3D Reconstruction of Underwater objects using wide-aperture Imaging SONAR

3D Reconstruction of Underwater objects using wide-aperture Imaging SONAR

Thomas Guerneve
Ocean Systems Lab
Heriot Watt University
SeeByte Ltd
Edinburgh, Scotland
tsg1@hw.ac.uk

Kartic Subr
University of Edinburgh
Edinburgh, Scotland
K.Subr@ed.ac.uk

Yvan Petillot
Heriot Watt University
Edinburgh, Scotland
Y.R.Petillot@hw.ac.uk

Abstract

The estimation of geometric structure of objects located underwater underpins a plethora of applications such as mapping shipwrecks for archaeology, monitoring the health of coral reefs, detecting faults in off-shore oil rigs and pipelines, detection and identification of potential threats on the seabed, etc. Acoustic imaging is the most popular choice for underwater sensing. Underwater exploratory vehicles typically employ *wide-aperture* Sound Navigation and Ranging (SONAR) imaging sensors. Although their wide aperture enables scouring large volumes of water ahead of them for obstacles, the resulting images produced are blurry due to integration over the aperture. Performing 3D reconstruction from this blurry data is notoriously difficult. This challenging inverse problem is further exacerbated by the presence of speckle noise and reverberations. The state of the art methods in 3D reconstruction from sonar either require bulky and expensive matrix-arrays of sonar sensors or additional narrow-aperture sensors. Due to its low footprint, the latter induces gaps between reconstructed scans. Avoiding such gaps requires slow and cumbersome scanning by the vehicles that carry the scanners.

In this paper, we present two reconstruction methods enabling on-site 3D reconstruction from imaging sonars of any aperture. The first of these presents an elegant linear formulation of the problem, as a blind deconvolution with a spatially-varying kernel. The second method is a simple algorithmic approach for approximate reconstruction, using a non-linear formulation. We demonstrate that our simple approximation algorithms perform 3D reconstruction directly from the data recorded by wide-aperture systems thus eliminating the need for multiple sensors to be mounted on underwater vehicles for this purpose. Additionally, we observe that the wide aperture may be exploited to improve the coverage of the reconstructed samples (on the scanned object's surface). We demonstrate the efficacy of our algorithms on simulated as well as real data acquired using two sensors and compare our work to the state of the art in sonar reconstruction. Finally, we show the employability of our reconstruction methods on field data gathered by an Autonomous Underwater Vehicle.

1 Introduction

1.1 Motivation

With the increasing importance of autonomous vehicles in underwater research and engineering applications, the detection, identification and localization of underwater objects is a central problem. In particular, the maintenance of off-shore infrastructure often relies on the ability to detect underwater man-made structures. The importance of obtaining accurate 3D representation of the sub-sea environment underpins a plethora of applications such as mapping shipwrecks for archaeology, monitoring the health of coral reefs, detecting faults in off-shore oil rigs and pipelines, detection and identification of potential threats on the seabed, etc. There has been considerable effort towards estimating 3D structure from optical images (Massot-Campos and Oliver-Codina, 2015). This can be very challenging underwater and even impossible in poor visibility conditions that are prevalent in many coastal environments. LASER based systems such as LIDARs are popular for accurate 3D scanning of objects in air but are rarely used on underwater vehicles due to their relatively high cost, limited range and small footprint. Acoustic sensors are ubiquitous in underwater environments due to the favourable properties of sound propagation in water with respect to other sensing modalities. Large-scale sonar sensing is typically performed using low to medium frequency acoustic waves and large antennas whilst high resolution sensing is achieved using smaller antennas and high frequency waves.

In this paper we are interested in using commodity acoustic systems, which can be mounted on small underwater platforms to obtain fine-scale 3D reconstruction of objects within short ranges in all weather conditions. This can be achieved using pencil-beam systems (*profilers*) exhibiting a very narrow aperture and a low spatial footprint or *imagers* which use a wider beam for general scene acquisition and understanding. Recent research work based on carving methods (Guerneve and Petillot, 2015; Aykin and Negahdaripour, 2015) have demonstrated the feasibility of using acoustic imagers to obtain a 3D map of the environment, recovering enough information to enable manual object recognition. These systems, such as the BlueView sonar or Didson acoustic cameras provide high-resolution imaging at short ranges, are widely available on existing underwater platforms and have traditionally been employed to perform other tasks such as collision avoidance (Y. Petillot and Lane, 2001), sea floor mapping (N. Hurtos and Salvi, 2014) or target tracking (J. Folkesson and Williams, 2007).

The large spatial footprints of wide-aperture sensors enable real-time imaging of considerable volumes of water. However their wider aperture results in a blur, via a more complex imaging model, thereby reducing spatial resolution. There are key benefits for the underwater community if 3D reconstruction can be achieved from such images. Perhaps most importantly, a single imaging sensor can be used for mapping, navigation and 3D reconstruction where currently two or more dedicated sensors would be required, therefore reducing the sensor payload on AUVs (Autonomous Inspection Vehicles). Additionally, the field of view of the sensor enables better coverage than traditional profiling systems when the vehicle navigation is not perfectly controlled.

3D reconstruction from wide-aperture sonar is an ill-posed inverse problem which is impossible to solve without prior assumptions. This is exacerbated by the process of acoustic sensing which is notoriously noisy, unable to distinguish multi-path and internal reflections as well as subject to seabed reverberation. The image in fig. 1-a exemplifies the poor quality of the acquired sonar images. We propose two different formulations to address this problem: first, in the form of classical blind deconvolution, with spatially-varying kernels, and later as a non-linear approximation (akin to voxel-carving) that lends itself to a simple implementation. While the first formulation assumes rectilinear motion of the sensor, the second approach allows unconstrained trajectories. We adapt a constrained-linear solver (sparse) to solve the former as a batch-processing while the second method offers a per-image iterative method. We perform quantitative analyses of the two methods and demonstrate their relative benefits in controlled experiments using simulations as well as with measurements in a tank, using two different sensors that are commonly embedded in underwater platforms. We verified the suitability of our reconstruction method on field data acquired by an AUV, exhibiting the ability to recover accurate 3D representation of the structures as illustrated in fig. 1-b-c.

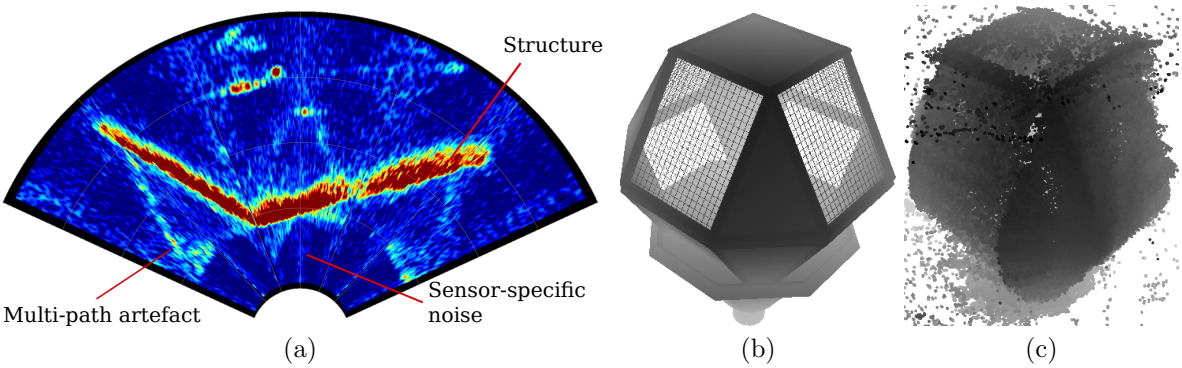


Figure 1: a) Typical BlueView P900-130 SONAR image acquired by Subsea 7 AIV (Autonomous Inspection Vehicle) during a structure inspection mission, exhibiting noise stripes and multipath artefacts. b) The model of the inspected structure exhibits a complex polygonal shape. c) 3D reconstruction of the structure based on a set of BlueView SONAR images and vehicle navigation data.

1.2 Related work

Underwater acoustic imaging has been widely researched because of inherent limitations of electro-optical systems (Massot-Campos and Oliver-Codina, 2015). The physical size of the acoustic antenna dictates the horizontal and vertical aperture of the imager whilst the frequency influences its range and its range resolution. A number of different sensors have been proposed to perform imaging and 3D reconstruction of underwater scenes. Pencil-beam sonars, akin to an acoustic laser pointer, have been widely used for 3D reconstruction (Jouvencel et al., 2012; **?**) and compare favourably to laser systems subsea (Georgios Papadopoulos and Patrikalakis, 2011). Although they provide results over medium-size areas at high resolution (cm), when used in a phased-array configuration the scanning and reconstruction of large areas requires very dense sampling and careful processing of the data. Since operation time is often a scarce resource in field applications, the ability of increasing the coverage per time spent on the field is of prime importance.

One dimensional antenna systems producing a narrow horizontal beam and a wide vertical beam (side-scan sonars) have been used to image the seabed by concatenating a series of narrow scan-lines to form an image. These systems have recently been superseded by SAS (Synthetic Aperture Sonar) systems (Hawkins, 1996; Callow, 2003) which create an arbitrarily long antenna by coherently integrating multiple views of the scene along a known trajectory (in a similar way to tomography). In both cases, the system produces a 2D image as obtained by structured lighting of a 3D scene.

Attempts have been made to extract the 3D information from the 2D imagery using shape from shading (Coiras and Groen, 2009; Enrique Coiras and Lane, 2007) or image segmentation (K. Bikonis and Moszynski, 2008). More recently, 3D reconstruction from SAS data has been demonstrated using multiple circular observations around a target and wider fractional bandwidth (**?**). While the 3D reconstructed targets exhibit good accuracy, the constraint of acquiring multiple observations along circles make its use time-consuming and potentially impractical on an offshore field due to the presence of other structures around the inspected elements. More importantly, SAS techniques crucially require very accurate navigation (typically a fraction of the wavelength (Hawkins, 1996)) which can only be obtained using expensive navigation sensors such as Inertial Navigation Systems.

These 2D imaging systems have been further extended to full 3D by using two dimensional arrays of receivers. The Echoscope technology (Davis and Lugsdin, 2005) demonstrated results with an angular resolution of 0.4° and a 1 cm range resolution. However, the sensor is extremely expensive and composed of large sensors making it impractical for integration on small underwater platforms. Following a similar idea, an ensemble of 54 pencil-beam sensors has been used in (Fairfield et al., 2007) providing direct 360° sensing capability but required custom integration on the vehicle.

