THREE-DIMENSIONAL SKELETON EXTRACTION BY POINT SET CONTRACTION

Riccardo Giannitrapani

Dipartimento di Matematica e Informatica
University of Udine
Via delle Scienze 206, 33100 Udine, Italy
Email: gtrapani@dimi.uniud.it

Vittorio Murino

Dipartimento Scientifico e Tecnologico
University of Verona
Ca' Vignal 2, Strada Le Grazie, 37134 Verona, Italy
Email: murino@sci.univr.it

ABSTRACT

In this paper, a skeleton extraction method for unstructured three-dimensional data is presented. The algorithm is based on a point set contraction procedure and is proved to be robust to noise and coarse resolution of original data. Preliminar experiments performed on both synthetic and real data are provided, showing the goodness of the proposed method.

1. INTRODUCTION

It is widely known the importance of the skeleton [1, 2] for shape analysis and object recognition. Roughly speaking, the skeleton is essentially a simplified structure summarizing the general form of an object. While a large literature addresses the problem of skeleton extraction from (two-dimensional, 2D) binary and gray-level images, a few researchers approached the same problem when sets of three-dimensional (3D) points are considered. Nowadays, three-dimensional images are become common as intensity images and this work proposes a method to estimate the skeleton from a cloud of 3D point. Actually, skeleton definition and extraction from a 3D sparse image, i.e. a distribution of points in $\mathbb{R}^3$ is still a challenging and quite novel 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 bidimensional patches is developed to extract axes of symmetry [3] 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 [4].

In this paper, a technique, that does not need the subdivision of space in regular cells, useful to extract the 3D skeleton from distributions of 3D points acquired by an acoustic camera sensing an underwater scene is proposed. Still related to the considered application, the reconstruction of a 3D underwater scene of an off-shore structure composed by pipes and joints is also proposed.

In Section 2, the main algorithm is described and the involved parameters are discussed. In Section 3, tests are performed on both synthetic and real underwater 3D images, and related results are shown, also discussing possible improvements of the techinque. Finally, some conclusions are drawn in Section 4.

2. SKELETON EXTRACTION

Three-dimensional data are obtained by a high resolution acoustic camera, the *Echoscope* [6] which provides a set of 3D points. The acoustic camera is formed by a two-dimensional array of transducers sensible to signals backscattered from the scene previously insonified by a high-frequency acoustic pulse. The whole set of raw signals is then processed to estimate signals coming from fixed steering directions (called beamsignals) and to attenuate those coming from other directions. Assuming that the beamsignals represent the responses of a scene from a 2D set of (steering) directions, a 3-D point set can be extracted detecting the time instant (t^*) at which the maximum peak occurs in each beamsignal. Images are formed by 64 × 64 3D points ordered according to an angular relation, as adjacent points correspond to adjacent beam signals. Their coordinates are expressed in a 3-D reference frame attached to the sensor. This kind of images are heavily affected by speckle noise, that may degrade the image quality and hence the reliability of the 3D measure. Further, due to the wavelength of the acoustic sound in water, the range resolution is quite coarse (a few centimeters in the best conditions), and also this makes more difficult the extraction of geometric characteristics from such kind of data. The subsequent problem we faced is the lack of a definition of 3D skeleton (see for exam-

ple [5]) adaptable to a distributions of points in a 3D space belonging to surfaces instead of volumes. In fact, general range images acquired by laser range finders or data obtained using an high resolution acoustic camera represent geometric information derived by object surfaces. In this case, it is not possible to use the concept of the locus of the center of maximal spheres contained in the object [2, 5], since we have no solid object to consider, but only 3D surfaces. Therefore, a rough definition of skeleton is first proposed, i.e., a skeleton is 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 and 4) have to be invariant to 3D rotation. The third point requires, essentially, that *the operations performed on the distribution to extract the skeleton preserve such topological information* like the number of "holes" in the distribution and their relative position. 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.

In short, the method for skeleton extraction is performed by applying the following phases. Let us define the 3D image as I, i.e., the ensemble of points in $\mathbb{R}^3$

$$I = \{\mathbf{a}_i\} \quad i = 1, ..., N$$

where $\mathbf{a} = (x, y, z)$.

