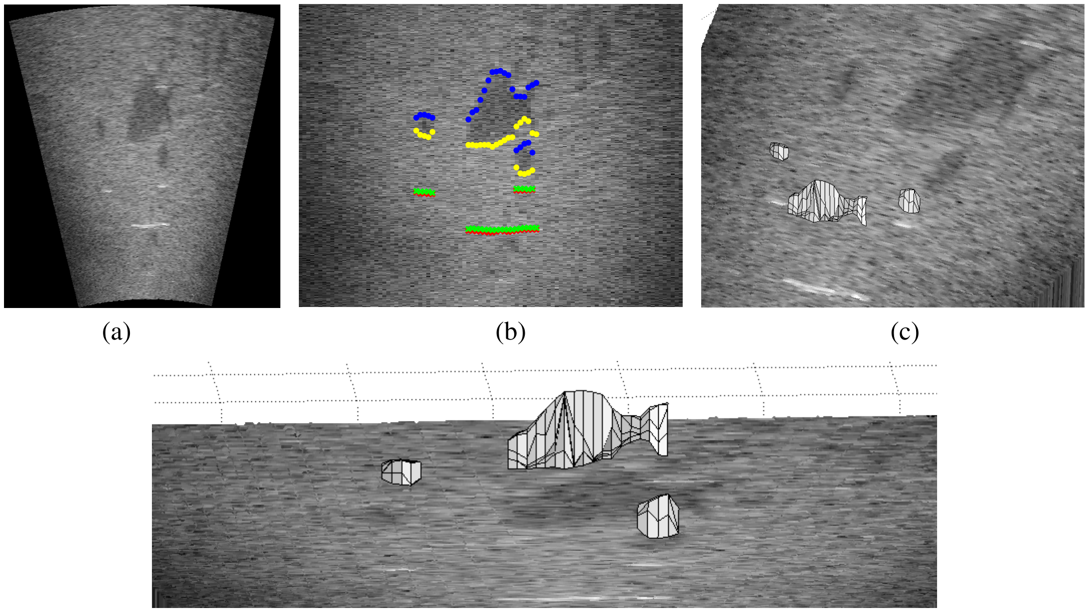

Fig. 10. Reconstruction of fish in the water column. (a) DIDSON image of three fish with distinct shadows. (b) Close and far ends of the highlights and shadows are marked. (c-d) Reconstructed image.

map from images, AI Memo AIM-744, Artificial Intelligence Laboratory, MIT, Citeseer, 1983.

[23] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, *Imaging sonar-aided navigation for autonomous underwater harbor surveillance*, Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, IEEE, 2010, pp. 4396–4403.

[24] K. Kim, N. Intrator, and N. Neretti, *Image registration and mosaicing of noisy acoustic camera images*, Electronics, Circuits and Systems, 2004. ICECS 2004. Proceedings of the 2004 11th IEEE International Conference on, IEEE, 2004, pp. 527–530.

[25] D. Langer and M. Hebert, *Building qualitative elevation maps from side scan sonar data for autonomous underwater navigation*, Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on, IEEE, 1991, pp. 2478–2483.

[26] J.L. Leonard, R.N. Carpenter, and H.J.S. Feder, *Stochastic mapping using forward look sonar*, Robotica **19** (2001), no. 5, 467–480.

[27] P. Mamassian, D.C. Knill, and D. Kersten, *The perception of cast shadows*, Trends in cognitive sciences **2** (1998), no. 8, 288–295.

[28] S.K. Nayar, K. Ikeuchi, and T. Kanade, *Surface reflection: physical and geometrical perspectives*, IEEE Transactions on Pattern Analysis and Machine Intelligence **13** (1991), no. 7, 611–634.

[29] S. Negahdaripour, *On 3-d scene interpretation from f-s sonar imagery*, IEEE/MTS Oceans'12 Conference, IEEE, October 2012.

[30] S. Negahdaripour, P. Firoozfam, and P. Sabzmeydani, *On processing and registration of forward-scan acoustic video imagery*, Computer and Robot Vision, 2005. Proceedings. The 2nd Canadian Conference on, IEEE, 2005, pp. 452–459.

[31] S. Negahdaripour and X. Xu, *Mosaic-based positioning and improved motion-estimation methods for automatic navigation of submersible vehicles*, Oceanic Engineering, IEEE Journal of **27** (2002), no. 1, 79–99.

[32] Y. Petillot, I.T. Ruiz, and D. Lane, *Auv navigation using a forward looking sonar*, Oceanology International 2000 (2000).

[33] A.D. Pierce and PW Smith, *Acoustics: an introduction to its physical principles and applications*, Physics Today **34** (1981), 56.

[34] S. Stanic, KB Briggs, P. Fleischer, RI Ray, and WB Sawyer, *Shallow-water high-frequency bottom scattering off panama city, florida*, The Journal of the Acoustical Society of America **83** (1988), 2134.

[35] J. Tong, J. Han, W. Shen, and P. Xu, *Mosaicing of acoustic video images for underwater structure inspection*, ASME, 2010.

[36] R.J. Urick and R.J. Urick, *Principles of underwater sound*, vol. 3, McGraw-Hill New York, 1983.

[37] J. Vaganay, M. Elkins, D. Esposito, W. O'Halloran, F. Hover, and M. Kokko, *Ship hull inspection with the hauv: Us navy and nato demonstrations results*, OCEANS 2006, IEEE, 2006, pp. 1–6.

[38] J. Vaganay, ML Elkins, S. Willcox, FS Hover, RS Damus, S. Desset, JP Morash, and VC Polidoro, *Ship hull inspection by hull-relative navigation and control*, OCEANS, 2005. Proceedings of MTS/IEEE, IEEE, 2005, pp. 761–766.

[39] M. Walter, F. Hover, and J. Leonard, *Slam for ship hull inspection using exactly sparse extended information filters*, Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on, IEEE, 2008, pp. 1463–1470.

[40] S. Xie, J. Luo, J. Chen, and Y. XU, *Stability augmentation and mosaic method of forward-scan sonar images*, Proceedings of SPIE, vol. 8372, 2012, p. 837214.

[41] Y. Yacoob and L. Davis, *Segmentation of planar objects and their shadows in motion sequences*, International journal of computer vision **67** (2006), no. 1, 53–69.

[42] S.C. Yu, *Development of real-time acoustic image recognition system using by autonomous marine vehicle*, Ocean Engineering **35** (2008), no. 1, 90–105.

[43] S.C. Yu, T.W. Kim, A. Asada, S. Weatherwax, B. Collins, and J. Yuh, *Development of high-resolution acoustic camera based real-time object recognition system by using autonomous underwater vehicles.*, OCEANS 2006, IEEE, 2006, pp. 1–6.