For all these systems, widening the aperture increases the coverage of the beam thereby reducing the required sampling rate. However, this also results in a blurring along the aperture thus hampering the precision with which objects can be scanned. Consequently wide-aperture (imaging) sonars have until recently been overlooked for applications requiring high-resolution 3D acquisition. Recovering 3D structure from the images blurred by the wide aperture is an ill-posed problem (Aykin and Negahdaripour, 2013a).

Various approaches exploit specific prior assumptions about the environment to make the problem more tractable. E. g. acoustic shadows (Zerr and Stage, 1996), initial coarse maps using a CAD (Computer Aided Design) model or the result of an optical reconstruction (Ozog et al., 2015), etc. Recent work (Huang and Kaess, 2015) proved the feasibility of applying structure from motion techniques to Didson sonars, using manually selected feature points with specified associations (correspondences). Following the same idea, inspired by approaches in computer vision, stereo matching techniques have been employed towards purely acoustic features (N. Brahim and Solaiman, 2011; Aykin and Negahdaripour, 2013b) as well as opti-acoustic features (Babaee and Negahdaripour, 2015). The latter is sensitive to lighting conditions, requires pristine visibility conditions underwater and constrains the trajectory to being circular around the object of interest. Inspired by computed tomography methods, elevation information has also been extracted by combining multiple scans (Zerr and Stage, 1996) acquired in a circular trajectory around the object of interest followed by classification of the returns into echo, shadow and background. Due to water currents and safety margins, such constrained trajectories are often impractical on a real field. Additionally, the use of shadows restrict the reconstruction to areas featuring a solid background such as a seabed which does not necessarily hold true, for example when inspecting a vertical structure such as a pipeline.

In summary, 3D reconstruction of underwater objects from acoustic imagery is a difficult problem that has only been partially addressed. Whilst 3D SONARs exist, their limited availability, large size and prohibitive cost means they are seldom used in practical systems. Similarly, SAS approaches provided good 3D reconstruction results using coherent processing but required expensive navigation sensors, strict path constraints and complex signal processing. In this paper, we propose simple algorithms to obtain 3D reconstructions from wide-aperture sonar with errors comparable to reconstructions using pencil-beam sonar. Based on incoherent processing, our approach does not require expensive navigation systems such as INS and is adapted to the limited computing resource available on traditional AUVs. Since wide-aperture imaging sonars are commonly embedded on vehicles to perform fundamental tasks such as collision avoidance (Y. Petillot and Lane, 2001), sea floor mapping (N. Hurtos and Salvi, 2014) and target-tracking (J. Folkesson and Williams, 2007), our proposed algorithm will eliminate the need for a dedicated sensor to perform 3D reconstruction of underwater objects.

2 Review: Imaging model

Multi-beam imaging SONARs capture reflectivity across a given interval of range (distance from sensor) and a fan of directions within a wide field of view. With the exception of Mechanically-Scanned Imaging SONARs (MSIS) where low imaging rates limit the speed of operations (**?**), these systems are usually composed of an array of hydrophones, where each element consists of an emitter and a receiver. The former emits beams of pulses and the latter records local pressure variation due to any reflected pulses. A second-order differential equation (wave equation) describes the acoustic pressure field in terms of the emitted signal. Using this equation, and the observed delay between the emitted and reflected signals, the distance to reflectors along each beam can be calculated. A vector of *intensities*, or acoustic power per unit area, is recorded at each element corresponding to reflections from points at different distances from the element. The intensities measured by a sensor-array can be arranged in the form of an image spanning directions and range (distance). As opposed to multi-beam profiling SONARs (or pencil-beam SONARs), imaging SONARs typically feature a large vertical aperture (7 to 20°) providing a 2D representation of a large volume of water.

Using cylindrical coordinates, the intensity at each pixel (r, ϕ) on the recorded image (see fig. 2), assuming Lambertian reflection (Aykin and Negahdaripour, 2013a), is approximately the integral of reflected intensities

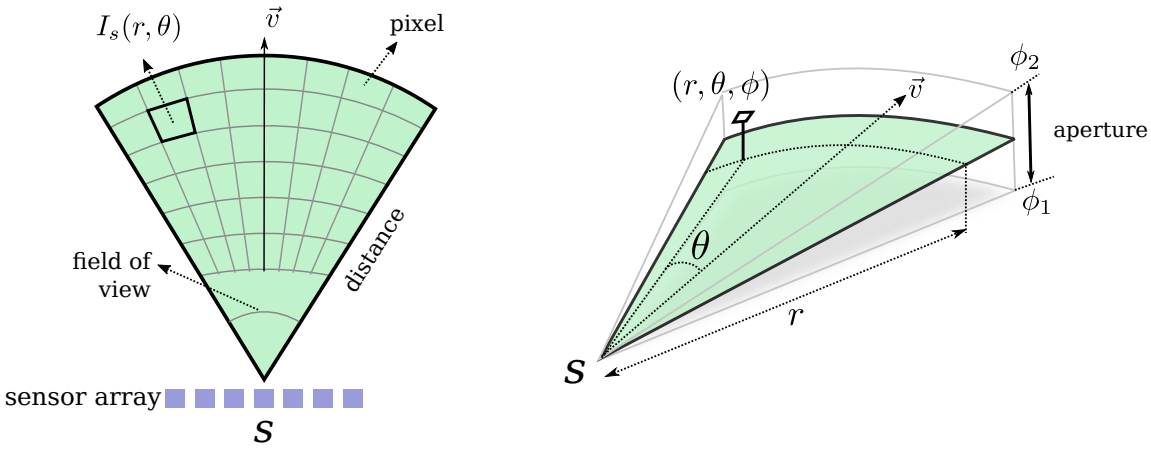


Figure 2: Notation used in the SONAR imaging model (see eq. 1).

from points (r, θ, ϕ), over the aperture $[\phi_1, \phi_2]$:

$$I_s(r,\theta) \quad = \quad \int_{\phi_1}^{\phi_2} \beta(\phi) \;\; V_s(r,\theta,\phi) \;\; \frac{\vec{v} \cdot \vec{n}_{r\theta\phi}}{\|\vec{v}\| \; \|\vec{n}_{r\theta\phi}\|} \;\; \mathrm{d}\phi. \tag{1}$$

Here $I_s(r, \theta)$ is the measured reflected intensity with the sensor-array at s, β characterises the angular distribution of energy in the beam which is known as the *beam pattern*, V_s represents the albedo of the reflector at (r, θ, ϕ) wrt the source location s and the last term in the integral is the cosine of the angle between the direction of propagation of the beam $\vec{v}$ and the surface normal $\vec{n}$ of the object at (r, θ, ϕ). In practice, the recorded values include contributions due to reverberations (multi-bounce paths with the same length) as well as measurement noise.

Wide-aperture systems record blurry projections of the reflected intensity, from within the view frustum, due to the integration over ϕ. While this hampers precise estimation of the distribution of V_S, it enables larger swathes of space to be covered by the scanner which is useful in detection of obstacles along the direction of motion. Reducing the aperture enables more precise estimation of the structure of reflectors, in the limit, but only provides information along a single 'scan-line'. In this paper, we propose to improve the estimation of 3D structure using wide-aperture systems.

2.1 Contributions

In this paper we unify our previous results (Guerneve and Petillot, 2015), considered state-of-the-art on approximate 3D reconstruction using carving with a more fundamental formulation for the ill-posed problem of 3D reconstruction from wide-aperture sonar images. Then we perform a rigorous evaluation of both approaches, comparing their relative benefits. In summary, our contributions are:

1. We present two formulations for 3D reconstruction from wide-aperture sonar images: blind deconvolution with a spatially-varying kernel and non-linear carving.
2. We adapt a sparse-linear solver to approximately solve the blind deconvolution.
3. We perform the first quantitative analysis of error and coverage of sonar carving.
4. We demonstrate using controlled experiments that our formulation as a blind-deconvolution leads to improved surface sampling coverage without an increase in error compared to state of the art.

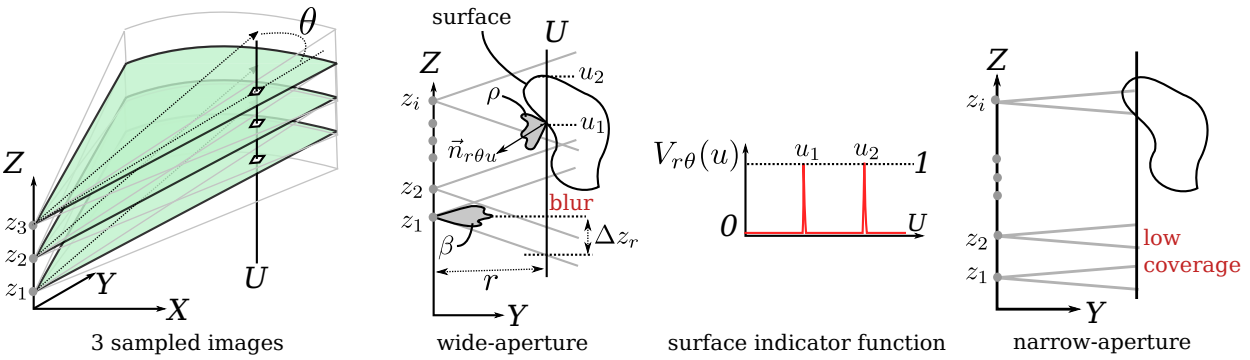


Figure 3: We show that 3D reconstruction amounts to blind deconvolution on the U-axis with a spatially-varying kernel. We approximate the estimation of the surface indicator function $V_{r\theta}(u)$ using a simple algorithm that requires the constrained solution of a sparse linear system at each r and θ. Our algorithm can be applied to wide-aperture as well as pencil-beam sensors.

5. We verify the effectiveness of our algorithms using real data, acquired in a specially-constructed tank environment with a programmable controller for positioning the sonar sensors.

6. We validate the applicability of our algorithms with respect to real field operations using trials data gathered by an AUV in two different scenarios. The first dataset was acquired in shallow water on known structures with no marine growth while the second was acquired on an offshore field, featuring advanced marine growth and stronger water current.

