3D acoustic image segmentation by a RANSAC-based approach

L. Tao, U. Castellani, A. Fusiello, and V. Murino
Dipartimento di Informatica
University of Verona
37134 Verona, Italy
Email: vittorio.murino@univr.it

***Abstract*— In this paper, a new technique for 3D acoustic image segmentation and modelling is proposed. Especially, in the underwater environment, in which optical sensors suffer from visibility problems, the acoustical devices may provide efficient solutions, but, on the other hand, acoustic image interpretation is surely more difficult for a human operator. The proposed application involves the use of an acoustic camera which directly acquires images structured as a set of 3D points. Due to the noisy nature of this type of data, the segmentation problem becomes more challenging and the standard algorithms for range image segmentation are likely to fail. The proposed method is based on a simplified version of the so called *recover and select* paradigm in which the seed areas, from which the segmentation starts, are generated by adopting a robust approach based on the RANSAC (RANdom Sample And Consensus) algorithm. Superquadric primitives are directly recovered from raw data without any pre-segmentation processing. Experimental trials using real acoustical images confirm the goodness of the method, and a large robustness of the resulting segmented images, associated to a relatively low computational load.**

I. Introduction

The image segmentation problem has been approached in many ways in the literature, concerning all types of images (optical, acoustical, etc.), but it has not been yet solved. Furthermore, one of the primary issues in the vision community is to extract primitive models from images that would bridge the gap between low-level features and high-level symbolic structures useful for further processing (i.e., image interpretation). In this paper, we focus on the segmentation and modelling of underwater acoustic images for which the problem becomes more challenging because of the very noisy nature of acquired data. Classical range image segmentation algorithms [1] are based on the extraction of local features such as surface normals, 3D edges, gaussian and mean curvatures, and others [2]. Unfortunately, these features are very sensitive to the noise so that it is not possible to use these methods for our particular application. To the best of our knowledge, the literature devoted to acoustic image segmentation is focused on quite complex techniques based on probabilistic approaches (e.g., Markov Random Fields) [3], [4], which take into account the type of noise heavily affecting such images. For example, in [3] the authors proposed to consider the intensity of the acoustic signal as a *reliability* value from which a *confidence* map was also obtained. Then, the proposed MRF model allowed to restore both the range and the confidence image by permitting the reconstruction and the segmentation of the scene at the same time. In [4], a hierarchical Markovian modelling approach is described. The output of the processing chain is a two class segmentation of the sonar image in shadow and sea-bottom reverberation regions. This Markov Random Field model takes into account the phenomenon of speckle noise modelled through the Rayleigh's law as well as a-priori information on the geometric object shadows. Although the probabilistic approach is effective for acoustic data, it is very computationally demanding. Less computational expensive solutions are preferred, especially for on-line applications.

Regarding models extraction, we consider that a specific set of volumetric models can be directly recovered from data [5]. This is in contrast to the common belief that the recovery of volumetric models is possible only after a pre-segmentation of the data. In particular, an interesting approach is the so called *recover and select* paradigm, proposed by Leonardis et al. [5]. The recover and select paradigm is based on the fact that as 3D data are considered, it is possible to cast the 3D segmentation as a surface fitting problem, exploiting such information by grouping points likely belonging to a single surface. In this way, segmentation also leads to estimated shape model parameters, so that the geometric models of the partitions are also obtained. Starting from a set of randomly selected seeds, the algorithm iteratively switches between the growing process by searching new compatible points (the *recover* phase) and the process of discarding models by solving the quadratic boolean problem [5] (the *select* phase, i.e., model selection).

As pointed out above, in order to manipulate our noisy data we need to exploit robust techniques based on a conservative approach. In particular, in our algorithm we proceed step by step by extracting from the images the most reliable information. The proposed method is based on a simplified version of the cited recover and select paradigm in which the seeds are generated by using a RANSAC-based algorithm. The RANSAC (RANdom Sample And Consensus) is an algorithm developed by Fischler and Bolles with the aim at reducing the influence of outliers (i.e., noise) for function fitting [6] problems. Several initial small data set of points are selected randomly on which a model is fitted; subsequently, these sets grab the votes from consistent data taken from the whole set. The model which is consistent with the majority of the points (i.e., that fits most of the points with respect to the other

models) is the winner model. Using this approach a very small set of highly reliable seeds is extracted from which, after the growing phase (recovery), only the best model is retained (selected). Points compatible to the extracted model are then discarded (i.e., they form a segment) from the original cloud of point and the process is repeated up to all models can be identified. Superquadric models have been chosen because of their compact representation and their well-known good properties [7].

