Reconstruction of Complex Environments by Robust Pre-aligned ICP

Vittorio Murino, Luca Ronchetti, Umberto Castellani and Andrea Fusiello
Dipartimento Scientifico e Tecnologico
University of Verona, Italy
{murino,castellani,fusiello}@sci.univr.it

Abstract

This paper proposes a technique for three-dimensional reconstruction of an underwater environment from range views acquired by an acoustic camera. The final target of the work lies in improving the understanding of a human operator driving an underwater remotely operated vehicle floating close to an offshore structure. Due to the narrow field of view and the absence of control of the sensor position, no information is available about the degree of overlapping between the range images; further, speckle noise and low resolution make more difficult the registration process. In this paper, we propose a preprocessing method which gives a coarse alignment of range images prior to running the Iterative Closest Point (ICP) algorithm for the accurate registration of views pairs. The pre-alignment is based on the matching between the three-dimensional skeletons extracted from the images. A comparative analysis is presented where our method is compared with plain ICP, and with a technique based on principal components analysis.

1. Introduction

Underwater exploration is nowadays growing due to both industrial and scientific needs. Fortunately, even technology is improving with the advent of smart sensors able to provide data with high visual quality, differently from only few years ago. Recently, computer vision scientists have also approached underwater scene understanding issues [17].

This work is have been carried out in the context of a project aimed at the three-dimensional (3D) scene reconstruction from a sequence of range data acquired by an acoustic camera. The final goal is to provide a 3D scene model to the human operator(s) of an underwater remotely operated vehicle (ROV), in order to facilitate the navigation and the understanding of the surrounding environment, namely an offshore rig, composed by pipes connected each others by joints (see Fig. 1).

The underwater environment is undoubtly a complex scenario for both the implicit limited accessibility and the difficulty to retrieve good quality data. Typically, acoustic systems are used to sense underwater scenes as they allow to achieve a larger visibility and to measure range distance, unlike the more used optical sensors. In the present case, our data are obtained by a high frequency acoustic camera, called Echoscope [10]. These data are affected by speckle noise, due to the coherent nature of the acoustic signals, which corrupts sensibly the visual quality and decreases the reliability of the estimated 3D measures. Moreover, there is a trade-off between range resolution and field of view. 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, the narrower the field of view. In addition, the sensor is mounted onboard an underwater vehicle with inherent limited capability of precise control. Therefore, as we are forced to operate with

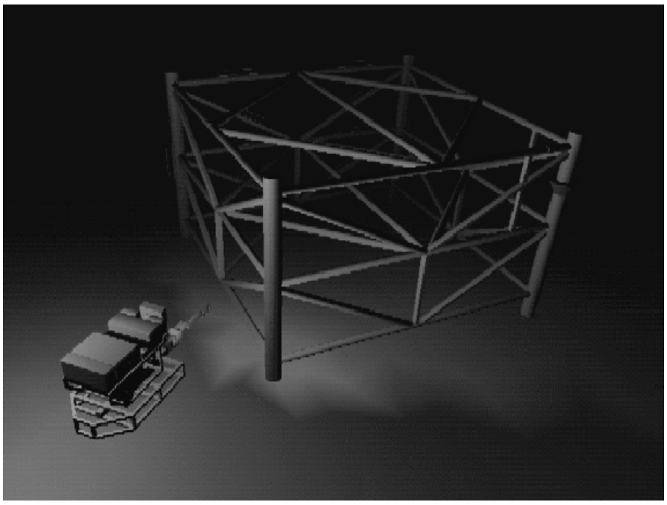


Figure 1. Rendering of the model of an oil rig with the ROV.

a limited field of view, range images are acquired with no precise idea of the position and heading of the sensor. Consequently, a technique to reconstruct progressively the scene while the sensor is moving is necessary.

In this context, a technique for registering pairs of range images having any initial pose has been explored and tested. An enhancement of the Iterative Closest Point technique (ICP) [2] is presented, in which the images are first pre-aligned using the essential structure (i.e., skeletons) of the observed objects. It is known that ICP is optimal when the two clouds of points to be registered actually correspond to the same distribution (or one is a subset of the other) seen from different viewpoints, and when they are not too far each other in the pose space. In our application, none of the above conditions can be met: actually, as the sensor is navigating around an object, some parts may disappear and other may be included, and even if the degree of overlapping is large, the point of view can be so different to preclude the convergence of the classic ICP.