3 3D reconstruction from SONAR images

To mitigate blur resulting from integration over a wide aperture, we utilise multiple images acquired with overlapping aperture-frusta. Inspired by synthetic aperture imaging, we propose to obtain the images by relocating the source along a direction orthogonal to the range as well as the sensor array. Without loss of generality, we depict this chosen direction as aligned with the vertical Z-axis, as illustrated in fig. 3. We present in this section our two 3D reconstruction approaches as well as a practical denoising method enabling the reconstruction from noisy images such as real SONAR observations.

3.1 Pre-reconstruction denoising

Sonar data being noisy, we propose to apply a denoising step on the input samples before the 3D reconstruction process. This pre-reconstruction step provides noise-free samples, allowing a formulation of the reconstruction problem under a noise-free assumption. In theory, the observed noise in multibeam images is typically a combination of electronic and eventual ambient acoustic noise. In practice, our experiments showed little variation in the noise pattern when changing environments. Since acoustic sensors on AUVs are rarely replaced, the estimation of noise models can be performed once and for all on each sensor in a water tank or during a travel of the AUV through open water. We estimate the parameters of a Gaussian model for noise from observations of an open-water sequence (without any objects to scan) in the same environment that we perform experiments in, i. e. we estimate $\mu(r, \theta)$ and $\sigma(r, \theta)$ across various values of range r and bearing θ. We then apply a simple Z-test using the same threshold for all experiments in this paper so as to reject 95% of the input noise.

3.2 Formulation as a blind deconvolution

We tailored eq. 1 to arrive at our formulation. First, we introduce the Z-coordinate z_i of the sensor array. We also assume that the SONAR wavefronts arriving at the scanned geometry are close to planar (far-field assumption (Aykin and Negahdaripour, 2013a)). Under this assumption, the integration domain is approximated to be parallel to the Z-axis. Finally, we replace the cosine term with a general reflectance function ρ (which subsumes the cosine). Thus, eq. 1 transforms into

$$I(z,r,\theta) \quad = \quad \int_{z-\Delta z_r}^{z+\Delta z_r} \beta_r(u-z)\ V(u,r,\theta)\ \rho(z,u,\vec{n}_{r\theta u})\ \mathrm{d}u$$

where $\Delta z_r \approx r \tan \frac{\phi_2-\phi_1}{2}$, β_r is the warped beam pattern over the wavefront at distance r, ρ is the reflectance function evaluated for retro-reflection from the source to the point at u where the surface has a normal $\vec{n}_{r\theta u}$ and $V(u,r,\theta)$ is a binary function that is unity if there is a surface at a distance of r from the source, along the bearing θ and offset by u in the Z-direction and zero otherwise. Although the coordinate system appears confusing due to a mixture of cylindrical and Cartesian coordinates, recall that r and θ are common to all images taken by relocating the source along the Z-axis. z represents the height at which the source is placed and u is the height at which the reflector is located.

The variation in the measured images at some fixed pixel, $I_{r\theta}$ as a function of the source position z is

$$I_{r\theta}(z) \quad = \quad \int_{z-\Delta z_r}^{z+\Delta z_r} \beta_r(u-z)\ V_{r\theta}(u)\ \rho(z,u,\vec{n}_{r\theta u})\ \mathrm{d}u \tag{2}$$

where $V_{r\theta}(u)$ corresponds to an indicator function that is unity if there is a surface at height u corresponding to the range r from the Z-axis and bearing θ wrt $\vec{v}$ (see the plot in fig. 3). For a given sensor, β_r is known but the following two terms in the integral depend on the geometry (surface indicator function and surface normal respectively).

Eq. 2 suggests that, for a set of SONAR images acquired along the Z direction, the values at each pixel over z represent a convolved version of the surface indicator function. The kernel of the convolution has two components: A fixed component (β_r) and ρ a spatially-varying data-dependent modulating term that depends on the surface normal and reflectance function. If the reflectance distribution is diffuse, ρ simplifies to a cosine term but remains spatially varying (over z). Reconstructing surface points, V, from observations of $I_{r\theta}(z)$ is *therefore challenging and equivalent to blind-deconvolution where the kernel is spatially varying* — even under simplifying assumptions of far-field insonification and Lambertian reflection.

3.2.1 Approximation using a sparse linear system

Let z_i $i = 1, 2, ..., N$ be the Z coordinate of the sensor location for the i^{th} acquisition and let $I(z_i)$ denote the corresponding image. Under the far-field assumption and under ideal conditions (no reverberation or noise), the vector of values $I_{r\theta} = \{I_{r\theta}(z_i)\}_{1 \le i \le N}$ corresponds to measured reflectivities of the surfaces along a line parallel to the Z-axis at a distance of r from it. We call this the u-axis, since it corresponds to the variable of integration u in eq. 2.

Let the u-axis be discretized into P segments. P may be used as a parameter that controls the approximation. For large enough P, we assume that the variation of all three components of the integrand in eq. 2 are negligible within each of these segments and that they are mutually uncorrelated. Then, the integral can be replaced by a summation:

$$I_{r\theta}(z_i) \quad = \quad \sum_{j=1}^{P} \tilde{\beta}_j^i\ \tilde{\rho}_{jr\theta}^i\ \tilde{V}_{jr\theta}. \tag{3}$$

where $\tilde{\beta}^i_j$, $\tilde{\rho}^i_{jr\theta}$ and $\tilde{V}_{jr\theta}$ are expected values of $\tilde{\beta}_r(u)$, $V_{r\theta}(u)$ and $\rho(z, u, \vec{n}_{r\theta u})$ over the j^{th} segment on the u-axis relative to the sensor placed at z_i. Given N images, we can then write

$$I_{r\theta} \quad = \quad A_{N\times P} \,.\, b_{r\theta} \tag{4}$$

where $I_{r\theta}$ contains N observations along the u-axis, A is a matrix with $A_{ij} = \tilde{\beta}^i_j \tilde{\rho}^i_{jr\theta}$, $b_{r\theta}$ is a vector with $b_{r\theta}(j) = \tilde{V}_{jr\theta}$, $j = 1, 2, 3, .., P$, which indicates the presence of a surface in the j^{th} segment along the u-axis. Since the beam pattern is only non-zero for an interval corresponding to the width of the aperture projected onto the u-axis, the matrix A is sparse. Reconstruction of the surfaces is then equivalent to estimating $b_{r\theta}$ given $I_{r\theta}$ and A at each (r, θ).

3.2.2 Deconvolution algorithm

The input to our algorithm is N 2D SONAR images acquired using the same beam direction (bearing), the same x and y coordinates but from depths z_i, $i = 1, 2, ..., N$. From this, we extract our point cloud in three simple steps: denoising (using the same method as detailed in section 3.1), constrained solution of a sparse linear system followed by de-ringing.

Linear solver Since $b_{r\theta}$ is positive, in eq. 4, we solve the linear system using a Non-Negative Least Squares (NNLS) formulation:

$$\underset{b_{r\theta}\geq 0}{\arg\min} \, \|Ab_{r\theta} - I_{r\theta}\|_2 \tag{5}$$

The reconstruction is then obtained with an open-source C++ implementation of a NNLS solver (Lawson and Hanson, 1974), allowing direct and computationally-efficient integration in the simulation and reconstruction framework. The elements in the matrix contain $\tilde{\rho}^i_{jr\theta}$ which are dependent on the reflectance distribution function of the object. Although, under the Lambertian assumptions, these terms would reduce to cosines, their calculation would require knowledge of surface normals. We observed that ignoring them (setting them to unity) yields results that are comparable to more complex treatment (see discussion in sec. 5).

De-ringing Since we solve for the discretized version b of a surface indicator function V across a slice in range (and bearing), the function we aim to reconstruct is typically composed of impulses where there is a surface. Consequently, the reconstruction of its discrete version via the above least squares formulation leads to severe ringing artifacts that cause 'halo' structures. To identify and eliminate these artifacts, we test segments in space where b is non-zero to ensure that they were not observed to be 'free space' in any of the input images. If they were observed to contain no reflectivity in any unoccluded view, we set the corresponding element of b to zero. The result of suppressing the ringing artifacts is a discrete representation, per range-bearing slice, of the geometry. We associate a point with every non-zero element of this representation.

Sparsity promotion In order to constrain the resolution of an underdetermined system, additional constraints on the solution can be added through regularization. A common way of imposing sparsity being the use of L_1 regularization, we include here results obtained adding a regularization term, resulting in the following formulation :

$$\underset{b_{r\theta}\geq 0}{\arg\min} \, \|Ab_{r\theta} - I_{r\theta}\|_2 + \lambda\|b_{r\theta}\|_1 \tag{6}$$

where λ is the regularization parameter aiming at weighting the sparsity term versus the data fidelity term. Our regularized results are obtained using a Matlab implementation of the truncated Newton interior-point method described in (Kim et al., 2007).

3.3 Carving method

As presented in our previous work (Guerneve and Petillot, 2015), the problem of 3D reconstruction from multiple observations can be solved using a non-linear approximation akin to a carving method.

3.3.1 Carving theory

Recalling the sonar imaging model in equation 1, one can consider vertical subdivisions in p uniform sections Φ_i of size $\Delta_\phi = (\phi_2 - \phi_1)/p$ such that: $[\phi_1, \phi_2] = \bigcup_{i=1}^{p} \Phi_i = [\phi_1, \phi_1 + \Delta_\phi] \cup ... \cup [\phi_1 + (p-1)\Delta_\phi, \phi_2]$ and

$$I_s(r,\theta) \quad = \quad \sum_{i=1}^{p} \int_{\Phi_i} \beta(\phi) \;\; f(r,\theta,\phi) d\phi = \sum_{i=1}^{p} \sigma_i(r,\phi) \tag{7}$$

with $f(r,\theta,\phi) = V_s(r,\theta,\phi) \;\; \frac{\vec{v}\cdot\vec{n}_{r\theta\phi}}{\|\vec{v}\| \; \|\vec{n}_{r\theta\phi}\|}$. Due to the positivity of f, each subsection integral $\sigma_i(r,\phi)$ is positive and for small enough subdivision Φ_i, one can assume (Riemann quadrature):

$$\sigma_i(r,\phi) \approx K_{\Phi_i} \; \beta(\phi_i) \; f(r,\theta,\phi)$$

with ϕ_i being Φ_i midpoint and K_{Φ_i} being a measure of the size of the integration domain Φ_i. For simplicity and since we aim at estimating voxels of equal sizes, we consider, without loss of generality K_{Φ_i} as a constant equal to 1. From this approximation and given a single measurement $I_s(r,\theta)$, results the following inequalities providing upper boundaries to f on each of the p subdivisions:

$$\forall i \in [\![1,p]\!]; \quad 0 \leq f(r,\theta,\phi_i) \leq \tfrac{I_s(r,\theta)}{\beta(\phi_i)} \tag{8}$$

Therefore in the case of N observations $\{I_k\}_{1\leq k\leq N}$ of the same point P from different elevation angles $\{\phi_k\}_{1\leq k\leq N}$, the upper boundary of the reflected intensity is given by the observation of lowest intensity:

$$0 \leq f(P) \leq \min_{1\leq k\leq N} \frac{I_k}{\beta(\phi_k)} \tag{9}$$

From this simple min-filtering rule results a non-linear behaviour akin to a carving behaviour where the observation of empty spaces (low intensities) decreases the upper boundaries of potential scatterers located within the vertical aperture of the sensor.