The rest of the paper is organized as follow. In Section II, we shortly describe the acoustic imaging process, and in Section III the superquadric models are defined. Section IV describes in detail the proposed algorithm. Results of the algorithm on real images are shown in Section V, and finally, in Section VI conclusions are drawn.

II. The acoustic imaging sensor

Three-dimensional acoustic data are obtained with a high resolution acoustic camera, the *Echoscope* 1600 [8]. An acoustic signal is transmitted and the echoes from the scene targets are collected and processed. Acoustical intensities and range information are retrieved for several viewing directions (beam directions). The camera has a 40×40 array of receiving transducers and operates on the pulse-echo principle [9]. The acoustic image is formed by using a sort of beamforming (BF) technique. In short, it is a spatial filter that combines linearly temporal signals spatially sampled by a discrete antenna [10]. In this way, when a scene is insonified by a coherent pulse, the signals representing the echoes backscattered from objects present in a specific direction contain attenuated and degraded replicas of the transmitted pulse. The resulting 3D image is formed by 64×64 points ordered according to a polar reference system, as adjacent points correspond to adjacent beam signals. A minimum and maximum detection range values define the viewing volume of the underwater scene. There is a trade-off between range resolution and the field of view. The resolution depends on the frequency of the acoustic signal (it is about 3 cm at 500 KHz): roughly speaking, the higher the frequency, the higher the resolution, and the narrower the field of view. The acoustic image is also affected by false reflections, caused by secondary lobes, and by acquisition noise, which is modelled as speckle [11]. Moreover, the intensity of the maximum peak can represents the reliability of the associate 3D measures, so that, generally speaking, the higher the intensity, the safer the estimated distance. A dramatic improvement of the range image quality is obtained by discarding points whose associated intensity is lower than a threshold depending on the secondary lobes [3], [11].

III. The superquadric models

The chosen basic shape models are represented by the superquadrics [7], which are an extension of the typical quadric solids, particularly useful as 1) they have psycho-physical justifications, 2) are general, and 3) can be used as part-based models that can be deformed to adapt to articulated shapes. Moreover, unlike the usual quadrics, they do not require to change the model when the error-of-fit measure becomes too large. This is a very useful characteristic which supersedes typical methods for acoustic images segmentation as those proposed in [12].

A superquadric surface is defined by the following implicit equation:

$$F(x,y,z) \equiv \left(\left(\frac{x}{a_1}\right)^{\frac{2}{\epsilon_2}} + \left(\frac{y}{a_2}\right)^{\frac{2}{\epsilon_2}} \right)^{\frac{\epsilon_2}{\epsilon_1}} + \left(\frac{z}{a_3}\right)^{\frac{2}{\epsilon_1}} = 1, \quad (1)$$

where a_1, a_2, a_3 define the superquadric size, and ϵ_1, ϵ_2 define a smoothly changing family of shapes, from a rounded one to a square. Nevertheless, the implicit function for a superquadric in a general position is given by:

$$F(x,y,z;\Lambda) = F(x,y,z;a_1,a_2,a_3,\epsilon_1,\epsilon_2,\phi,\theta,\psi,p_x,p_y,p_z), \quad (2)$$

where ϕ, θ, ψ define the orientation in space, and p_x, p_y, p_z define the position in space. Given a set of points $\mathbf{x_i}$, the superquadric's parameters Λ are recovered applying the Levenberg-Marquardt algorithm for least squares estimation of non-linear parameters [2].

