AUGMENTED REALITY BY INTEGRATING MULTIPLE SENSORY MODALITIES FOR UNDERWATER SCENE UNDERSTANDING

Vittorio Murino and Andrea Fusiello
Dipartimento Scientifico e Tecnologico, University of Verona
Ca' Vignal 2, Strada Le Grazie, 37134 Verona, Italy
Email: {murino,fusiello}@sci.univr.it

Abstract This paper proposes a method for the integration of acoustic and optical data to enhance the perception of an underwater environment in teleoperation tasks. Off-shore applications are addressed, in which an underwater remotely operated vehicle is approaching an oil rig for inspection, maintenance and repairing tasks. A technique is presented which takes advantage of optical features to segment an acoustic three-dimensional image. Cylindrical surfaces are than extracted from 3-D points, and whole cylinders are reconstructed. The final step is to present useful information to the human operator, by displaying the superposition of measured acoustic data and geometric primitives fitted to parts of it, i.e. an augmented reality view. Experimental results are reported showing the effectiveness of the proposed approach.

1. INTRODUCTION

This paper is devoted to the construction of an augmented reality view that can help a human operator of an underwater remotely operated vehicle (ROV) to better perceive and understand the surrounding environment. Two sensing channels are available, optical and acoustic. The former gives an image easier to read by a human, but visibility is very limited due to low illumination and clutter presence. On the other hand, acoustic data are not affected by illumination problems and provide inherently 3-D information, but are more complicated to understand for a human operator. From these considerations it arises the need to integrate, whenever possible, the two channels in order to exploit the

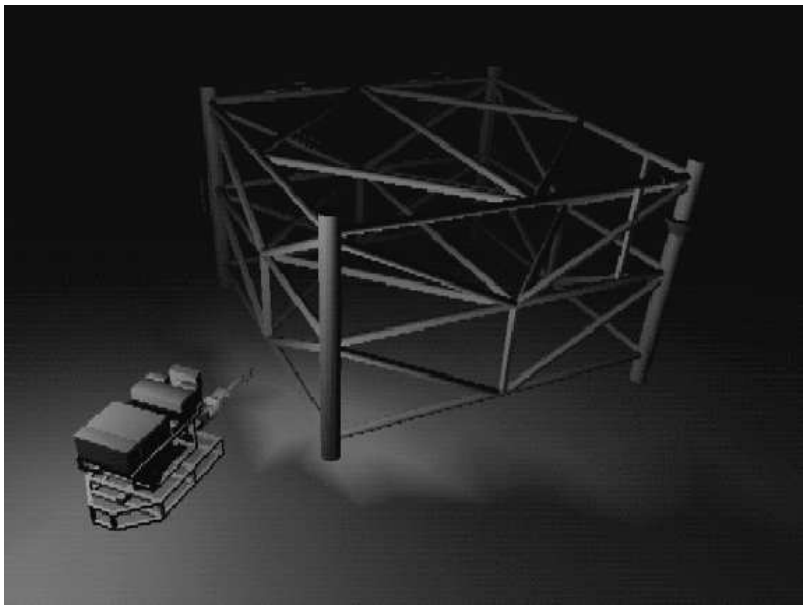

Figure 1.1 Rendering of the VRML model of the oil rig with the ROV.

best of both, so as to compensate their lacks. Moreover, presenting a synthetic model of the scene superimposed on data, i.e., generating an augmented reality image, is much more useful and readable for a human operator.

The applicative scenario consists in an ROV approaching an oil rig made up of connected pipes (see Fig. 1.1). The ROV is equipped with an optical and an acoustic camera. The optical camera provides classical gray-levels images and the acoustic one provides (acoustic) intensity images associated with a set of 3-D points [17].

A virtual reality view is obtained by displaying the superposition of measured 3-D points and of a synthetic model automatically constructed from data. The key issue for automatic modeling is the segmentation of the range data into subsets of points corresponding to the desired primitives, cylinders in our case. Due to the noisy and low-density nature of the acoustic range data, segmentation using differential geometry or step-edges is infeasible. Therefore we propose to integrate acoustic and optical images, and to use optical edges to segment the acoustic data.

First we extract pipes boundaries in the optical image. Then, assuming that the mutual position of the two cameras is known, acoustic (3-D)

data points are projected onto to the image plane. The points falling inside pipes boundaries are segmented. Pipe direction and radius are then estimated and a synthetic model is generated by fitting cylinders.

Fusion and integration of different kinds of data is actually a matter of active research. When available information is of different nature, probabilistic, heuristic, or fuzzy methods are typically used [15, 3]. In case of visual data, a straightforward approach consists in recovering symbolic information separately from the several types of data, and then perform data fusion at the highest (symbolic) level. Classical symbolic Artificial Intelligence technique [1] are applied in this case. In our case, due to the similarity of the data at hand (they are both images), we would like to integrate them at a lower processing level, possibly to improve or facilitate the recognition procedure on either sensorial channel.

Some works are present in literature about data fusion and integration of the different sensors functionalities. Among these ones, some interesting papers can be considered concerning the fusion of intensity and range data, mainly derived by a laser range finder [10, 23, 22].