A lot of works on the registration of image pairs or the integration of a set of range images are present in literature, but none dealing with the particular kind of 3D data we are using.

Among the works related to registration, the Iterative Closest Point (ICP) procedure [2] has been already quoted, and its earlier variants [6, 28] are worth to be mentioned. The original algorithm is based on the search of pairs of nearest points in the two sets, so that the rigid transformation which align them can be estimated [2]. Variants include the use of closest points in the direction of the local surface normal [6], and the use of a robust statistics technique [28] to limit the maximum distance between points. As said above, these works assume that one point set is a subset of the other. Indeed, [28] foresees the possibility to have partially overlapped data, but some thresholds must be set, and this can be a difficult and tricky operation, critically affecting the quality of the registration.

In [24] and [8], a force-based optimization approach is proposed. Assuming known the points' correspondences among the data sets, interconnections using springs between corresponding points is simulated.

A quite different approach has been proposed in [13]. Pairwise registration is here accomplished starting from triangulated meshes, and finding points' correspondences using an alternate representation called *spin image*. Then, a rigid transformation is estimated using such correspondences and an ICP variant is used to refine the registration.

Another approach to the fusion of multiple range images consists in the use of the *level set* theory. This is a statistical approach which maximize the posterior probability of a surface. In [27], this approach is utilized to reconstruct an indoor scene, mainly based on the segmentation and registration of planar surfaces. Baillard at al. [5] use the level set theory in cooperation with 3D registration to automatically segment a volumetric brain image.

Other works deal with registration of range images using surface models, mesh representations alignment, and volumetric approaches [20, 11, 26, 7, 12] using methods closer to the computer graphics.

Several other work deals with the possibility to register more range images by incrementing pairwise registration, or assuming available all the views performing global registration in one step [6, 3, 1, 23, 24, 19, 15].

An interesting approach, alternative to the proposed method, consists in segmenting and modeling the individual range images, and then, try to align the segmented regions characterized by the same parametric surface. In [14], an efficient yet accurate and robust algorithm for range image segmentation and modeling is proposed. In short, the algorithm is subdivided in two main steps. The first one is essentially a region growing method in which superquadrics are used as fitting surfaces, and the second one is a model selection phase in which insignificant models are discarded on the basis of an objective function suitably devised. The generality of superquadrics and the simultaneous segmentation and model estimation make this method interesting to be tested also in our case.

All these works assume range images from a laser range finder looking at a single, albeit complex, object. Unlike these ones, the proposed algorithm should deal with cluttered, uncertain, noisy, and low resolution data, in which problems of filtering and segmentation should be all considered in order to get a reliable reconstruction of the scene. Although these problems are also quoted as critical in other papers, the fact that they typically operate in aerial environments using laser range finders, makes the above assumptions much less critical than in our case.

Moreover, previous papers do not consider too much a pre-alignment phase, i.e., they assume range views quite close or not affecting the registration. In [4] the alignment is addressed by using the Principal Component Analysis (PCA), but the problem is backed up by a few sentences. A similar approach to the alignment of couple of range images is proposed in [22]. The problem is here the registration of archeological ceramic fragments, derived from a rotationally symmetric object, for purposes of reconstruction and classification of the original object. The idea is simple, acquire the front and back 3-D views of the fragment and align them using the (same) axis of symmetry. Axes are estimated using 3-D Hough transform under the hypothesis that all surface normals intersect in points belonging to the axis. This approach is interesting due to the inherent robustness of Hough-based technique, but the relative optimal conditions of the application (fixed object, no clutters, high resolution data) make questionable the applicability of this method in our case.

Our work deals with the pre-alignment problem proposing a robust yet simple procedure for closing the poses of the two range images. Three-dimensional skeletons are estimated from both images and matched each other in order to find a good alignment between the two views. A point set contraction algorithm [16] is used to extract 3D skeletons.

This approach is quite robust as overlapping object parts can be safely matched so that the rotation matrix and the translation vectors can be estimated and the views can be correctly aligned prior the application of ICP to accurately register the images.