In order to define a *goodness of fit* measure, an effective estimation of the Euclidean distance of a general point x_k from the the superquadric surface is given by:

$$D = (a_1 a_2 a_3)^P \cdot (F^{\epsilon_1}(x_k;\Lambda) - 1) \quad (3)$$

where $(a_1 a_2 a_3)$ is related to the volume of the bounding box of the superquadrics and P is an arbitrary parameter [13]. Although in [12] the authors proposed to fix $P = 0.5$, an empirical value derived after several trials on different objects, we need to tune this coefficient because of the noisy data, especially for the growing phase.

IV. The segmentation algorithm

The proposed algorithm is based on the iteration between two main phases. The first phase is aimed at extracting very reliable seeds. The second phase is focused on the growing and selecting models. From these two phases a single model is extracted and the process is repeated until all the superquadrics are recovered.

A. Reliable seeds selection

The initial seeds are generated by randomly selecting M subsets of points $S_i \subset X$ from the whole 3D data set X. Then, a superquadric is fitted on each seed. Since superquadrics are defined by 11 parameters each seed area must be composed of at least 11 points. To face numerical problems and by carrying out empirical experiments, we define the cardinality of each subset as $\sharp(S_i) = 30$. We also introduce a distance constraint on the randomly chosen points in order to avoid the generation of large seeds. Furthermore, the number of seeds M can range from some hundreds to ten thousands. According to the RANSAC approach, the method for selecting good seeds is based on a consensus processing. There is a trade-off between computational load and the quality of seeds. By generating

more starting seed we are more likely to obtain good seeds but we increase the computational load. In order to avoid the estimation of the tolerance threshold [6], we reverse the voting scheme: instead of fixing the model and counting their compatible points (i.e., the points for which their distance to the model is closer than the tolerance), we define a score of the most compatible models for each point. The distances from the point to each superquadric are calculated by using eq.(3), so that the closest n superquadrics receive a vote. After this step, for all the superquadrics, we calculate a) the number of compatible points, and b) the mean value of the distances from each compatible point to the superquadric (mean error).

The best superquadrics are selected in two stages. In the first selection, only the 10% of superquadrics having the majority of the votes survive. In the second selection, the N superquadrics having the least mean error are chosen.

The selected superquadrics SQ_i $(i = 1, ..., N)$ are used as seeds in the successive expanding procedure.

B. Expanding and segmentation

After the seeds selection process is completed, the growing phase proceeds as follows:

Let SQ_k be a superquadric seed, S_k the set of compatible points, and ME_k the mean error. The method for expanding seed superquadrics is described as follows:

1) For all of the points $x_i \in X$:
   a) calculate the distance D between each point and SQ_k using eq.(3).
   b) if $D < (CS \cdot ME_k)$ the point x_i is joined to S_k. CS is a parameter that controls the expanding speed.
2) Re-fit SQ_k on S_k.
3) Repeat step 1 and 2 until there are no more points that can be joined to the superquadrics.

All of the seed superquadrics SQ_i are expanded by employing the above procedure. The best superquadric $SQ_{\hat{k}}$, i.e., that having the least mean error, survives, and the final set $S_{\hat{k}}$ is the extracted 3D segment.

C. Summary of the algorithm

Finally, the overall segmentation and models extraction algorithm can be summarized as following:

1) Set $X_{fit} = X$.
2) find $SQ_i, S_i, ME_i (i = 1, ..., N)$ from X_{fit}.If there are no more seeds then STOP.
3) find $SQ_{\hat{k}}$, $S_{\hat{k}}$ and store them.
4) Set $X_{fit} = \{X_{fit} \setminus S_{\hat{k}}\}$.
5) goto 2.

V. Experimental results

In this section, we show some results on real images. The scene is composed of a oil rig composed by pipes converging in a joint. Figure 1 shows the image as it appears from the acoustic camera. Data are structured as unorganized set of 3D points. The images are very noisy and it is quite difficult for a user to understand the observed scene.