In [10], a Markov Random Field (MRF) model is proposed for the fusion of registered range and intensity images aimed at image segmentation. An extended weak membrane model is utilized as prior knowledge devoted to enforce the line process, so improving edge detection. The fusion occurred by means of a coupled term in the energy function that penalized different edge configurations in the two kinds of images. A similar method for the fusion of range and intensity images was followed in [23] by integrating in a single framework edge detection, semantic labeling and surface reconstruction. Initial edge labeling and classification is based on a physical analysis of the local behavior of intensity and range data. Then, an MRF model is used to relax the edge configuration while performing concurrently the reconstruction of the surfaces. In [22], an intensity-guided range sensing technique is presented.

Concerning specifically on 3-D scene modeling, there are several work on robotic applications mainly devoted to decontamination and decommissioning tasks in hazardous environments [16]. The closest to our work are [9, 12, 13]. In [9], segmentation of range data in pipes and torii is proposed by using a procedure estimating local centers of curvature. Locally fitting a bi-quadratic function the locus of centers of curvature is estimated using a robust least squares method. Then, these centers of curvatures are used to discriminate between straight and curved cylinders, so allowing the accurate reconstruction of these parts for CAD modeling. A-priori information is utilized to set some algorithms' parameters in order to increase the precision of the segmentation. In [12], quadric surface parameters (representing cylinders) are used to estimate

radius, axis and position, so that the resulting cylinders are displayed to an operator, without performing an actual recognition phase. Generalized cylinders are fitted to range data in [13]. The extraction of axis points is done by computing midpoints between two contour points, then the axis curve is represented as a third degree polynomial

In our work the acoustic range data and the optical intensity image are used in a cooperative way to extract useful (topological and geometrical) information allowing the construction of a virtual environment. Our goal is to automatically model significant objects present in a cluttered scene and facilitate human interpretation by displaying such objects in an augmented reality view. In another work [6] we deal with the similar problem of fitting (a portion of) a known model of the rig to the sensed data.

The rest of the paper is organized as follows. Section 2 describes the processing of intensity and range data and their integration, in order to obtain a segmentation of range data. In Section 3 the extraction of cylinders from range data is outlined and in Section 4 the augmented reality view is obtained. Finally, Section 5 shows some results of the method applied on real data and, in Section 6, conclusions are drawn.

2. SEGMENTATION

The first processing step consist in filtering and segmenting both acoustic and optical data.

2.1 ACOUSTIC DATA PROCESSING

Three-dimensional data are obtained by a high resolution acoustic camera, the *Echoscope* [11]. The scene is insonified by a high-frequency acoustic pulse and a two-dimensional array of transducer gathers the backscattered signals. The whole set of raw signals is then processed in order to enhance those coming from fixed steering directions (called *beamsignals*) and to attenuate those coming from other directions. The distance of a 3-D point can be measured by detecting the time instant at which the maximum peak occurs in the beamsignal (see Fig 1.2. A range image is formed by 64×64 points ordered according to an angular relation, as adjacent points correspond to adjacent beamsignals. Moreover, the intensity of the maximum peak can be used to generate another image, representing the reliability of the associated 3-D measures: the higher the intensity, the safer the associated measure.

The acoustic image is affected by false reflections, caused by secondary lobes, and by acquisition noise, which is modeled as speckle noise. The intensity image turns out to provide very useful information to discrim-

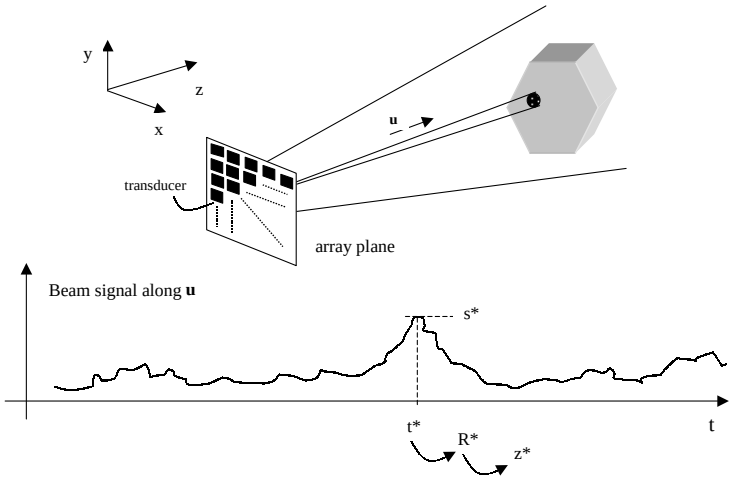

Figure 1.2 Acoustic camera.

inate between "good data" and noise. A dramatic improvement of the image quality is obtained by discarding points whose associated intensity is lower than a threshold. Then, the connected components are extracted by a percolation technique: a sphere of radius R is drawn around each point, and two points are considered to be connected if their spheres intersect. Finally, a size filter eliminates the small blobs caused by noise and clutters. The radius R, the threshold on the intensity, and the threshold on the blob size are chosen basing on *a priori* knowledge of the spatial resolution and directivity characteristics of the sensor [8].