As should be more clear now, this case is quite different from the registration of a couple or more range images, proposed in many previous papers [2, 6, 28] as the operative conditions are different. Actually, resolution is never better than some centimeters, sensor position is unavailable, and the motion of the sensor is quite unstable, and cannot be controlled with precision in any real case, so acquired images from a fixed position may be different due to speckle and sensor floating. As a consequence, some previous solutions based on estimation of surface parameters cannot be taken into account due to the high uncertainty of the data.

In Sec. 2 , the proposed approach is described: 3D skeletons are first extracted from both images, skeleton branches are classified, and a matching phase is finally carried out to associate skeleton branches in in the two images so that the rigid transformation between the views can be estimated. Finally, pre-aligned range images are pairwise registered using an ICP algorithm with outliers rejection [15]. Examples on real images of an underwater oil rig are presented in Sec. 3. A comparative analysis is performed with respect to simple ICP and PCA pre-alignment, showing the better performances of the proposed algorithm. Finally, conclusions are drawn (Sec. 4).

2. The method: ICP with pre-alignment

Input images are partial views of a tubular underwater rig (like the one depicted in Fig. 1) taken from different points of view.

Acoustic data points, which lie on the surface of cylinders, are expressed in the mobile reference frame attached to the sensor. The unknown rigid transformation that links two reference frames is obtained by registering the two views.

In their paper Besl and McKay [2] describe a general purpose method for the registration of rigid 3-D shapes which they refer to as the Iterative Closest Point (ICP) algorithm. This approach eliminates the need to perform any feature extraction, or to specify feature correspondence. While the ICP algorithm is only guaranteed to converge to a local minimum, there is no guarantee that this local minimum will correspond to the actual global minimum, therefore a good initial guess is mandatory. In our case, pre-alignment is obtained by matching skeletons, using the following procedure.

First, we extract the skeletons of the two cloud of points. From the skeletons we obtain a three dimensional segmentation of the image through which we can classify each point as belonging to a pipe or to a joint. We extract every pipe of the scene and calculate its orientation and centroid. We also extract every joint location and its connected pipes. Thanks to the match between correspondents joint and pipes in two images we obtain a good pre-registration, sufficient for ICP to converge. Please note that a good pre-alignment can be found with only a single joint match (and at least two connected pipes) and so the registration is possible also with very small overlap between images. In summary, the main stages of this work are skeleton extraction, segmentation, matching and pre-alignment, and, finally, ICP registration.

It is worth noting that, although the rig structure formed by pipes variously joined may seem a quite simple object, this is not actually the case. In fact, if we consider convex (even complex) objects, PCA can be sufficient for a good pre-alignment. Conversely, when rig-like objects are considered, such techniques may fail as the 3D structure is quite complicated and may be very different in the two views. Skeletons, instead, mimic thoroughly the 3D structure of the observed object parts and a good matching algorithm can find a good pre-alignment, disregarding the non overlapping parts.

2.1. Skeleton extraction

Skeleton extraction from two-dimensional binary and gray-level images is a well-known topic, but this is not the case when sets of three-dimensional (3D) points are considered. Actually, skeleton definition and extraction from a 3D sparse image, i.e. a distribution of points in $\mathbf{R}^3$ is still a challenging field of investigation. To the best of our knowledge, a few works address the problems of skeleton extraction from range images, i.e., from a set of points distributed on a surface in 3D space rather than points distributed in 3D volumes. For instance, a method using a discrete subdivision of 3D space in regular cells and the approximation of range data by two-dimensional patches is developed to extract axes of symmetry [21] in order to get skeletons from range data. Similar techniques operating on a voxel space: an algorithm to extract skeletons via a thinning algorithm is presented in [18]. In our technique [16], the set of 3D points is iteratively contracted up to form a thin distribution that can be assimilated to a skeleton .

Let us define a *skeleton* as a distribution of 3D points that 1) must be thinner than the original one, 2) must be located in the neighbors of the median lines of the original point set, 3) must have the same *homotopy group* (for example, a torus must have as a skeleton a circle, while a bitorus must have two circles connected at a point of their perimeter) and 4) have to be invariant to 3D rotation. Let X be a 3D image, i.e., the set of points in $\mathbf{R}^3$,

$$X = \mathbf{x}_i \qquad i = 1, ..., N$$

where $\mathbf{x} = (x_x, x_y, x_z)$.