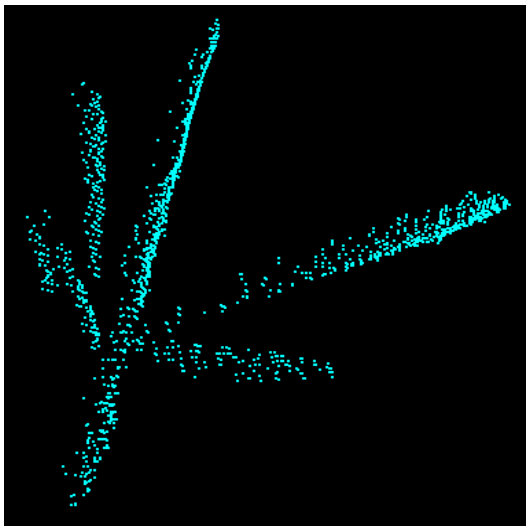
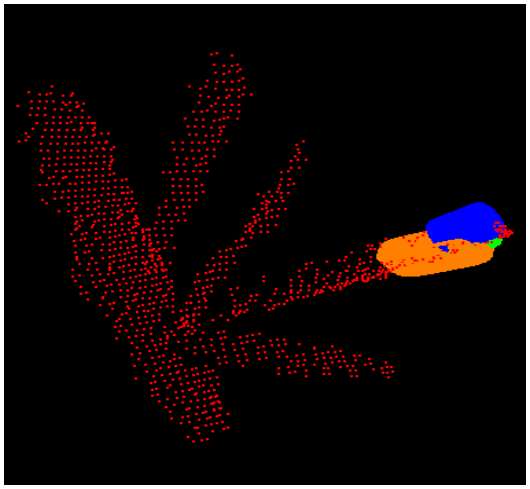


Fig. 1. 3D acoustic image

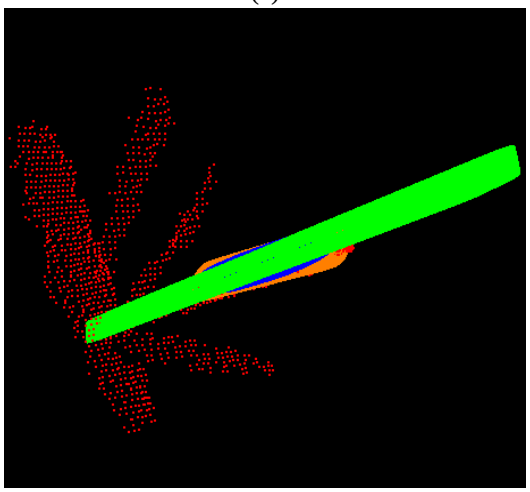
The initial subsets are randomly chosen and the consensus process is carried out. We verify empirically that the free parameters M and n described in section IV-A are not very sensitive for the algorithm performances, but they only influence the speed. After the seed selection, the best starting superquadrics are chosen (Figure 2.a). For this experiment, we retained only $N = 3$ models. The growing phase is carried out and the first superquadric is recovered (Figure 2.a). We deduced from experiments that introducing an expanding criterion based on mean error is more reliable than fixing a global threshold, even if an arbitrary parameter CS must be estimated. Proceeding with the algorithm, all the models are recovered and the final segmentation (Figure 3.a) and the modelling (Figure 3.a) of the scene are obtained. Regarding performance aspects, we verified that the algorithm takes only a few seconds on a PC Pentium IV 1.6Gb having a memory of $256Mb$ for about one thousand 3D points.

VI. Conclusions

In this paper, a new technique for 3D acoustic image segmentation and modeling is proposed. Due to the very noisy nature of the images the segmentation is made more challenging than for optical 3D images, and a robust approach is necessary, aiming at extracting more reliable information from the observed data. The described algorithm is able to extract both segments and models at the same time from a set of unorganized 3D points acquired by an acoustic camera. The performance are computationally acceptable and it is very simple to estimate the best values of the few free parameters. The results show the effectiveness of the obtained partitions and models, allowing the extraction of interesting information which is useful for improving the understanding of an underwater scene.