3.3.2 Carving algorithm

As presented in fig. 4, an online version of the carving algorithm is composed of three main steps: a 2D to 3D spherical expansion of the sonar image is performed to account for the sensor 3D footprint. Importantly, the intensities are scaled by the beam pattern value following eq. 8. Once this 3D estimation of the sonar image is obtained, a rescaling is operated. Due to the spherical imaging model of a 2D SONAR, the 3D scene is sampled anisotropically. In order to yield memory-efficient storage of the map and enable further 3D processing at sensor resolution, each measurement is scaled anisotropically: no rescaling is applied in the direction of highest resolution while the two other components are scaled down to match the voxel size of the direction of highest resolution. The resulting 3D representation is then stored as a regular grid in an Octree. Each time a new image is available, the current map is updated with its expanded 3D representation using a nearest neighbour gridding approach. Following eq. 9, when multiple observations are available (when the current image 3D footprint overlaps with the current map), the minimum intensity over all the observations at this point is kept and stored in the map. Finally and because the occlusions cannot be estimated from single images when performing the 2D to 3D spherical expansion, an occlusion resolution step is performed once all the observations have been processed. This occlusion resolution step aims at selecting in the map only the 3D points representing the front surface of the objects and removing occluded 3D points such as inner parts of the objects. The 3D footprint of each input image is projected on the map and only 3D points that are observed directly in at least one image are kept.

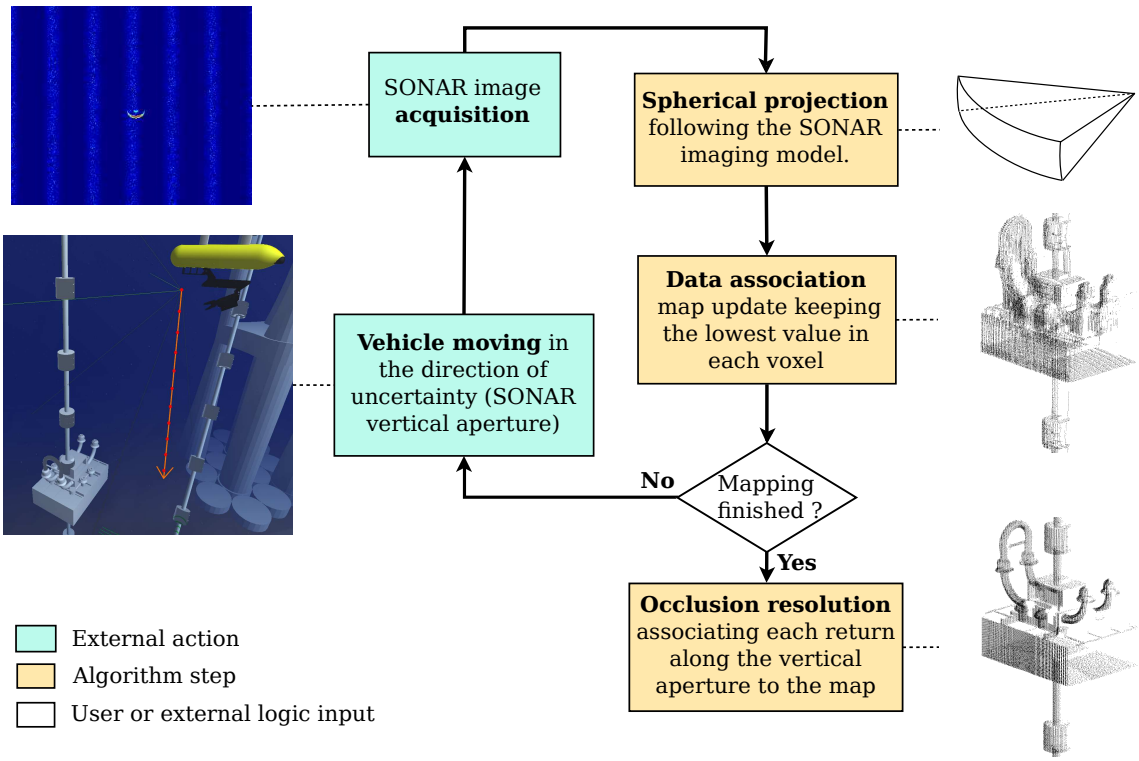


Figure 4: Online carving algorithm diagram. Following the acquisition of data from a new point of view, the 2D image is expanded to its 3D footprint representation with intensities scaled by the beam pattern. The 3D map of the scene is then updated with each new observation until the whole scene has been covered. The last step operates an occlusion resolution to retain only the front surface of the observed scene and generate the final map.

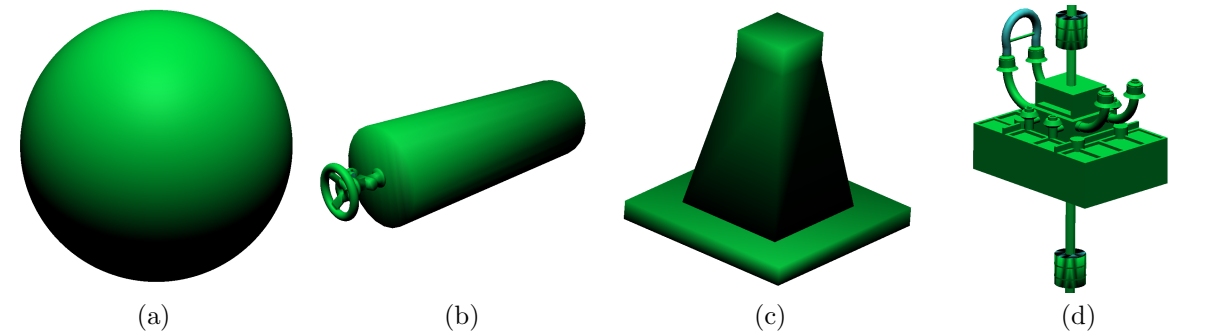
4 Results

We tested each of our two algorithms (deconvolution and carving) on two aperture settings that are typically used by sensor manufacturers (1° and 10° apertures). We performed each test on a set of four test 3D models (fig. 5.a-d) using simulated data and a different set of four objects using measured SONAR data (fig. 5.e-h). We present snapshots of the point clouds reconstructed using our two algorithms.

4.1 Experiments

Simulated data We ran simulations using the open-source underwater simulator UWSim (Prats et al., 2012). We implemented realistic sonar simulation by including the imaging process characteristics. The 3D geometries were embedded into the simulation environment and the virtual vehicle (with virtual SONAR sensors) was used to sense the scene from a distance of 2m. The images generated by the simulator were of 2cm range resolution and 0.2° bearing resolution. For each model, two sets of noise-free data were generated via simulation of SONAR sensors with 1° and 10° vertical aperture respectively. The results of reconstruction performed on simulated data are summarised in fig. 7. The figure shows the 3D point clouds (black points on white background) obtained using different combinations of models (columns), algorithms and types of sensor (rows).

Tank data We gathered real data using two different SONAR heads to scan objects that were submerged in a concrete water tank. A 2D plotter mechanism, with a programmable interface for controlling its motion,



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 5: Reference models. The top row shows 3D models used in simulation and the bottom row shows real objects in our tank. a) Buoy b) Bottle and valve c) Bridge pillar d) Oil field riser pipeline e) Aluminium sphere f) cylinder g) parallelepiped and ROV shell h) Hyball ROV

was used (see fig. 6) to accurately position the sonar head. We collected data on a variety of objects immersed in the tank. For each object, we repeated the experiment with two different sensors mounted on the plotter head: a BlueView MB2250 featuring a small (1°) vertical aperture and an ARIS Explorer 3000 offering a large aperture (14°). In each case, the sonar was mounted in a downward-looking configuration enabling the acquisition of vertical slices of objects placed on the bottom of the tank. For each sensor, a sequence with no object placed in the tank was recorded to provide material for the background noise estimation. In order to provide significant overlap between successive images footprints, the sampling resolution used was 4mm for the BlueView and 1.2cm for the ARIS. As shown in fig. 5.e-g, the objects consisted of an aluminium sphere of radius 15cm, a metal cylinder with 10cm radius and 1.5m long, a rectangular parallelepiped ($1 \times .2 \times .2$m) on which an ROV shell was placed irregularly and a Hyball ROV ($1 \times .5 \times .5$m). The results of reconstruction performed on real data are summarised in fig. 8. The figure shows the 3D point clouds obtained using different combinations of models (columns), algorithms and sensor used (rows). As presented in fig. 14, the addition of L_1 regularization to the deconvolution process leads to denser surface reconstructions. In practice, our experiments showed that the choice of the regularization parameter λ is not straight forward and appears to depend on the level of noise in the reconstructed sequence. We therefore chose to present all our experimental deconvolution results obtained with un-regularized non-negative least square optimization, including some results obtained with regularization (see fig. 14) as a comparison.