We also define for every point *i* and every $R \in \mathbf{R}^+$ (ray of a sphere about the point *i*) the subset O_i^R of X and the 3D point $\mathbf{y}_i^R$ defined in the following way:

$$O_i^R \equiv \{\mathbf{x}_j \in R : |\mathbf{x}_j - \mathbf{x}_i| < R\}$$

$$y_i^R \equiv \frac{\sum_{x_j \in O_i^R} \mathbf{x}_j}{dim\{O_i^R\}}$$

where $\dim\{O_i^R\}$ is the cardinality of O_i^R (i.e., $\mathbf{y}_i^R$ is the centroid of O_i^R).

Defining the following image transformation:

$$X \longrightarrow X^R = \{y_i^R\} \tag{1}$$

and indicating with the symbol $X^{R,n}$ the iterative application of it for n times, our skeleton extraction is simply the construction of the image $X^{R,n}$ for a suitable choice of R and n. The overall effect of this transformation is to shift points on the border toward the center, while leaving unaltered the points well inside an object. The iterative application of such transformation shifts all the points of the distribution towards its skeleton. More details can be found in [16].

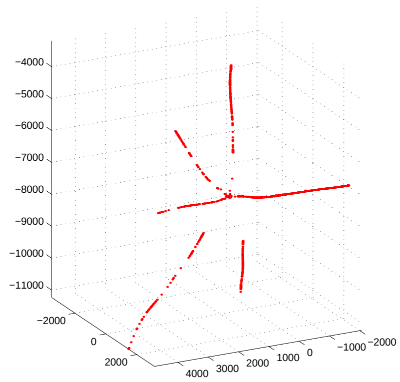


Figure 2. Skeleton extracted from the data cloud of Experiment 1

2.2. Segmentation

In order to match the two views we need to extract the pipes in each view and to group pipes belonging to the same joint.

2.2.1 Pipe extraction

In our images, we are interested classifying between points belonging to a single pipe, and points belonging to the intersection of two or more pipes (joints). Skeleton points belonging to a pipe segment are very close to the pipe axis and so they lie on a straight line. For each point we consider all its neighbors in a spherical region: if all of them belong to an approximately straight segment, the point is classified as a pipe, otherwise it is a joint. On the points recognized as belonging to the skeleton of a pipe, a Least Squares straight line is fitted, representing the pipe axis.

2.2.2 Joint extraction

In general, the axis of pipes belonging to a joint will not intersect exactly in one point or may not intersect at all. In order to extract a joint we calculate an approximate intersection of the pipes with 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 center of the joint as the center of mass of these midpoints, i.e.:

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

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

When there are more than one joint in the scene, 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. The pipes grouping is carried out symbolically, building the *incidence graph* of the pipes, i.e. a graph whose nodes are the pipes and in which two nodes are connected by an arc if the distance between the corresponding pipes is below a given threshold. The search for the joints corresponds to the search of the maximal complete subgraphs of the incidence graph, i.e., subgraphs that are complete and that are not contained in any larger complete subgraph. Two maximal subgraphs can have no more than one node in common (corresponding to the pipe that connect two distinct joints). The algorithm can be summarized as follows:

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. 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.

2.3. Matching

The matching phase is performed by associating both joints and pipes of one image with joints and pipes in the second image.

First we have to identify the same joint in the two images. A joint with n pipes ($n > 3$) is completely identified (for the affine geometrical properties) by the list of the $n(n-1)/2$ angles between all the pipes; it therefore sufficient to build such a list for the data image and to compare it with the analogous lists for the model image. Unfortunately, some pipes may be missing in one of the two images, therefore, such a comparison will be done also with all the sub-joints of dimension n. The output is a list of model joints or sub-joints that could be matched with the joints in the data image. The procedure is repeated by exchanging the role of data and model and the list is updated.

The second task is the matching of the pipes belonging to the same joint: for every model joint contained the above list one has to find a one to one correspondence between its pipes and the ones of a data joint. To each match a cost is associated which takes into account the differences in the matching angles and, possibly, the number of unmatched pipes. Since the objects we are dealing with are quite simple, an exhaustive method is feasible, consisting in trying all the possible matches and choosing the one with the minimum cost.