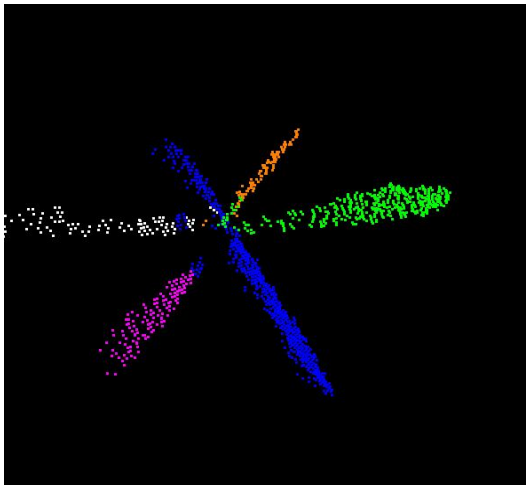


(a)

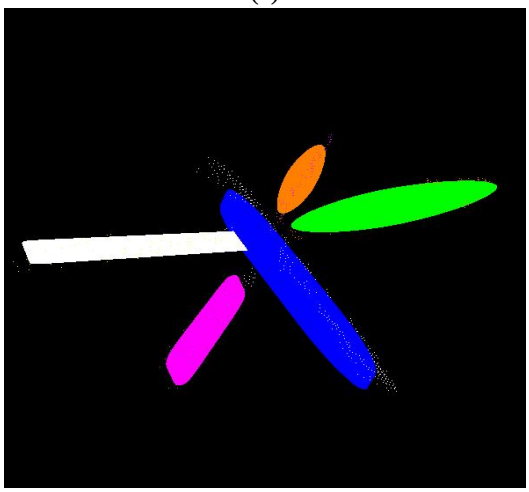


(b)

Fig. 2. Best seeds (a) and extracted superquadric (b)



(a)



(b)

Fig. 3. Extracted segments points (a) and superquadric models (b)

ACKNOWLEDGMENTS

This work was supported by the European Commission under the project no. GRD1-2000-25409 ARROV (Augmented Reality for Remotely Operated Vehicles based on 3D acoustical and optical sensors for underwater inspection and survey). Echoscope images are courtesy of Dr. R.K. Hansen of CodaOctopus Omnitech (Norway).

REFERENCES

[1] A. Hoover G. Jean-Baptiste X. Jiang P.J. Flynn H. Bunke D. Goldgof K. Bowyer D. Eggert A. Fitzgibbon and R. Fisher, "An experimental comparison of range segmentation algorithms," *IEEE Trans. Pat. Anal. and Mach. Intel.*, vol. 7, no. 18, pp. 673–689, 1996.

[2] Emanuele Trucco and Alessandro Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice-Hall, 1998.

[3] 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*, vol. 20, no. 1, pp. 9–22, January 1998.

[4] C. Collet, P. Thourel, P. Perez, and P. Bouthemy, "Hierarchical mrf modelling for sonar icture segmentation," in *IEEE Int. Conf. on Image Processing*, Los Alamitos USA, 1996, vol. 3, pp. 979 – 982.

[5] A. Leonardis, A. Jaklic, and F. Solina, "Superquadrics for segmenting and modeling range data.," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 11, pp. 1289–1295, 1997.

[6] M.A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *CACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[7] F. Solina and R. Bajcsy, "Recovery of parametric models from range images: the case for superquadrics wigh global deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 2, pp. 131–147, 1990.

[8] R. K. Hansen and P. A. Andersen, "A 3-D underwater acoustic camera - properties and applications," in *Acoustical Imaging*, P.Tortoli and L.Masotti, Eds., pp. 607–611. Plenum Press, 1996.

[9] R. J. Urik, *Principles of Underwater Sound*, McGraw-Hill, 1983.

[10] V. Murino and A. Trucco, Eds., *Special Issue on Underwater Computer Vision and Pattern Recognition*, Computer Vision and Image Understanding, July 2000.

[11] A. Trucco V. Murino, "Three-dimensional image generation and processing in underwater acoustic vision," *Proceeding of the IEEE*, vol. 88, no. 12, pp. 1903–1946, December 2000.

[12] R. Hoffman and A. K. Jain, "Segmentation and classification of range images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 9, no. 5, pp. 608–619, 1987.

[13] P. Whaite and F.P. Ferrie, "From uncerainty to visual exploration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 10, pp. 1038–1049, 1991.