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

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

$$\mathbf{b}_i^R \equiv \frac{\sum_{\mathbf{a}_j \in O_i^R} \mathbf{a}_j}{\dim\{O_i^R\}}$$

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

Let us define the *interior* λ_i^R for the point i in the following way:

$$\lambda_i^R \equiv \frac{R - |\mathbf{a}_i - \mathbf{b}_i^R|}{R}$$

One can easily verify that

$$0 \leq \lambda_i^R \leq 1 \quad \forall i$$

The interior is a measure of how much a point is "inside" the object specified by I. If $\lambda_i^R \simeq 1$, the point is inside a homogeneous zone of radius at least equal to R; on the contrary, for $\lambda_i^R \simeq 0$, the point is very

near to a border of the three-dimensional distribution of points.

Defining the following image transformation:

$$I \longrightarrow I^R = \{\mathbf{b}_i^R\} \tag{1}$$

and indicating with the symbol $I^{R,n}$ the iterative application of it for n times, our skeleton extraction is simply the construction of the image $I^{R,n}$ for a suitable choice of R and n.

The overall effect of this transformation is to shift points on the border, i.e. points with a low λ^R, toward the center, while leaving points well inside an object, i.e., points with a almost unitary value of λ^R, unaltered. The iterative application of such a procedure tends to shift all the points of the distribution towards its skeleton. Such an algorithm may be seen from a physical point of view like a short range interaction process between physical points.

The choice of the number of iterations, n, is not easy, and actually it is performed with an *ad hoc* criterion. Iterations are stopped when the mean value of λ (computed considering all of the points of the distribution) is under a fixed threshold experimentally determined. In this way, the set contraction algorithm is stopped when almost all points are mostly contracted, meaning that the distribution is very thin. Obviously, the iterations must be stopped before interesting features of the distribution are lost due to the contraction: this is a difficult problem and must be addressed by an adaptive local threshold criterion that is under investigation (see next section). For the acoustic images, the problem does not seems to rise and our empirical method works quite well and this is due to the tubular form of the surfaces observed.

The choice of the free parameter R is very important. If it is too small, with respect to the average mutual distances of the points, the distribution would likely collapse in a certain number of disconnected punctual regions, preventing the extraction of the skeleton. If it is too large, the value of λ may be small also for points well inside the object and the skeleton will still not be properly extracted. The parameter R plays a role similar to the dimension of the structuring element in the classical mathematical morphology [1]. It is straightforward to verify that the skeleton so extracted is invariant for three dimensional rotation (condition that is never exactly satisfied for skeletons of a 2D bitmap image[1]).

In summary, our algorithm is the following:

---

[1]This is obviously trivial since we use the continuous Euclidean metric that is invariant under rotation, while for a 2D image there is a lattice structure that imply a discrete geometry that is never invariant under a generic rotation.

1. Choice of the free parameter R

2. Contraction of the set of points using eq. (1)

3. Calculation of the mean value of λ, $\overline{\lambda}$

4. *If* $\overline{\lambda}$ < *threshold* stop *else* goto 2

The skeleton presented here is just a particular distribution of points, i.e., a transformation of original data. It is possible to extract from it a symbolic representation in the form of a graph with two kind of nodes: *branch* nodes and a *joint* nodes. The former is a node representing a subset of points belonging to the same branch in the skeleton; the latter represents a subset of points belonging to joints in the skeleton. The symbolic skeleton is extracted from the skeleton simply by looking at the neighbours of each point. If such a neighbour has some geometrical properties, mainly it is elongated, than the point belongs to a branch, otherwise it belongs to a joint. In other words, if the neighbours of a point are distributed along a quite straight segment than the point belongs to a branch of the skeleton, otherwise it is in a zone where more branches intersect and so it is a joint point. This check for the point i is easily performed by calculating the inertial tensor of the set O_i^r, that is the set of neighbours that are at a distance less than r (also in this case, this parameter has to be checked experimentally). The structure of its eigenvalues gives information on the shape of the set. A symbolic skeleton example will be shown in the following section.

3. RESULTS AND OUTLOOK

We have tested our technique on different kind of synthetic and real 3D images. Since we are interested in the analysis of tubular underwater structures, our algorithm behaves well with such kind of images, as one can see from the following examples.