2.4. Registration

In order to align two 3D views it is sufficient to compute a rigid transformation (translation T and 3D rotation R). One image is called *model-image* and the other is called *data-image*. In our work, pairwise registration is achieved in two stages: (i) skeleton-based pre-alignment, in which a direct alignment between the joints of the two images is made; (ii) fine registration with ICP, where, starting from a couple of very close images, the registration is refined using a variation of the classic ICP algorithm .

2.4.1 Pre-alignment

After the matching phase the pre-alignment is simply a matter of computing the rigid transformation. The translation T is obtained by subtracting the model-joint position to the data-joint position. The 3D rotation R is obtained by rotating the data-joint in such a way to align two pipes of it with the two matching pipes of the model. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be the versors giving the direction of two pipes in the data set and $\mathbf{w}_1$ and $\mathbf{w}_2$ be the corresponding versors in the model set. Let $\mathbf{r}_1 = \mathbf{v}_1 \wedge \mathbf{v}_2$ and $\mathbf{r}_2 = \mathbf{w}_1 \wedge \mathbf{w}_2$. If $\mathbf{r}_1$ and $\mathbf{r}_2$ are not already aligned, the rotation that brings the plane spanned by $\mathbf{v}_1$ and $\mathbf{v}_2$ in the plane spanned $\mathbf{w}_1$ and $\mathbf{w}_2$ is given by

$$R_1(\mathbf{r}_1 \wedge \mathbf{r}_2, \mathrm{acos}(\mathbf{r}_1 \cdot \mathbf{r}_2))$$

where the first parameter is the rotation axis and the second is the rotation angle. Let $\mathbf{v}'_1 = R_1\mathbf{v}_1$ and $\mathbf{r}'_1 = R_1\mathbf{r}_1$, then the rotation

$$R_2(\mathbf{r}'_1, \mathrm{acos}(\mathbf{v}'_1 \cdot \mathbf{w}_1))$$

eventually align the axis. The total rotation is $R = R_2R_1$.

2.4.2 The ICP algorithm

The ICP algorithm is based on on the search of pairs of nearest points in the two sets, and estimating the rigid transformation which align them, assuming that the point correspondence provided by sets of closest points is a reasonable approximation to the true point correspondence. Then, the rigid transformation is applied to the points of one set, and the procedure is iterated until convergence.

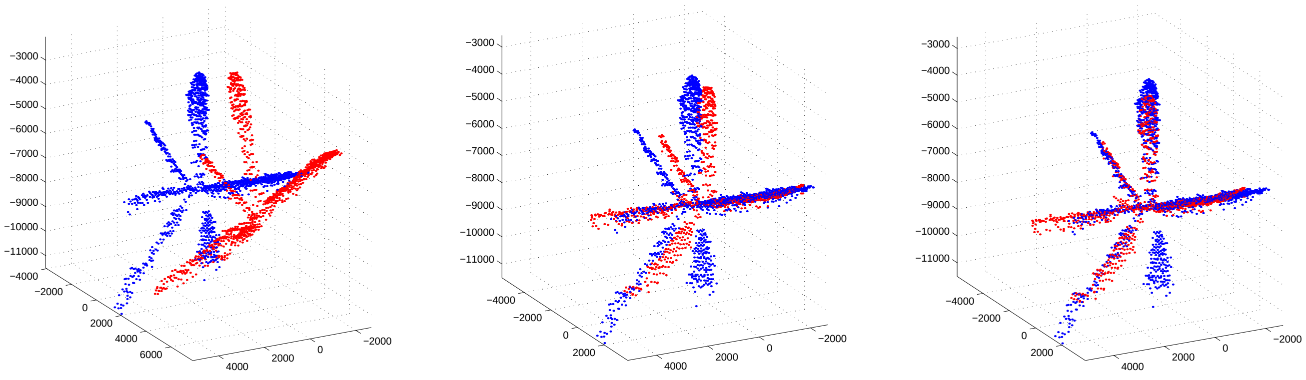
Modifications to the original ICP are now widely used to achieve accurate registration of pairs of *partially* overlapping range images [28, 26, 25]. We implemented a variation similar to the one proposed by Zhang [28], using a thresholding based on robust statistics to limit the maximum distance between closest points [15]. As pointed out by Zhang, the distribution of the residuals for two fully overlapping sets approximates a Gaussian, when the registration is good. Non-overlapped points skew the distribution of the residuals, hence the threshold on the distance must be set using a robust statistics.