Field data We applied our reconstruction techniques on field data gathered by Subsea 7 AIV (Autonomous Inspection Vehicle) prototype during validation trials in Loch Eil. The vehicle was following a vertical movement providing a sequence of horizontal slices recorded at different altitudes. Although non strictly uniform, the sampling period was around 4cm at a standoff distance to the structure of about 2m. The images were acquired using a BlueView P900-130 featuring a wide vertical aperture (20°) and a large horizontal field of view (130°). The SONAR head was mounted in a forward-looking configuration with a small downward pitch angle of about 5°. The vehicle position was provided by a navigation module integrating readings from multiple sensors. Bottom lock velocities were provided by a DVL (Doppler Velocity Log), depth was read from a pressure sensor while orientation estimation was obtained using a compass and a gyroscope. On the

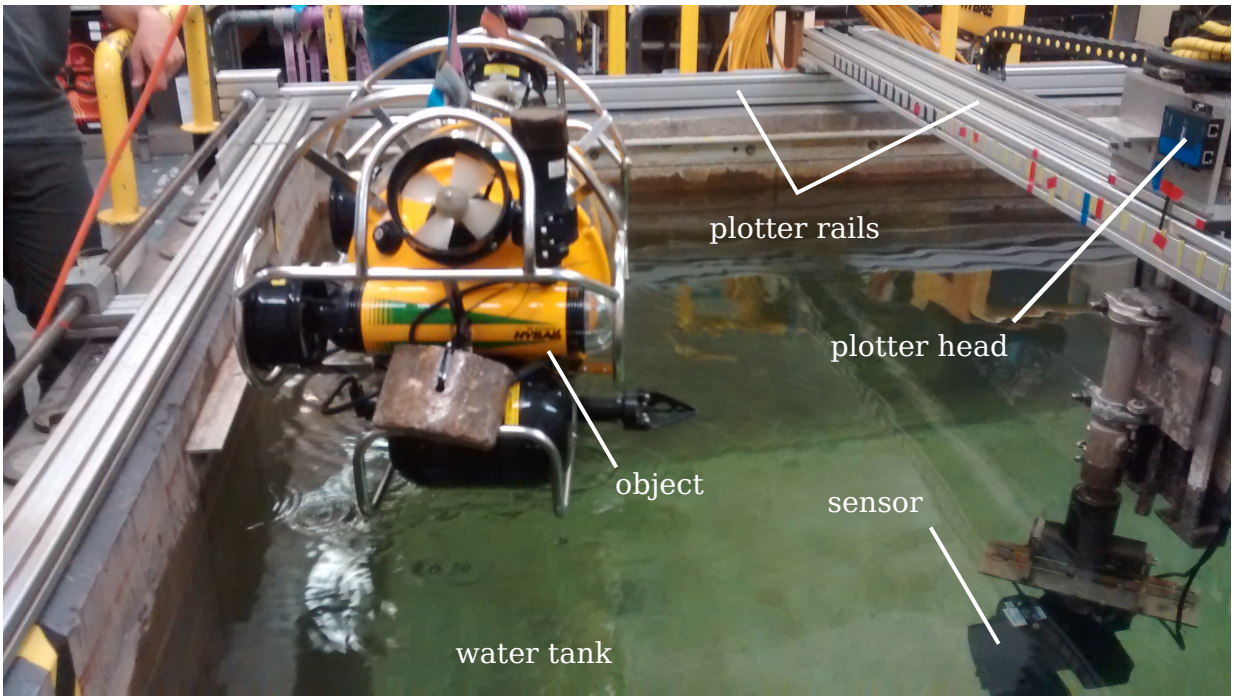


Figure 6: Our experimental setup for gathering data from sensors. A programmable plotter head is moved on guide rails to scan a water tank. A BlueView MB2250 is shown mounted on the head and an object (Hyball ROV) is about to be submerged for scanning.

site of the trials, a few metallic structures were placed on the seabed, effectively simulating a real oil field. The first inspected structure is a 3 meters large square resembling a commonly used structure in the oil industry for which a small number of observation was acquired ($N = 30$). In addition to this and using the same vehicle and experimental conditions but in a different location, we present reconstruction results from real oil-field inspection data. This structure is a 6m large and 8m tall polygonal structure with advanced marine growth inspected vertically from 5 lateral positions around the structure ($N = 5 \times 200$). The last experiment consisted in a top inspection of the first structure using a downward-looking pencil-beam SONAR (BlueView MB2250) with a sampling period of 3cm ($N = 100$). In these conditions and due to the small aperture of the sensor (1°), the overlap between successive samples is limited resulting in very similar results with both reconstruction methods. The results of our reconstructions obtained using our field dataset are shown in fig. 9. The figure shows reconstructed results obtained with both the deconvolution (b,e,h) and the carving approach (c,f,i). As a reference, rough CAD models of the inspected structures are provided (a,d,g).

4.2 Quantifying error and coverage

Simulation results The 3D models used in simulation were used for a quantitative comparison of the performance of our two algorithms on the two different sensors. In particular, we compared the error and coverage achieved in each reconstruction. For measuring error, we computed the median value of the distances between the reconstructed points to the finely tesselated surface of the object and call this the *median error value*. For assessing how much of the model's surface was scanned, we define *coverage* as the percentage of points on the reference model (for an absolute and fixed tesselation resolution) that contain at least one reconstructed point within a 1 cm radius.

For each test model, we reconstructed the 3D point cloud using our proposed deconvolution approach and estimated the resulting error and coverage for different sampling rates. Then we plotted (fig. 10 top row)

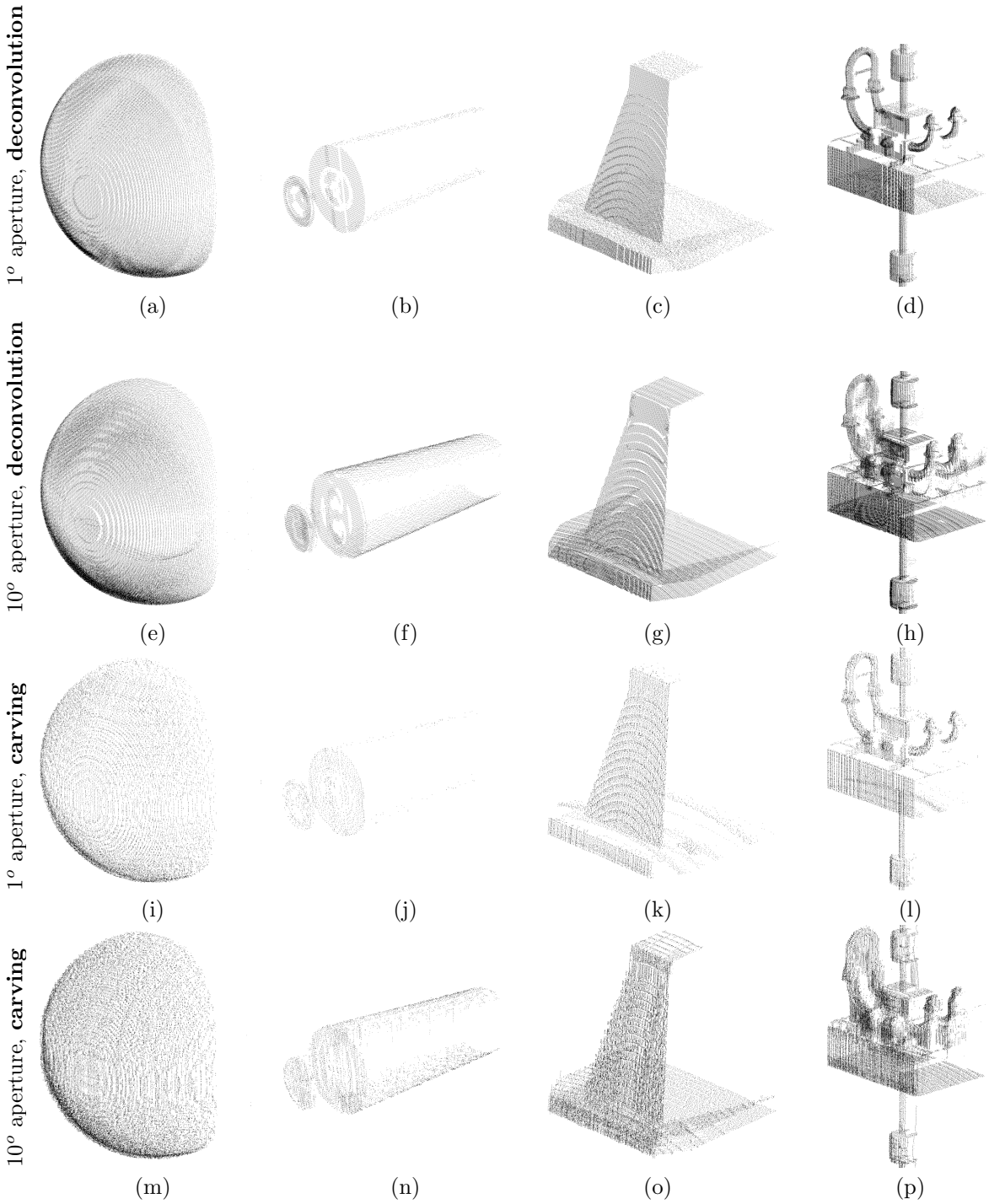


Figure 7: 3D reconstruction results from simulated SONAR images using our two algorithms: deconvolution (a-h) and a fast approximation akin to a carving method using the minimum observed intensity per voxel (i-p). Our algorithms reconstruct point clouds from wide-aperture (e-h and m-p) images that are of comparable quality to those from narrow-aperture scanners (a-d and i-l) and improve coverage of surfaces that are parallel to the scanning beam's view direction. e. g. top of the conical bridge pillar model and the long surface of the cylinder.