2.2 OPTICAL DATA PROCESSING

The image, obtained by a conventional optical camera is first filtered with an edge preserving anisotropic smoothing [18], that is a smoothing operator whose strength depends on the local gray-level gradient.

Straight lines are extracted by combining Canny's edge detector [5] and Burn's *Plane Fit Algorithm* [4]. First edge points are extracted with the Canny edge detector, that allows to find very sharp edges (often one pixel large) thanks to the non-maxima suppression. Then, pixels are clustered in support regions if they are spatially adjacent and if their gradient orientation is roughly the same. The line parameters are computed with intersections of the weighted fit plane to the intensity values and the horizontal average pixel intensity plane, within a support region. The weight favors intensity values of pixels with high gradient

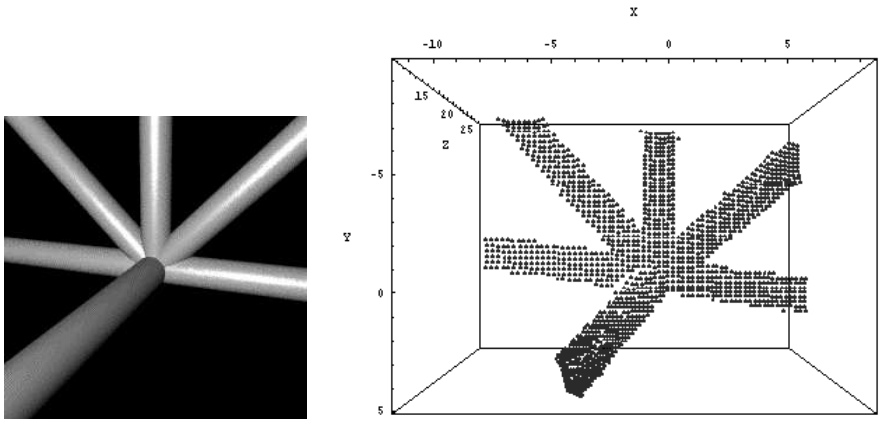

Figure 1.3 Synthetic optical (left) and acoustic (right) images of a joint.

magnitude. Taking mainly the gradient orientation as evidence for a line and using the plane fit method, the algorithm extracts long, straight lines as well as shorter lines and is effective in finding low-contrast lines.

Each extracted segment is then labeled, and its attributes (midpoint, length, etc ...) are computed. In order to find pipes in the image, pairs of segments are grouped together, which are possibly the projection of the boundaries of a pipe. Grouping is based on proximity and parallelism criteria: two segments are paired if the distance between their midpoints is less than a threshold (that is related to the expected distance of pipes boundaries in the image), and if their angle is in the range 180 ± 30 deg. Finally, the convex hulls of all the paired segments are computed.

2.3 INTEGRATION

Optical and acoustic data are integrated by projecting 3-D points obtained by the acoustic camera onto the image plane of the optical camera. Points falling inside a convex hull are deemed to belong to a candidate pipe. Points that lie outside every convex hull are discarded. In such a way a segmentation of the acoustic image is obtained.

In order to project 3-D points onto the image plane, the relative pose (i.e., position and orientation) of optical and acoustic cameras is needed. This information is obtained *off-line*, once and for all, by means of a semi-automatic calibration procedure. Both acoustic and optical data are registered to the same known model of a given object in the scene, thereby obtaining the relative pose of optical and acoustic cameras.

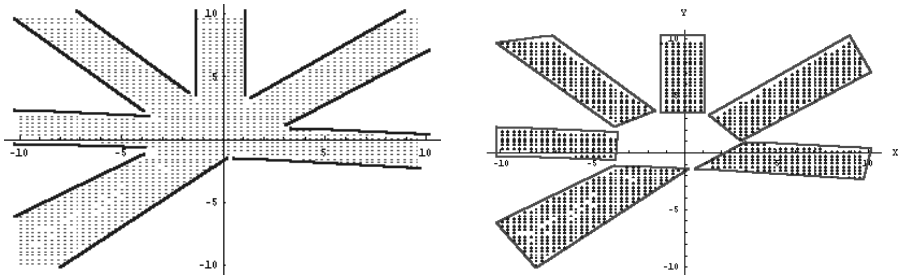

Figure 1.4 Projection of the acoustic points onto the image plane, where the optical edges are also depicted (left). Segmented points after computing the convex hulls (right).

In our approach, we used the oil rig itself as a calibration object. It must be stressed that this is the only point in this work where we use the CAD model of the oil rig. In the rest of the paper, describing the *on-line* functioning, only generic assumptions will be made (namely, knowing that the rig consists of pipes). The procedure can be summarized as follows (for more details refer to [7]):

- calibrate camera intrinsic parameters, using Roberts' algorithm [19] and a suitable calibration rig;
- register 3-D data points to the model by using the Iterative Closest Point algorithm [2], thereby obtaining the pose of the acoustic camera;
- match image segments and model segments in the image, using an algorithm due to Scott and Longuett-Higgins [21];
- register optical segments to the model, using Lowe's algorithm [14] to find the pose of the camera.