Following the X84 rule [9] we discard those points whose residual differ more than 5.2 MAD (Median Absolute Deviations) from the median. The value 5.2 corresponds to about 3.5 standard deviations, which encloses more than 99.9% of a Gaussian distribution.

Please note that, having matched pipes of the data with pipes of the model, only points belonging to matching pipes are given in input to ICP.

3. Experimental Results

Our technique have been tested with both synthetic and real images of an offshore rig acquired with the Echoscope



(a) Initial pose (b) Pose after skeletons alignment (c) Pose after ICP, starting from (b)

Figure 3. Experiment 1: registration with skeletons pre-alignment.

acoustical camera. Only the results of the latter are reported here, for reasons of space.

We compared our skeleton-based pre-alignment with the well-known *Principal Component Analysis* (PCA) pre-alignment, consisting in aligning the centroids and the principal axes (ordered by the magnitude of the eigenvalues) of the two distribution of points.

When the two clouds of points are very similar the PCA alignment is sufficient to make ICP converge. As soon as new parts become visible in the second view, the principal axes are skewed by the non-overlapping parts, and PCA alignment fails, like in both the experiments reported here.

Figure 3 shows a registration using ICP with skeletons pre-alignment. [1]

The final registration is visually good, indeed the average distance between closest points is about 7 cm (sensor resolution is about 3cm). This means that ICP converged to the global minimum.

Algorithm	Average distance (mm)		Iterations
	Initial	Final	
ICP	2288.82	217.305	55
P.C. + ICP	761.169	175.386	51
Skeletons + ICP	466.32	71.8153	48

Table 1. Experiment 1: comparison of average closest points distance and ICP iterations.

Figure 4 shows the results of plain ICP and PCA registration on the same clouds of points. The first fails, as one might expect, because the two cloud of points are too much displaced. PCA fails because the data cloud have one pipe that the model does not have, and this causes the principal axes to be different. The figures in Table 1 confirm that ICP got stuck into a local minimum in both cases.

Algorithm	Average distance (mm)		Iterations
	Initial	Final	
ICP	3521.93	859.011	53
P.C. + ICP	452.149	237.106	55
Skeletons + ICP	980.336	67.848	54

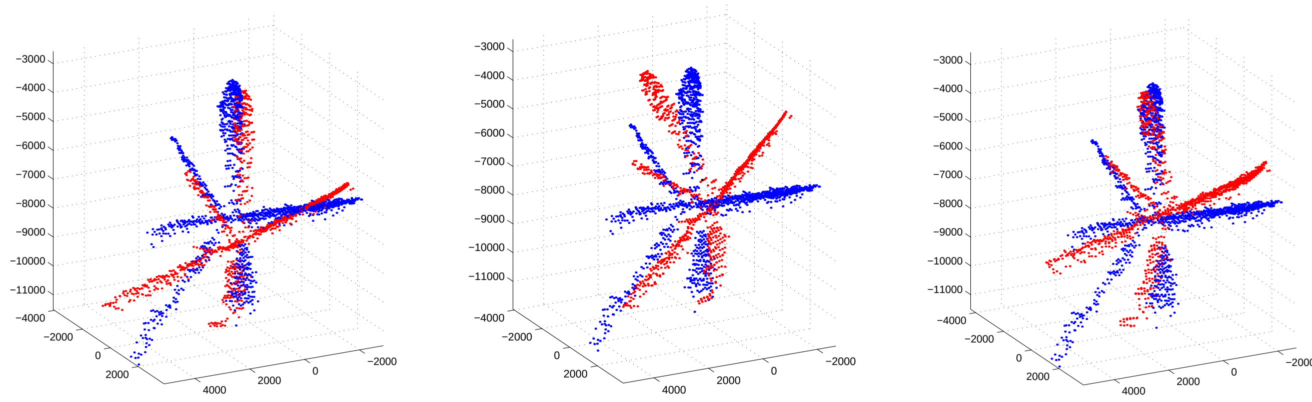
Table 2. Experiment 2: comparison of average closest points distance and ICP iterations.