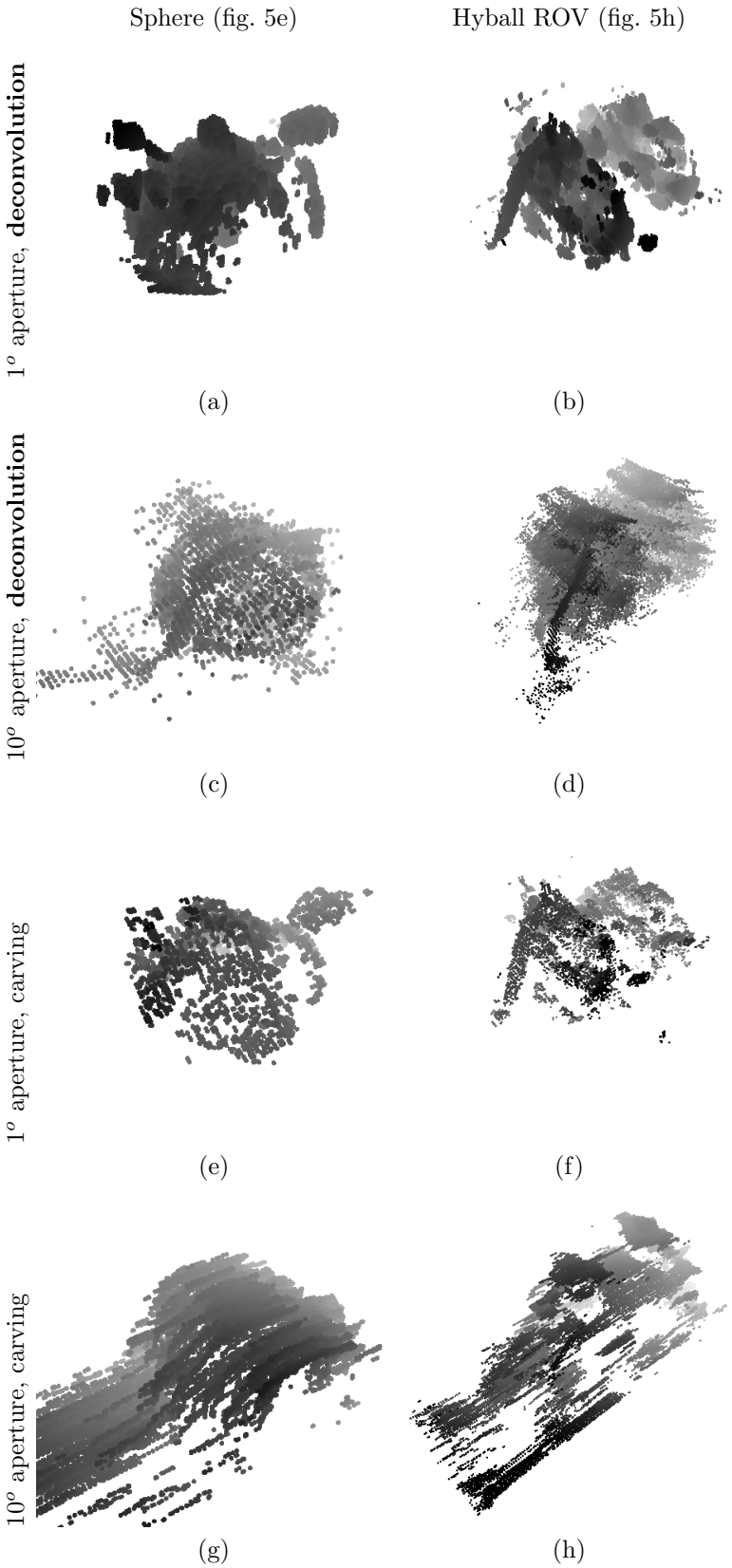


Figure 8: 3D reconstruction results from real experimental measurements in our tank environment. The rope used to secure the object underwater is reconstructed and evident in (c) and (d). Our deconvolution method (top two rows) produces discernible reconstructions in comparison to the carving approach (bottom two rows).

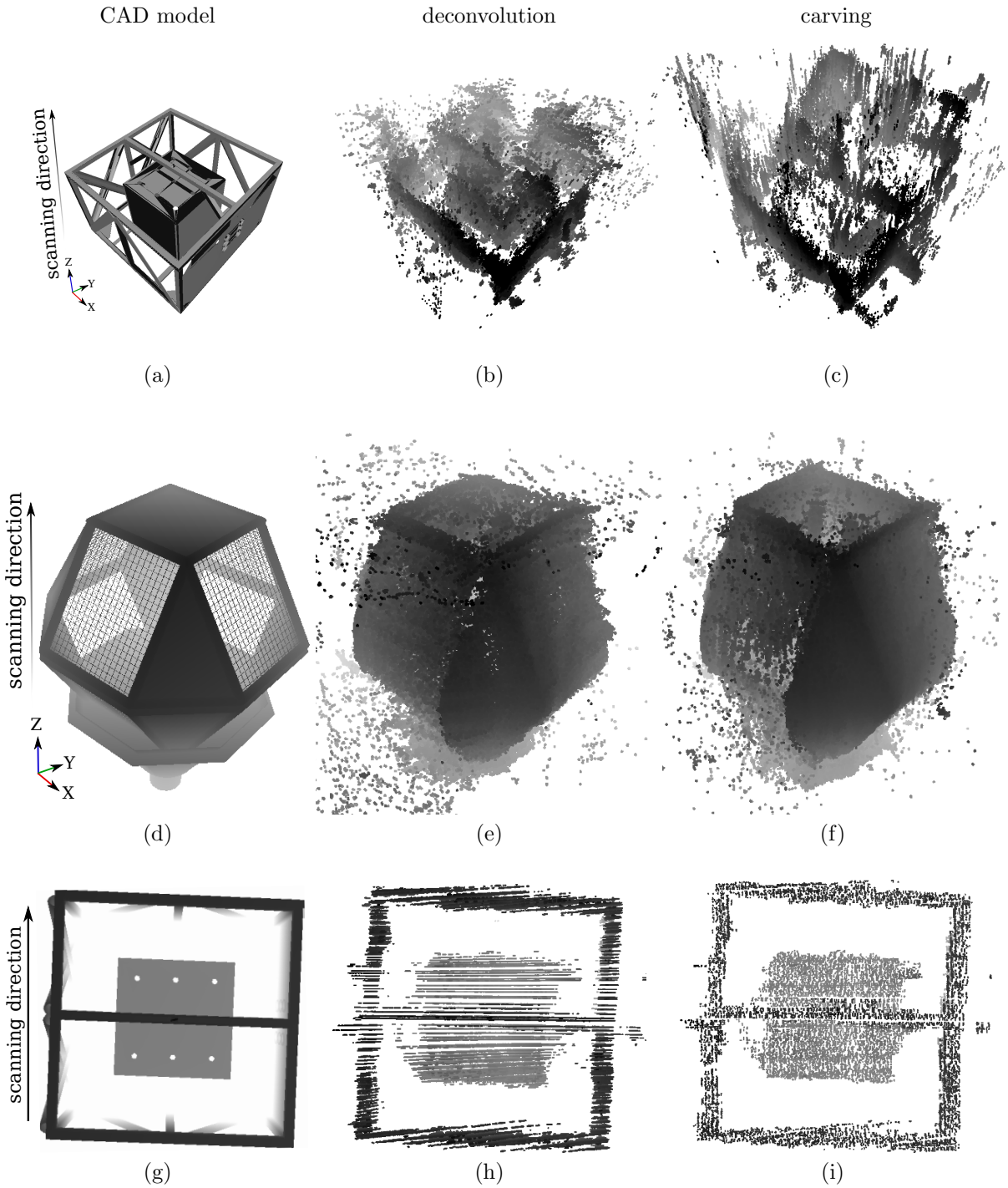


Figure 9: 3D reconstruction results from field experimental measurements during AIV trials. Using a wide-beam SONAR, the point clouds obtained using our deconvolution method (b,e) exhibit more details and structural coherence than the carving method (c,f). In particular, the inner box appears clearer in (b) than in (c), some points appear to be missing on the bottom left part of (f) which is not the case in (e). While both methods suffer from noise artefacts, the outliers from the deconvolution method only tend to appear at the extremities of the reconstruction, where less observations are provided whereas carving reconstructions are more uniformly corrupted by noise, resulting in more difficult outlier filtering. Using a pencil-beam sensor, the reconstructions with the two methods (h,i) appear very similar due to the reduced need for deconvolution (small aperture). CAD models (a,d,g) are included for reference.

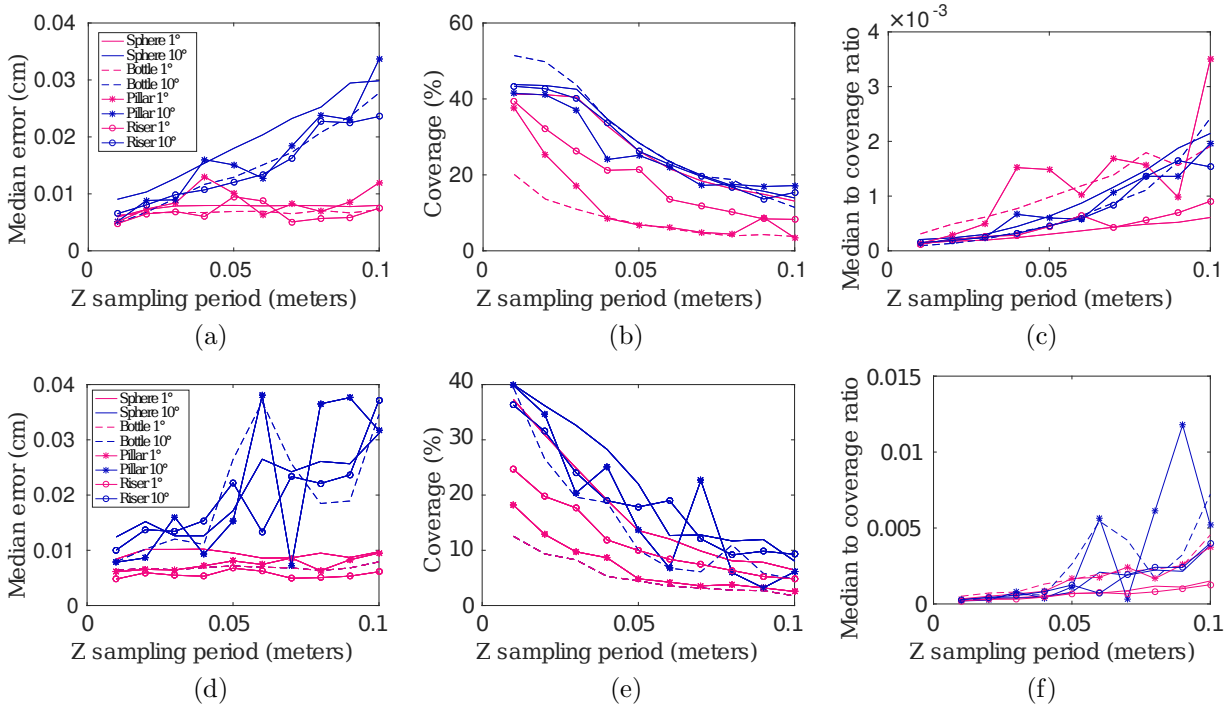


Figure 10: Plots showing median error (first column), coverage (second column) and the ratio median error:coverage (third column). Top row corresponds to data from using the deconvolution approach while the bottom row corresponds to results of using the carving method. Wide-aperture results are plotted in blue, narrow-aperture results are drawn in magenta.

error (first column), coverage (second column) and their ratio (third column) for all the models (see legend) against the sampling resolution in the Z-axis (X-axis of the plots). We generated a similar plot for error and coverage corresponding to the carving algorithm (fig. 10 bottom row).