Figure 1.4 shows an example of segmentation of synthetic data.

3. CYLINDERS EXTRACTION

Regions segmented in the optical plane are now back-projected into the original 3-D frame where the we look for cylindrical surfaces. The subsequent phase consists in estimating the axis and the radius of the pipes in order to reconstruct a synthetic representation of the objects.

3.1 FINDING CYLINDRICAL REGIONS

Pipe-like regions are detected by analyzing their *inertial tensor* (this criterion is related to the so called *principal component analysis* as discussed in [8]). The inertial tensor I of a set of 3-D points $\{\mathbf{x}_i\}$ of unit mass is defined as:

$$J = \sum_i (\mathbf{x}_i - \mathbf{o}) \sqcap (\mathbf{x}_i - \mathbf{o})$$

where $\mathbf{o}$ is the center of mass of the distribution and the symbol $\sqcap$ denotes the following operator:

$$\mathbf{a} \sqcap \mathbf{b} \equiv \begin{pmatrix} (a_y b_y + a_z b_z) & -a_x b_y & -a_x b_z \\ -a_y b_x & (a_x b_x + a_z b_z) & -a_y b_z \\ -a_z b_x & -a_z b_y & (a_x b_x + a_y b_y) \end{pmatrix}$$

We denote with $\{\lambda_i\}$ $i = 1, 2, 3$ the eigenvalues of J ordered by magnitude and with $\mathbf{e}_i$ the respective eigenvectors. For a cylindrical distribution of points, it can be shown that the eigenvector $\mathbf{e}_1$ points in the axis direction and the following relations holds for the eigenvalues:

$$\lambda_1 \ll \lambda_2, \qquad \lambda_2 \simeq \lambda_3.$$

Therefore, if one eigenvalue is much smaller with respect to the others, the region is classified as a cylinder, otherwise it is discarded. This algorithm needs a threshold to decide to what extent λ_1 has to be smaller with respect to the other two eigenvalues. If this threshold is too small, elongated regions can be misclassified as cylinders. On the other hand, if it is too high, some pipes could be lost.

3.2 FITTING CYLINDERS

In order to fit a cylinder to the cylindrical regions extracted in the previous step we need to find the axis and the radius. Axis direction is given by $\mathbf{e}_1/||\mathbf{e}_1||$, axis length is obtained by projecting the points belonging to the cylinder onto a plane parallel to the axis and computing the height of the bounding box of the points.

In order to find the radius of the pipe, we project the points belonging to the cylinder onto a plane perpendicular to its axis. 3-D points are not distributed on a cylindrical surface, but only on a portion of it, as only the sector of the pipe facing the camera backscatters the sonar signal. Hence, their projections lie approximately on a circular sector. The center of the circle fixes the position of the cylinder and the radius gives the radius of the cylinder. The problem of fitting a circle to the points is a classical parametric regression problem, that we solved using the

robust Least Median of Squares (LMedS) technique [20]. The principle behind LMedS is the following:

1. given a regression problem, in which d is the minimum number of points determining a solution (three, in our case)
2. compute a candidate model based on a randomly chosen d-tuple from the data;
3. estimate the fit of this model to *all* the data, measured by the median of the squared *residuals*;
4. if the current fit is better than the previous one update the model;
5. repeat from step 2.

The optimal model represents the majority of data. Data points that do not fit into this model are *outliers*. The *breakdown point*, i.e., the smallest fraction of outliers that can yield arbitrary estimate values, is 50%. Although, in principle, all the d-tuples should be evaluated, in practice a Monte Carlo technique is applied, in which only a random sample of size m is considered. Assuming that the whole set of points may contain up to a fraction $\epsilon = 0.5$ of outliers, and requiring that the probability of missing the optimal solution be $P = 0.1$, the sample size m is [24]:

$$m = \frac{\log(P)}{\log(1-(1-\epsilon)^d)} = 17. \tag{1.1}$$

Although LMedS is usually a computationally intensive method it is practicable in our case, due to the low dimensionality of the problem. Moreover, the following observation helps in reducing the number of evaluations. When the three points of the sample are very close to each other the estimation of the circle from such points is instable, and it is a waste of time to evaluate such a sample. In order to achieve better efficiency we used a bucketing technique, analogous to the one developed in [24], which works as follows. The rectangle containing the n points is partitioned in three regions (buckets) along the major dimension, each of them containing $n/3$ points. Each triple to be fitted with a circle is built by taking one random point from each bucket. This technique does not change the probability of a point to be selected, since each bucket has the same number of points.

An example of robust circle fitting is shown in Figure 1.5, where some of the circles that have been fitted in the LMedS process are depicted, and the selected one is drawn in bold line.