Finally, Fig. 5 shows another example of successful registration using skeletons pre-alignment where PCA fails (not shown here) because a large part of the second view is not visible in the first. By analyzing the figures in Table 2 one may note that the initial average distance for PCA is lower that for skeletons, yet ICP converged to a local minimum when started from the PCA alignment. This reminds us that a good alignment is not necessarily the one with the smallest average closest points distances, because the closest points may be wrong in the first ICP iteration.

4. Conclusions

In this paper we tackled the problem of registering two clouds of points with large non-overlapping parts and gen-

[1]In the original color pictures, the two clouds of points are more distinguishable, being blue and red. Please refer to the electronic proceedings.



(a) Pose after ICP without pre-alignment

(b) Pose after PCA alignment

(c) Pose after ICP, starting from (b)

Figure 4. Experiment 1: comparison with other techniques

eral starting pose. Driven by an underwater application, we considered objects composed by pipes and we demonstrated that 3D skeletons can be successfully used to obtain a coarse alignment sufficient to make ICP converge.

Moreover, thanks to the skeletons matching, we feed the final ICP only with the matching sub-set of the data-image, thereby alleviating the outliers conditioning.

Preliminary results are satisfactory, since as soon as the skeletons are correctly matched, the ICP always converges to the global minimum. Future work will address the improvement of the skeletons matching, by making it more efficient and reliable.

Acknowledgements. This work was 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). Thanks to R. Giannitrapani for many contributions, including the skeletonization and segmentation code. Echoscope images are courtesy of Dr. R.K. Hansen of Omnitech A/S. Thank to the anonymous reviewers, whose comments helped to improve the paper.

References

[1] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau. Towards a general multiview registration technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):540–547, May 1996.

[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] G. Blais and M. D. Levine. Registering multiview range data to create 3-D computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, 1995.

[4] D. Brujic and M. Ristic. Analysis of free-form surface registration. In *Proceedings of the IEEE International Conference on Image Processing*, volume II, pages 393–396, Lausanne, Switzerland, 1996.

[5] C. B. C. Baillard, P. Hellier. Cooperation between level set techniques and dense 3d registration for the segmentation of brain structures. In *15th Int. Conf. on Pattern Recognition, ICPR '00*, pages 991–994, Barcelona (Spain), 2000.

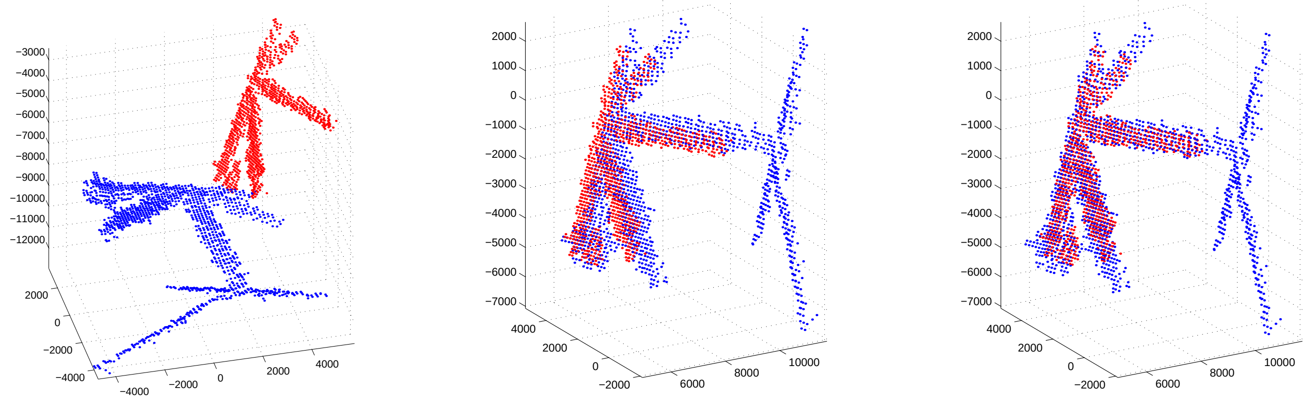
[6] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[7] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In H. Rushmeier, editor, *SIGGRAPH 96 Conference Proceedings*, pages 303–312. ACM SIGGRAPH, Addison Wesley, August 1996. held in New Orleans, Louisiana, 04-09 August 1996.