Field results Similarly to the quantitative analysis performed on simulated data, table 1 presents median errors, coverages and median to coverage ratios of each field reconstruction. Unlike our simulated experiments, the ground truth position of the CAD model in field reconstructions was only known to a limited level of precision, typically ±20cm, ±30°, therefore requiring position estimation through registration. For each test scenario, the associated CAD model (see fig. 9-a,d,g) was carefully registered to the 3D reconstructed point cloud using manual overlaying followed by an ICP-based method (**?**). This method, implemented in CloudCompare (**?**), enabled the registration using point clouds subsets by discarding a given fraction of

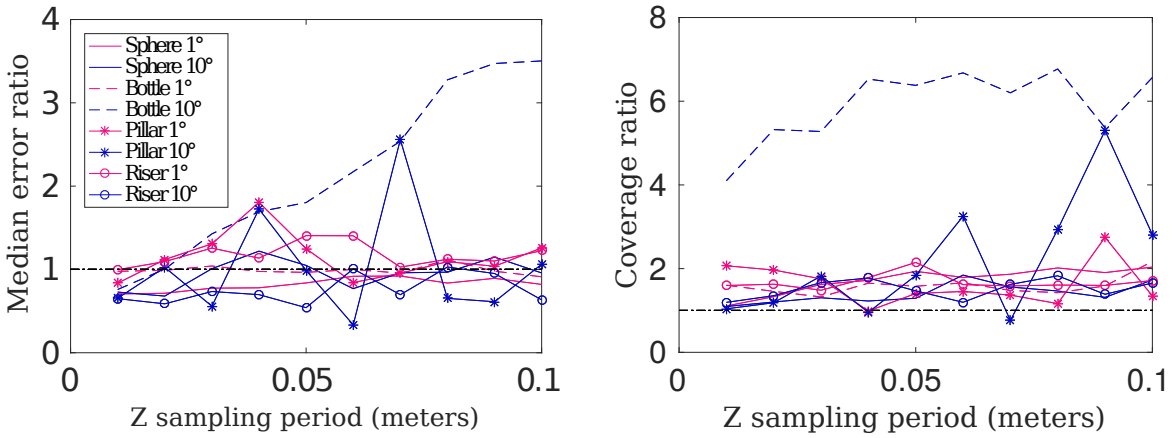
Test scenario	**Sensor aperture**	**Reconstruction method**	**Median error (m)**	**Coverage (%)**	**Median to coverage ratio**
Box structure partial side inspection	20°	deconvolution	**0.067**	**20.68**	**0.00664**
		carving	0.073	20.07	0.00701
Polygonal structure 360° side inspection	20°	deconvolution	0.112	**56.50**	0.00199
		carving	**0.089**	50.79	**0.00176**
Box structure top inspection	1°	deconvolution	0.029	**23.70**	**0.00121**
		carving	**0.020**	12.62	0.00157

Table 1: Quantitative results of field reconstructions. For each scenario, the best metric value is displayed in bold font.

points (outliers). Once the registration performed, the median error and coverage values were computed. In order to account for the limited accuracy of the registration as well as navigation error, the coverage was computed using a larger look-up radius than in simulation: 3cm for the two short inspections (box structure) and 5cm for the complete (360°) side inspection of the polygonal structure.

5 Discussion, comparison and limitations

Reduced memory requirements Sonar carving using the min-filter (Guerneve and Petillot, 2015) operates on a discretization of the 3D volume. Our deconvolution approach performs a discretization in 1D (along the U axis) for each chosen r and θ. Since there is no overlap across different r and θ, we only need to store elements where the indicator function is non-zero. Our algorithm is output-sensitive with respect to memory requirements, allowing scalability of sampling (and hence reconstruction) resolution. This is an important consideration in underwater vehicles with limited on-board memory resources.

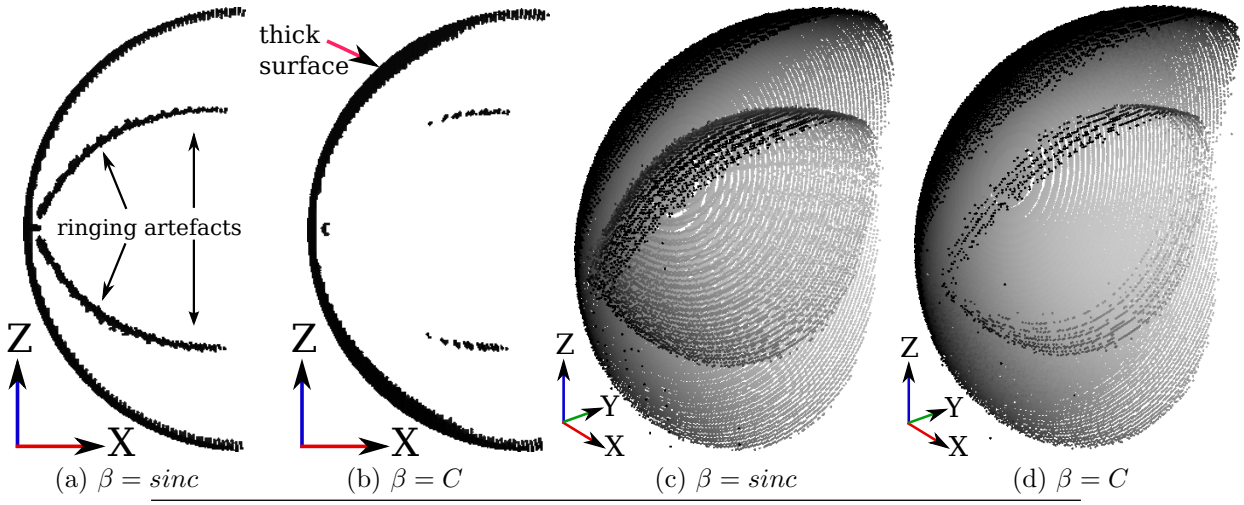


(a) (b)

Figure 11: Comparison of error and coverage between the deconvolution method and the carving approach. (a) ratio of median error using the deconvolution to median error using carving and (b) ratio of coverage using the deconvolution to coverage using sonar carving, for all models (different point markers) using narrow-(magenta) and wide-aperture (blue) sensors. Black, dashed lines are shown for reference at 1.

3D reconstruction error and coverage Most sonar scanning is limited to heightfields (2.5D). Notable exceptions include a few prohibitively expensive direct 3D sonar sensors (Davis and Lugsdin, 2005) and recent methods in sonar carving (Guerneve and Petillot, 2015). However, their errors and coverage have not previously been analysed quantitatively. We perform this quantitative analysis (bottom row of fig. 10) and also compare it to results using our deconvolution method. From the plots of median error, we observe that the deconvolution algorithm produces reconstructions with similar quality for narrow-aperture (magenta curves in fig. 10 a and d) and better quality for wide-aperture simulations (blue curves in fig. 10 a and d). Similarly, we observe that our deconvolution yields improved coverage (blue curves in fig. 10 b and e) over the carving method. Finally, the ratios of error to coverage are lower for all models than that obtained with carving (fig. 10 c and f). To facilitate comparison between our two methods, we plot the ratio of the errors as well as the coverages for the corresponding curves in fig. 10. a, b, d and e. Curves that stay below 1 (for error in fig. 11. a) and above 1 (for coverage) in fig. 11.b are better using the deconvolution approach. This result, visible in fig. 7 is justified by the difficulty to perform occlusion resolution on surfaces observed at grazing angles due to the limited resolution in elevation angle. As opposed to the space carving method, the deconvolution approach does not feature any occlusion resolution but rather selects the *a-posteriori* most likely solution considering the input observations. In spite of the difficulty of performing quantitative analysis

on field data (due to navigation drift, registration error, the presence of multipath effects as well as the limited fidelity of the CAD model due to modelling error and marine growth), the results presented in table 1 confirm the results obtained in simulation. When the navigation drift remains negligible (data acquired over a small distance as in the first and last scenarios), the error and coverage results match the results observed in simulation: the deconvolution provides improved errors and coverages over the carving method when using the wide-aperture sensor (first scenario) and better coverage when using the pencil-beam sensor (last scenario). When considering larger sensor displacements (second test scenario), the median error grows considerably (9 to 11cm) and the difference between the two reconstruction methods becomes less noticeable.



(a) $\beta = sinc$ (b) $\beta = C$ (c) $\beta = sinc$ (d) $\beta = C$

Figure 12: Thresholding the beam pattern to impose sparsity (b,d) produces a better reconstruction in practice than using quantized approximations of the warped beam pattern (a,c). Vertical slices of the reconstructed sphere (a,b) exhibit thicker surface, 3D views of the inside of the sphere (c,d) shows the reduced presence of ringings.

Approximations The function β ideally contains a step profile in the Fourier domain (low-pass filter). In practice the beam patterns for the sensors used are Bessel functions. While these patterns are known a priori, for any given sensor, we observed that they do not play a major role in the approximation. We observed that providing no beam-pattern adjustment produced lower ringing artefacts than when quantized approximations of the beam pattern were used (see fig. 12). Since computational resource and power are precious on unmanned underwater vehicles, we recommend that the term containing the beam-pattern can safely be thresholded to impose sparsity for on-line reconstruction. To accommodate the reflectance term in A_{ij}, we initially performed alternative minimization as is commonly done for blind deconvolution (Chan and Wong, 2000). We alternated between estimating normals and using them to obtain the cosine term that modulates the surface indicator function. We found this approach to be sensitive to the method and parameters chosen for estimating the normals. In addition, we observed that the method did not converge empirically. This points to an interesting direction of future work. We obtained better results with experiments where we simply replaced the cosine term with unity and ignored the surface normal. All results shown in this paper are obtained by thresholding the beam-pattern term and ignoring the cosine term due to surface normals.