Two synthetic examples are shown in Figs. 1 and 2, where skeletons are correctly derived. It should be noted that such preliminar experimental results show also that the homotopy group is preserved by the algorithm.

Using a priori knowledge about the structure represented by the $3D$ image and the sensor resolution, it is possible to estimate proper values for R and n aimed at improving the extraction of a accurate skeletons. Figures 3 and 4 show two real examples where, despite the noisy nature of data, skeleton are correctly identified.

However, it should be noted that a single real joint can be splitted in two near joints, as is the case of the example in Fig. 3, due to the noisy nature of data.
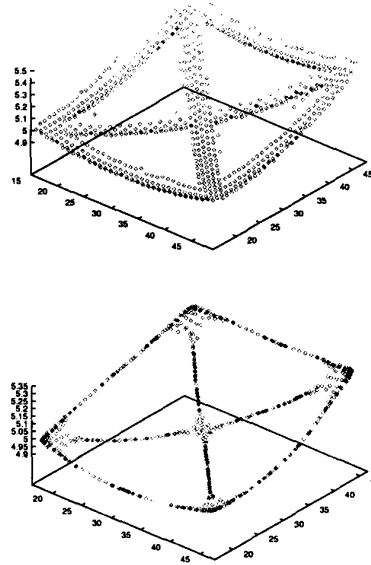

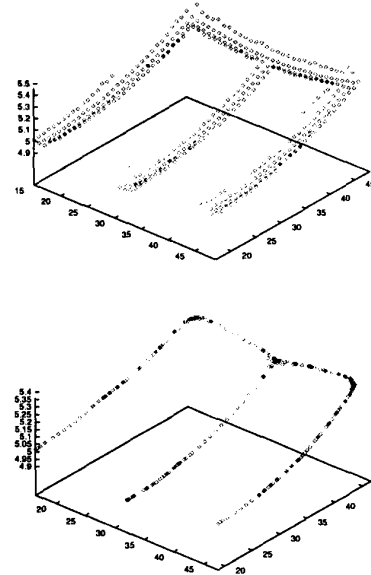Figure 1: Synthetic 3D data (up, original data; down, skeleton extracted using R = 3 cm and n = 2).

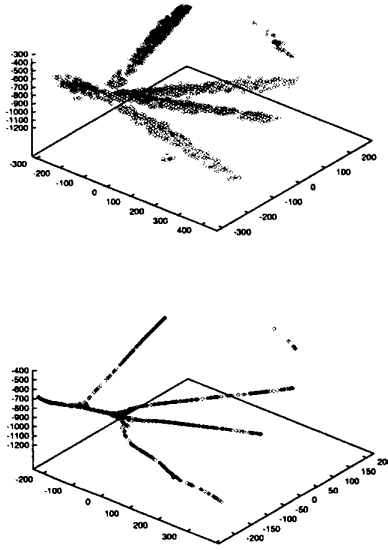Figure 2: Synthetic 3D data (up, original data; down, skeleton extracted using R = 3 cm and n = 2).

Figure 3: Real 3D acoustic underwater image representing a joint formed by 4 pipes (up, original data; down, skeleton extracted using R = 80 cm and n = 2).
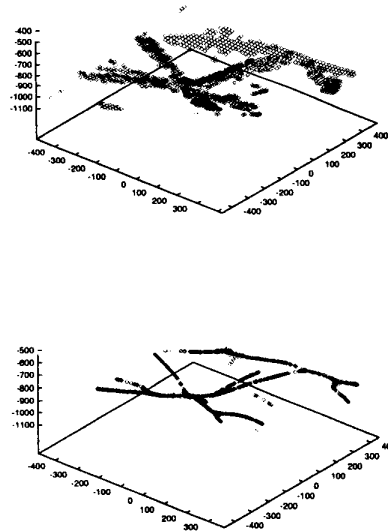

Figure 4: Real 3D acoustic underwater data representing two joints between pipes (up, original data; down, skeleton extracted using R = 80 cm and n = 2).
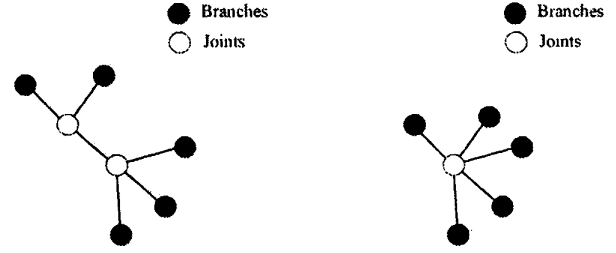

Figure 5: The symbolic skeleton of the joint of figure 3 is depicted (left) along with its reduction (right) due to the a priori knowledge of the scene observed.