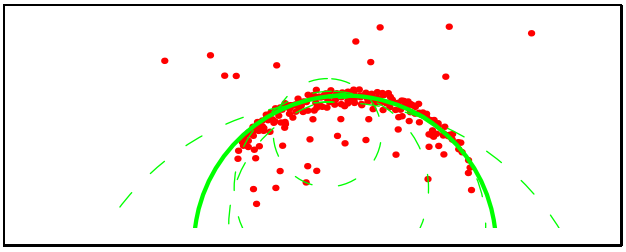

Figure 1.5 Some fitted circles and, in bold, the LMedS one.

3.3 FINDING INTERSECTIONS

In general, the axis of pipes belonging to a joint will not intersect exactly in one point or may not intersect at all. To extract an approximate intersection we use the following simple algorithm: for every axes pair i, we compute the midpoint $\mathbf{m}_i$ of the unique segment that connect the two lines defined by the axes and that is perpendicular to both of them.

If the number of axes is n, the number of possible pairs is $n(n-1)/2$. We define the centre of the joint as the center of mass of these midpoints, i.e.:

$$\frac{\sum_{i=1}^{n(n-1)/2} \mathbf{m}_i}{n(n-1)/2} \tag{1.2}$$

Since we consider the *line* containing the axis, we retain only intersections that are close enough too the axis endpoints.

This method works straightforward if there is only one joint in the scene; if this is not the case it is necessary to preliminary subdivide the set of extracted pipes in subsets containing pipes that belong to the same joint. To do this, it is sufficient to group pipes whose distance, defined as the distance between the lines passing through the axis, is below a threshold that depends on the radius of the pipes. This can be done by building the *incidence graph* G of the pipes, i.e. a graph whose nodes are the pipes and in which two nodes are connected if the distance between the corresponding pipes is below the given threshold. A joint correspond to a maximal complete subgraph of G i.e, a complete subgraph that is not contained in any larger complete subgraph. Two distinct joint can have no more than one node in common, corresponding to the pipe that connect them. The algorithm can be summarized as follows:

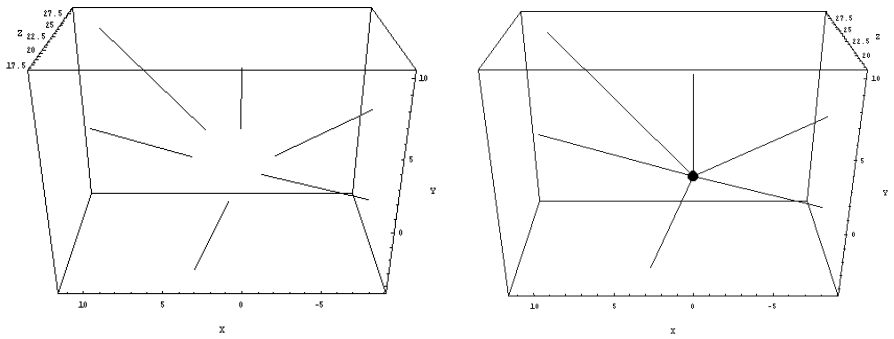

Figure 1.6 Cylinders axes as extracted from data (left) and after computing the intersection (right).

1. start with the graph G of order n (the total number of pipes) and with an empty list of joints;
2. while $n > 1$ repeat the following steps:
3. search for a complete subgraph of G of order n that is not contained in a subgraph of the list of joints.
4. if the latter exists, add it to the list of joints. Otherwise decrement n.

A complete subgraph of order three may not represent a real joint, but a triangle formed by three pipes (See Figure 1.7). This a degenerate case which is easily handled. It is sufficient to calculate the three midpoints $\mathbf{m}_i$ defined above for the three pairs of pipes and discard those for which the distance is greater than a threshold.

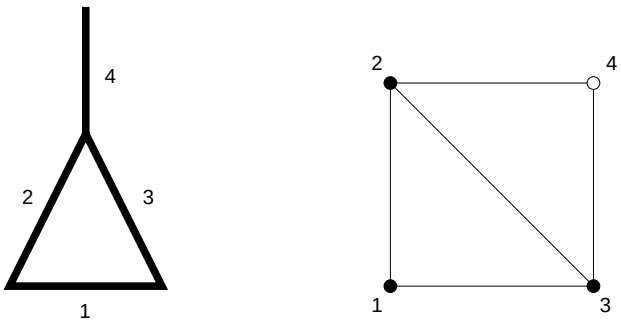

Figure 1.7 Example of a degenerate case. The rig depicted on the left has a proper joint and a false one, as its graph (right) has two complete subgraph of order three. The proper joint (2,3,4) shares two pipes with the false one (1,2,3).

For each of the remaining joints its center is computed using Formula 1.2.