Computation A sequential implementation of our algorithm typically takes about 15 seconds for a real dataset consisting of 128 frames on an Intel(R) Core(TM) i7-4700MQ CPU @ 2.40GHz processor with 16GB RAM. In comparison, carving takes about 32 seconds for the same data. Both denoising and deringing are important steps in our pipeline (see fig. 13) but they only require about 1 second of computation each. The least squares sparse solver at the crux of our deconvolution formulation is simple and efficient. Our method is parallelizable, as is carving using the min-filter, since we solve independent systems at each (r, θ) along the U-axis.

(a) plain reconstruction (b) only denoising (c) only deringing

Figure 13: The importance of denoising and deringing. Compare these results with the output when both are applied (fig. 8c).

(a) No regularization ($\lambda = 0$) (b) $\lambda = 0.1\lambda_{max}$ (c) $\lambda = 0.5\lambda_{max}$

Figure 14: The influence of regularization. The use of an L_1 regularization term in the optimization process imposes sparsity: the surface of the aluminium sphere appears to be more dense when giving more importance to regularization (b) and (c) whereas in absence of regularization (a), the reconstruction exhibits a few holes on the surface.

Sparsity promotion While L_1 regularization is commonly employed in optimization as a way to impose sparsity, recent study (Slawski et al., 2013) suggests that applying a non-negativity constraint promotes sparsity when the model matrix A has certain properties which appear to be satisfied in the case of a typical deconvolution problem. In this context, adding further sparsity through regularization is therefore optional and situation-dependent. In particular, sparsity is desired when reconstructing rectangular functions which is only true when observing a surface in a non-orthogonal way: L_1 regularization is therefore not suited for reconstructing a planar section such as a seabed but remains desirable when observing surfaces such as spheres. In any case, the choice of the regularization parameter was made empirically, restricting its use to human-assisted situations.

Limitations We do not address the phenomenon of multipath returns (reverberations). Especially in enclosed environments such as our tank setup, reverberations severely pollute the input signal hampering reconstruction. Similarly multipath phenomena were frequently observed in the first field dataset, producing artificial returns in the image resulting in non-existing 3D points (outliers) in the reconstruction as can be observed in fig. 9. Thanks to their smaller aperture, multipath effects are less frequent in pencil-beam observations. As can be observed in our last field experiment, using a large number of pencil-beam observations and low-drift navigation, the box structure reconstruction exhibits less outliers but lower surface coverage than its wide-beam equivalent (first field experiment). In addition to multipath effects, the reconstruction accuracy of field structures suffers from additional perturbations which include some natural navigation drift, suboptimal positioning of the sensor (small pitch angle), non-uniform sampling period and small yaw angle variation ($\pm 2°$). While these elements are expected to affect the reconstruction results, most of them will be addressed in future work to allow for less constrained inspection paths (non-rectilinear and non-uniform sampling) as is often the case when operating AUVs. In spite of these limitations, our reconstruction results exhibit enough detail for human-eye structure recognition, showing promising results towards automated underwater scene understanding from standard acoustic sensors.

Online reconstruction Our two reconstruction methods feature complementary pros and cons. The carving approach is an iterative reconstruction method where each new sample is used to update the map with the possibility to use the map at each time as an estimate of the occupancy around the vehicle. Due to its simple processing, the computation usage remains moderate but its memory usage remains high, requiring a fragmentation of the map into medium-size patches (of around 50m^2). On the other hand, the deconvolution method achieves reconstruction with an increased level of detail but requires (in its current implementation) rectilinear motion and uniform sampling. In comparison to the carving method, its memory usage is lower but the reconstruction can not be made iteratively, since all the samples need to be fed into the optimizer at the same time. A reconstruction per patch is therefore mandatory too. Both methods are applicable in an online configuration, providing a 3D representation of the world in an operation time.

6 Conclusion and future work

In this paper we have presented new approaches to reconstructing underwater objects using imaging SONAR sensors of any vertical aperture. We have proposed two ways of formulating the problem: as a spatially-varying blind deconvolution, and using a non-linear approach akin to space carving. We showed that the use of simple approximations enabled to solve the a-priori ill-posed problem of 3D reconstruction. Quantitative analysis of our reconstruction method was compared to state of the art carving method. We reported a gain in accuracy with both narrow and wide-aperture sensors with significantly improved rates of sample coverage. We stressed that wide-aperture SONARs, although rarely used in this context, can offer better overall reconstruction results when considering their higher coverage properties, effectively eliminating the need for embedding multiple acoustic sensors of different apertures. We presented results obtained with both simulated data and real data. Tests performed using real field data further validated our algorithm to be beneficial in the context of an inspection carried with an AUV.

Our current formulation is restricted to rectilinear and uniform sampling. In future work, we expect to

focus on extending this method to less-constrained trajectories, rotations and lateral movements. Another prospective direction will be to incorporate uncertainty in sampling locations within the model, to account for inaccurate navigation of the vehicle.

Acknowledgments

The authors would like to thank Subsea 7 for providing access to the AIV data. This project was funded under the Marie Curie ITN program Robocademy FP7-PEOPLE-2013-ITN-608096. Kartic Subr was supported by the Royal Society's University Research Fellowship.

References

Aykin, M. D. and Negahdaripour, S. (2013a). Forward-look 2d sonar image formation and 3d reconstruction. *Oceans - San Diego, 2013*.

Aykin, M. D. and Negahdaripour, S. (2013b). On feature matching and image registration for two-dimensional forward-scan sonar imaging. *Journal of Field Robotics*, 30(4):602–623.

Aykin, M. D. and Negahdaripour, S. (2015). On 3-d target reconstruction from multiple 2-d forward-scan sonar views. In *OCEANS 2015-Genova*, pages 1–10. IEEE.

Babaee, M. and Negahdaripour, S. (2015). 3-d object modeling from 2-d occluding contour correspondences by opti-acoustic stereo imaging. *Computer Vision and Image Understanding*, 132:56–74.

Callow, H. J. (2003). Signal processing for synthetic aperture sonar image enhancement.

Chan, T. F. and Wong, C.-K. (2000). Convergence of the alternating minimization algorithm for blind deconvolution. *Linear Algebra and its Applications*, 316(1):259–285.

Coiras, E. and Groen, J. (2009). 3d target shape from sas images based on a deformable mesh. *Proceedings of the 3rd International Conference on Underwater Acoustic Measurements (UAM)*.

Davis, A. and Lugsdin, A. (2005). High speed underwater inspection for port and harbour security using coda echoscope 3d sonar. *OCEANS, 2005. Proceedings of MTS/IEEE*.

DeMarco, K. J., West, M. E., and Howard, A. M. (2015). A computationally-efficient 2d imaging sonar model for underwater robotics simulations in gazebo. In *OCEANS'15 MTS/IEEE Washington*, pages 1–7. IEEE.

Enrique Coiras, Y. P. and Lane, D. M. (2007). Multiresolution 3-d reconstruction from side-scan sonar images. *IEEE Transactions on image processing, vol 16*.

Fairfield, N., Kantor, G., and Wettergreen, D. (2007). Real-time slam with octree evidence grids for exploration in underwater tunnels. *Journal of Field Robotics*, 24(1-2):03–21.

Georgios Papadopoulos, Hanna Kumiawati, A. S. B. M. S. L. l. W. and Patrikalakis, N. M. (2011). 3d-surface reconstruction for partially submerged marine structures using an autonomous surface vehicle. *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

Guerneve, T. and Petillot, Y. (2015). Underwater 3d reconstruction using blueview imaging sonar. In *OCEANS 2015-Genova*, pages 1–7. IEEE.

Hawkins, D. W. (1996). Synthetic aperture imaging algorithms: with application to wide bandwidth sonar.

Huang, T. A. and Kaess, M. (2015). Towards acoustic structure from motion for imaging sonar. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 758–765. IEEE.

J. Folkesson, J. Leonard, J. L. and Williams, R. (2007). Feature tracking for underwater navigation using sonar. *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems San Diego*.

Jouvencel, B., Xiang, X., Zhang, L., and Fang, Z. (2012). 3d reconstruction of seabed surface through sonar data of auvs. *Indian Journal of Geo-Marine Sciences (IJMS)*, 41(6):509–515.

K. Bikonis, A. S. and Moszynski, M. (2008). Computer vision techniques applied for reconstruction of seafloor 3d images from side scan and synthetic aperture sonars data. *Acoustics 08 Paris*.

Kim, S., Koh, K., Lustig, M., Boyd, S., and Gorinevsky, D. (2007). A method for large-scale l1-regularised least squares problems with applications in signal processing and statistics. *IEEE Journal on Selected Topics in Signal Processing*, 1:606–617.

Lawson, C. L. and Hanson, R. J. (1974). *Solving least squares problems*, volume 161. SIAM.

Massot-Campos, M. and Oliver-Codina, G. (2015). Optical sensors and methods for underwater 3d reconstruction. *Sensors*, 15(12):31525–31557.

N. Brahim, D. Gueriot, S. D. and Solaiman, B. (2011). 3d reconstruction of underwater scenes using didson acoustic sonar image sequences through evolutionary algorithms. *OCEANS, 2011 IEEE - Spain*.

N. Hurtos, S. Nagappa, N. P. and Salvi, J. (2014). Real-time mosaicing with two-dimensional forward-looking sonar. *2014 IEEE International Conference on Robotics and Automation (ICRA)*.

Ozog, P., Troni, G., Kaess, M., Eustice, R. M., and Johnson-Roberson, M. (2015). Building 3d mosaics from an autonomous underwater vehicle, doppler velocity log, and 2d imaging sonar. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 1137–1143. IEEE.

Prats, M., Pérez, J., Fernández, J. J., and Sanz, P. J. (2012). An open source tool for simulation and supervision of underwater intervention missions. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 2577–2582. IEEE.

Slawski, M., Hein, M., et al. (2013). Non-negative least squares for high-dimensional linear models: Consistency and sparse recovery without regularization. *Electronic Journal of Statistics*, 7:3004–3056.

Y. Petillot, I. T. R. and Lane, D. (2001). Underwater vehicle obstacle avoidance and path planning using a multi beam forward looking sonar. *IEEE Journal of Oceanic Engineering*, 26:240–251.

Zerr, B. and Stage, B. (1996). Three-dimensional reconstruction of underwater objects from a sequence of sonar images. *Proceedings of International Conference on Image Processing*, 3:927–930.