This is reflected in the symbolic skeleton derived from it, as depicted in Fig. 5.

A suitable reduction procedure for the symbolic skeleton, using a priori information, can be, in that case, the elimination of every branches whose set of associated points does not satisfy some geometrical constraints. In the proposed example, we impose that these sets must have a tubular shape (this is easily verified by means of inertia moments' estimation). If the branch connecting two joints does not satisfy such a constraint, it is eliminated, so implying the merging of the joints, for the rules discussed in the preceding section. The reduced graph has now the right structure to represent the joint observed in the scene (see Fig. 5).

The same problems of false branches formation are visible also in Fig. 4 and can be dealt with the same symbolic manipulation.

We have used the skeleton and its associated symbolic graph to reconstruct tubular structures from acoustic images in an underwater environment. As an example, the reconstruction, obtained with inertia moments' estimation, of the pipes belonging to the joint in Fig. 3 is depicted in Fig. 6 (the image is from a different point of view with respect to Fig. 3 and is obtained with a description in VRML language and it can be fully manipulated in 3D).

We are now working to a procedure that is automatically able to set the free parameter R depending on 3D distribution characteristics (like density of points, moments, etc.). Actually, such selection should be local, so that the skeleton will be extracted in different ways in different image regions, allowing the application of the method to any kind of data without prior knowledge of the scene investigated. Moreover, we are also planning the implementation of an estimator to evaluate the "quality" of the skeleton obtained, in order
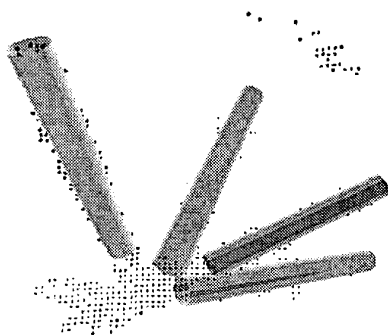
Figure 6: Final reconstruction of a joint obtained with the skeleton of the distribution.

to have an automatic stop criterion of the algorithm. Several ideas are currently investigated to solve these problems.

4. CONCLUSIONS

Apart the particular application discussed in this paper, the reconstruction of 3D skeletons from unstructured data is an interesting and challenging problem. The definition of skeleton and the design of an algorithm able to derive it from 3D point distributions having no particular constraint of geometric regularity, resolution, and affected by strong noise are not straightforward tasks.

In this paper, one possible definition and one possible algorithm for skeleton extraction have been proposed and discussed. Our work was mainly devoted to the particular application concerning the reconstruction of underwater scenes by acoustic 3D data analysis, and a more extensive and deeper theoretical analysis is in progress to validate the method.

5. REFERENCES

[1] I. Pitas, A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Academic Press, San Diego, USA, 1990.

[2] C.Arcelli and G.Sanniti Di Baja, "Euclidean Skeleton Via Centre-of-Maximal-Disc Extraction", *Image and Vision Computing*, vol. 11, pp. 163-173 (1993).

[3] R.A.Grupen, T.C.Henderson and C.D.Hansen, "Apparent Symmetries in Range Data", *Pattern Recognition Letters*, vol. 7, pp. 107-111 (1988).

[4] P.K.Saha, B.B.Chaudhuri and D.Dutta Majumder, "A New Shape Preserving Parallel Thinning Algorithm for 3D Digital Images", *Pattern Recognition* vol. 30, n. 12, pp. 1939-1955 (1997).

[5] D.Attali, A.Montanvert, "Computing and Simplifying 2D and 3D Continuous Skeletons", *Computer Vision and Image Understanding*, Vol. 6, No. 3, pp. 261-273 (1997).

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