4. AUGMENTED REALITY

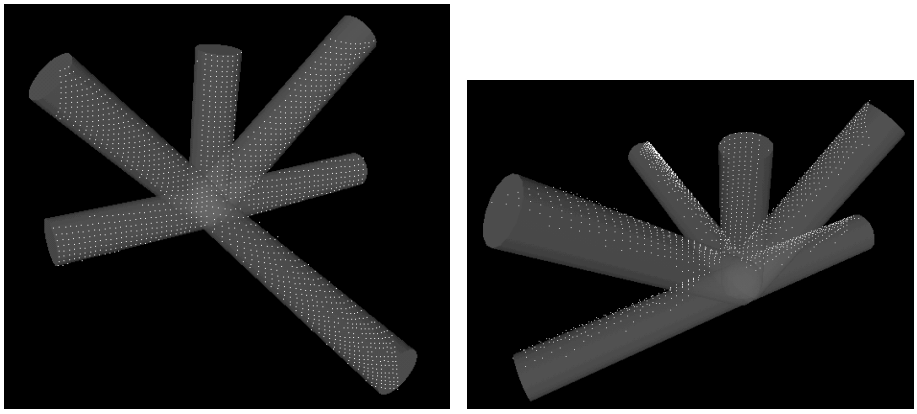

Figure 1.8 Augmented reality: virtual reconstruction of the joint with 3-D points superimposed, from two different viewpoints.

Once pipe axes have been estimated together with their radius and their reciprocal intersection, it is possible to build a VRML (Virtual Reality Modeling Language) representation of the scene observed. Owing to the registration of optical and acoustic data to the model, the synthetic representation can be superimposed on actual data, to support the ROV operator.

5. EXPERIMENTAL RESULTS

We performed experiments with real and synthetic images. Figures 1.3,1.4,1.6,1.8, show an example of our technique applied to a synthetic case. Due to the nature of the data, this case is not particularly interesting. In this section we describe results obtained in a real case.

Figure 1.10 shows a real image of a joint between four pipes, and the segments extracted from the image as described previously. Note that, due to the low quality of the image, only some segments have been detected.

Figure 1.9 shows the corresponding 3-D data, as returned by the Echoscope and the the result of pre-processing.

Using the algorithm illustrated in Section 3., cylinders are fitted to 3-D data. As one might expect, the axis direction is estimated with

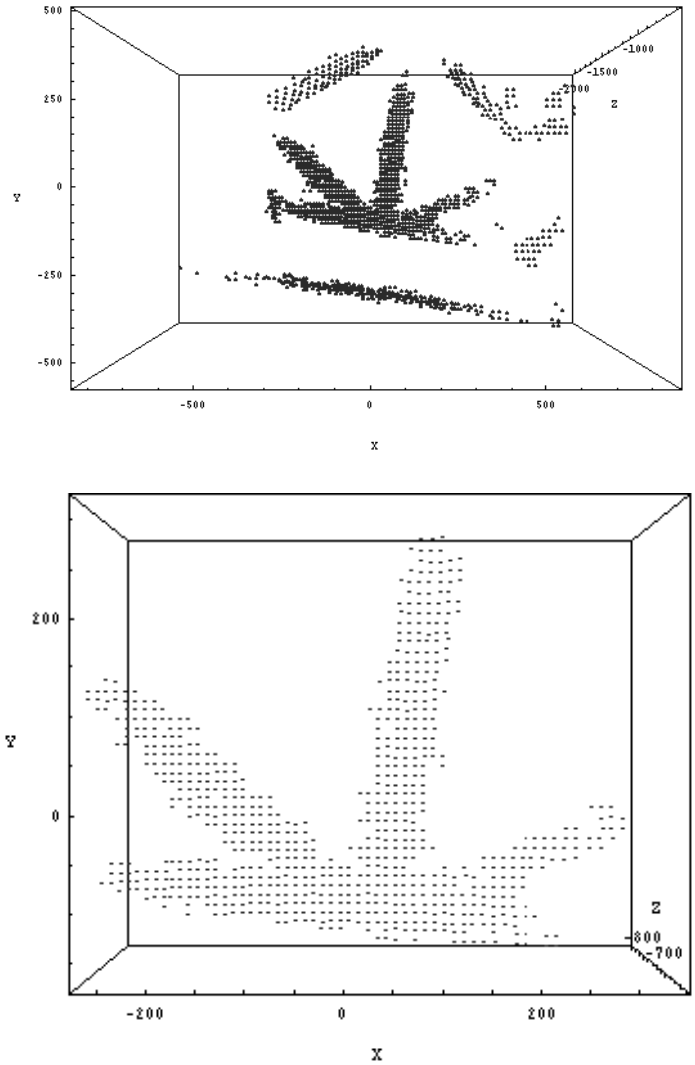

Figure 1.9 Acoustic 3-D data. Raw, from the Echoscope (top) and processed (bottom). Please note that the scale is different in the two images.