[8] D. W. Eggert, A. W. Fitzgibbon, and R. B. Fisher. Simultaneous registration of multiple range views for use in reverse engineering. In *Proceedings of the International Conference on Pattern Recognition*, pages 243–247, Vienna, 1996.

[9] F. Hampel, P. Rousseeuw, E. Ronchetti, and W. Stahel. *Robust Statistics: the Approach Based on Influence Functions*. Wiley Series in probability and mathematical statistics. John Wiley & Sons, 1986.

[10] 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.



(a) Initial pose (b) Pose after skeletons alignment (c) Pose after ICP, starting from (b)

Figure 5. Experiment 2: registration with skeletons pre-alignment.

[11] A. Hilton and J. Illingworth. Multi-resolution geometric fusion. In *Int. Conf. on Recent Advances on 3D Digital Imaging*, pages 181–188, Ottawa (Canada), 1997.

[12] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. *Computer Graphics*, 26(2):71–78, July 1992.

[13] A. Johnson, R. Hoffman, J. Osborn, and M. Hebert. A system for semi-automatic modeling of complex environments. In *International Conference on Recent Advances in 3-D Digital Imaging and Modeling*, pages 213–220, May 1997.

[14] A. Leonardis, A. Jaklic, and F. Solina. Superquadrics for segmenting and modeling range data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(11):1289–1295, 1997.

[15] V. Murino, A. Fusiello, N. Iuretigh, and E. Puppo. 3D mosaicing for environment reconstruction. In *International Conference on Pattern Recognition (ICPR 2000)*, volume 3, pages 362–366, Barcelona, September 2000. IAPR, IEEE Computer Society Press.

[16] V. Murino and R. Giannitrapani. Three-dimensional skeleton extraction by point set contraction. In *IEEE International Conference on Image Processing*, pages 565–569, Kobe, Japan, October 1999.

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

[18] B. P.K.Saha and D. Majumder. A new shape preserving parallel thinning algorithm for 3d digital images. *Pattern Recognition*, 30(12):1939–1955, 1997.

[19] K. Pulli. Multiview registration for large data sets. In *IEEE Int. Conf. on 3-D Imaging and Modeling, 3DIM '99*, pages 160–168, Ottawa (Canada), 1999.

[20] K. Pulli, T. Duchamp, H. Hoppe, J. McDonald, L. Shapiro, and W. Stuetzle. Robust approximate meshes from a collection of range maps. In *3DIM97*, pages 8 – Object Modeling, 1997.

[21] T. R.A.Grupen and C.D.Hansen. Apparent symmetries in range data. *Pattern Recognition Letters*, 7:107–111, 1988.

[22] M. Sablatnig, R .and Kampel. On registering front- and backviews of rotationally symmetric objects. In F. Solina and A. Leonardis, editors, *8th Intl. Conf. on Computer Analysis of Images and Patterns*, number 1689 in Lecture Notes in Computer Science, pages 339–346. Springer, 1999.

[23] M. Soucy and D. Laurendeau. A general surface approach to the integration of a set of range views. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 4(17):344–358, April 1995.

[24] A. Stoddart and A. Hilton. Registration of multiple point sets. In *Int. Conf. on Pattern Recognition*, pages 40–44, Vienna, 1996.

[25] E. Trucco, A. Fusiello, and V. Roberto. Robust motion and correspondences of noisy 3-D point sets with missing data, 1998. Submitted for publication in *Pattern Recognition Letter*.

[26] G. Turk and M. Levoy. Zippered polygon meshes from range images. In A. Glassner, editor, *Proceedings of SIGGRAPH '94 (Orlando, Florida, July 24–29, 1994)*, pages 311–318. ACM SIGGRAPH, ACM Press, July 1994.

[27] R. Whitaker. A level set approach to 3d reconstruction from range data. *International Journal on Computer Vision*, 29(10):203–231, October 1998.

[28] Z. Zhang. Iterative point matching of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119–152, 1994.