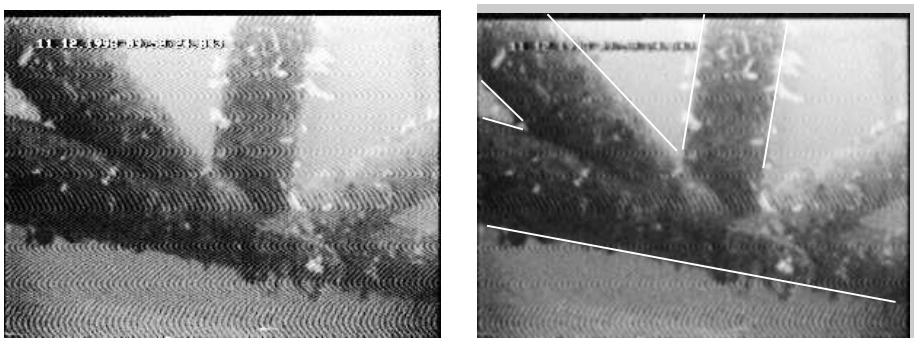

Figure 1.10 Real image (from the underwater camera) of a joint (left) and the processed image with extracted segments (right) .

far better accuracy than the radius. Indeed, we obtain, on the average, a relative error of 1% on the axis direction and of 20% on the radius. Figure 1.11 shows the projection onto the image plane of the 3-D points together with the boundaries of the extracted cylinders.

Finally, the reconstructed joint along with the original 3-D data are shown in Figure 1.12. Note that some pipes are missing in this joint, namely, the ones corresponding to the missing segments in the image. We are *not* assuming here that a high level description of the imaged portion of the rig is available. *The number, position and radius of the cylinders are obtained from the data only.*

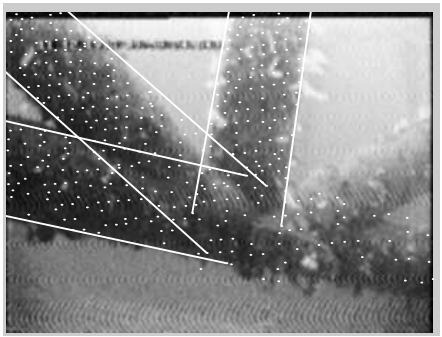

Figure 1.11 3-D points and pipes boundaries in the image plane.

6. SUMMARY

In this paper, the integration of optical and 3-D acoustic data for virtual scene reconstruction is addressed.

This work, carried out within the VENICE project (http://www.disi.unige.it/project/venice/), is aimed at presenting an integrated and informative view of the working environment to an underwater ROV operator. The ROV is equipped with an acoustic camera and an optical camera, and its task is the inspection, maintenance and repairing of an oil rig. The only a-priori information that we exploit is that the rig consists of connected pipes. No high level (CAD) description of the portion of the rig in the view frustum is available.

Our method can be summarized as follows:

1. extract pipe boundaries in the optical image;
2. project 3-D acoustic points onto the optical image plane;
3. segment points using pipes boundaries in the image and back-project them into the 3-D frame;
4. build a virtual reconstruction by fitting cylinders to the segmented data.

This is one of the few attempt to integrate different sensor modalities and actually fuse data having different nature and physical characteristics.

Presently, there is only a one-way influence of optical features on the analysis of 3-D acoustic data. We plan to investigate other schemes incorporating backtracking and mutual influence.

Acknowledgments

This work is supported by the European Commission under the BRITE-EURAM III project no. BE-2013 VENICE (Virtual Environment Interface by Sensor Integration for Inspection and Manipulation Control in Multifunctional Underwater Vehicles). The authors would like to thank Dr. R.K. Hansen of Omnitech A/S (http://www.omnitech.no) for kindly providing the images acquired by the Echoscope acoustic camera, Riccardo Giannitrapani for the fruitful discussions, and Claudio Miatto, who wrote part of the code used in experiments.

References

[1] D. H. Ballard and Brown C. M. *Computer Vision.* Prentice-Hall Inc., 1982.

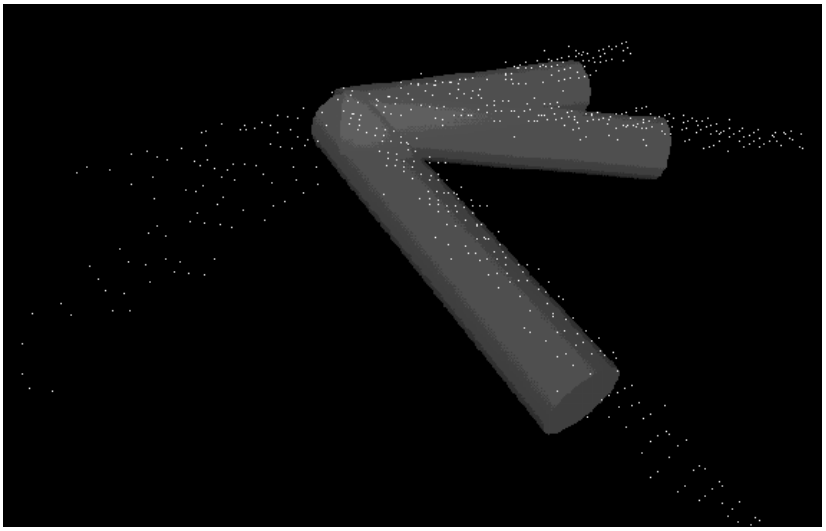

Figure 1.12 Augmented reality: virtual reconstruction of the joint with 3-D points superimposed, from two different viewpoints

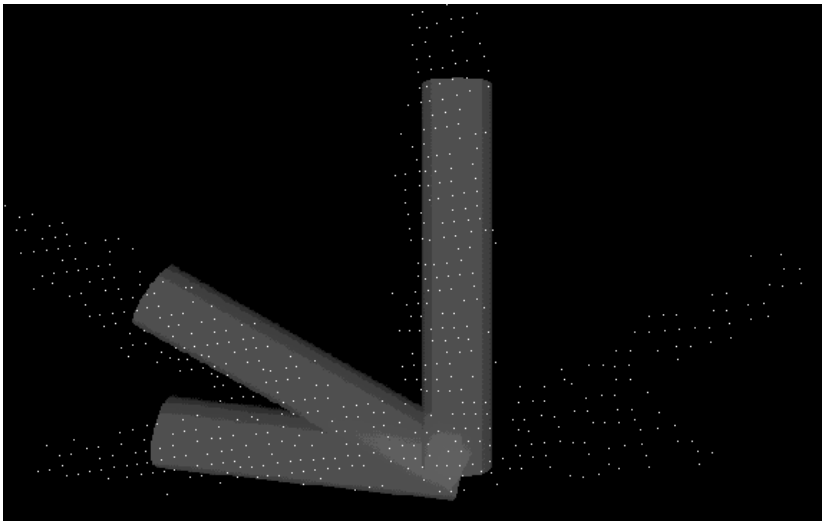

Figure 1.13 Augmented reality: virtual reconstruction of the joint with 3-D points superimposed, from two different viewpoints

[2] P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.

[3] R.R. Brooks and S.S. Iyengar. *Multi-Sensor Fusion.* Prentice Hall, Upper Saddle River, USA, 1998.

[4] J.B. Burns, A.R. Hanson, and E.M. Riseman. Extracting straight lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(4):425–456, 1986.

[5] J.F. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.

[6] A. Fusiello, R. Giannitrapani, V. Isaia, and V. Murino. Virtual environment modeling by integrated optical and acoustic sensing. In *Second International Conference on 3-D Digital Imaging and Modeling (3DIM99)*, pages 437–446, Ottawa, Canada, 4-8 October 1999. IEEE Computer Society Press.

[7] A. Fusiello and V. Murino. Calibration of an optical/acoustic sensor. In *6th International Conference on Computer Graphics and Image Processing (GKPO2000)*, 2000. Submitted.

[8] R. Giannitrapani, A. Trucco, and V. Murino. Segmentation of underwater 3D acoustical images for augmented and virtual reality applications. In *Proceedings of the OCEANS'99 Conference*, pages 459–465, Seattle (USA), September 1999. MTS/IEEE.

[9] F. Goulette. AutomaticCAD modeling of industrial pipes from range images. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 229–233, May 1997.

[10] B. Gunsel, A.K. Jain, and E. Panayirci. Reconstruction and boundary detection of range and intensity images using multiscale MRF representations. *CVGIP: Image Understanding*, 63(2):353–366, March 1996.

[11] R. K. Hansen and P. A. Andersen. A 3-D underwater acoustic camera - properties and applications. In P.Tortoli and L.Masotti, editors, *Acoustical Imaging*, pages 607–611. Plenum Press, 1996.

[12] M. Hebert, R. Hoffman, A. Johnson, and J. Osborn. Sensor based interior modeling. In *American Nuclear Society 6th Topical Meeting on Robotics and Remote Systems (ANS '95)*, pages 731 – 737, February 1995.

[13] D. Dion Jr. and D. Laurendeau. Generalized cylinders extraction in a range image. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 141–147, May 1997.

[14] D.G. Lowe. Fitting parameterized three-dimensional models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):441–450, May 1991.

[15] R. C. Luo and M. G. Kay. Multisensor integration and fusion in intelligent systems. *IEEE Transactions on Systems, Man and Cybernetics*, 19(5):901–931, September-October 1989.

[16] M. Maimone, L. Matthies, J. Osborn, E. Rollins, J. Teza, and S. Thayer. A photo-realistic 3-D mapping system for extreme nuclear environments: Chornobyl. In *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robotic Systems (IROS '98)*. IEEE, 1998.

[17] V. Murino, A. Trucco, and C.S. Regazzoni. A probabilistic approach to the coupled reconstruction and restoration of underwater acoustic images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):9–22, January 1998.

[18] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[19] L Robert. Camera calibration without feature extraction. *Computer Vision, Graphics, and Image Processing*, 63(2):314–325, March 1996.

[20] P. J. Rousseeuw and A. M. Leroy. *Robust regression & outlier detection*. John Wiley & sons, 1987.

[21] G. Scott and H. Longuet-Higgins. An algorithm for associating the features of two images. In *Proceedings of the Royal Society of London B*, volume 244, pages 21–26, 1991.

[22] C. Yu W. Lie and Y. Chen. Model-based recognition and positioning of polyhedra using intensity-guided range sensing and interpretation in 3D space. *Pattern Recognition*, 23:983–997 1990.

[23] G.H. Zhang and A. Wallace. Physical modeling and combination of range and intensity edge data. *CVGIP: Image Understanding*, 58(2):191–220, September 1993.

[24] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1-2):87–119, 1995.