Image Segmentation by Label Anisotropic Diffusion

Raphaëlle Chaine[1] and Saïda Bouakaz[1]

LIGIM, bât 710
Université Claude Bernard LYON 1
69622 Villeurbanne cedex, FRANCE

Abstract. Weighing the difficulties of a symbolic description of 3D surface based scattered data, this article propounds a formalisation of the segmentation in discrete labelling terms. Global consistency of the result is expressed as a constraint satisfaction problem. To solve this problem, the method we present is based on an anisotropic diffusion principle along two structures respectively denoted minimal and maximal escarpment trees. These structures are drawn from the graph theory. Novel aspect of our method is its ability to work on non organised points and to detect arbitrary topological types of features, as crease edge or boundaries between two smooth regions. The proposed approach makes possible an hybrid segmentation, involving the duality between regions and boundaries. The method has proven to be effective, as demonstrated below on both synthetical and real data.

1 Introduction

Most vision systems need a symbolic description to represent data. In particular, the goal of three dimensional reconstruction is the production of a geometric model from discrete information. This reconstruction entails the analysis of an object surface, the topology of which is unknown. Segmentation represents one of the main steps of this analysis. It aims at extracting elements relevant to the representation of an object surface, based on a coherence criteria. More precisely, the segmentation consists in the subdivision of a points set into subsets that are homogeneous according to a common property [18, 3]. The method and the criterion used to make this partitioning depend on the outer description and treatments we intend to do on the data. The problem of the segmentation has long been treated: an excellent survey can be found in [13] and some new methodologies are underlined in this paper. In particular, Hoover et al [10] propose an experimental evaluation of range image planar segmentation methods. Most of these existing segmentation techniques are available on regular images only. The data at our disposal consists of a non-organised collection of points on or near a given surface in R^3. This kind of data occurs in diverse application fields, such as range scanning an object from multiple view points. Gathering of the data gives a non organised points set and our process cannot use any regularity in the distribution.

Points must be treated with the intuitive idea they belong to the same surface. The information on a point of the set, such as coordinates or locally inferred properties, remains quantitative and it only becomes meaningful when confronted with the information relative to other points. The emergence of qualitative information also involves the use of contextual elements. Formally, this idea corresponds to consistency relations modelled by topological and geometrical constraints to be satisfied.

Generally, the notion of geometrical homogeneity can be expressed as the identity or similarity of differential properties up to a given order of derivation. Among others, Besl & Jain [5], Sander & Zucker [15, 16], R.Hoffman & A.K.Jain [9] and Faugeras [7, 8] use the information of differential properties up to the second order to characterise the points in their process. A study of the curvature calculus stability can also be found in [1]. As far as J.P.Thirion [17] is concerned, he also uses third order differential properties.

At an higher level, at the manner of Mohan & Nevatia [12], we want to coordinate local constraints to issue a single consistent interpretation of the image. Berthod [4] classifies the different parts of one fact to be explained. Local level involves observations and representations relative to distinct elements, for example points coordinates and/or differential properties. Then, there are constraints on each element property. These constraints express the belonging of the data to a given universe. This semantic information makes possible to improve representation derived from local information.

The main contribution of our work is to show that a contextual process can be achieved by an anisotropic diffusion of label. In this paper, we formalise the segmentation problem as a constraint satisfaction problem, the solution of which is a labelling function. We underline the importance of an anisotropic resolution of this CSP and we introduce two original structures denoted minimal and maximal escarpment trees. They embody the notion of anisotropy on the points set. The segmentation is made by diffusion/inhibition of label along these structures. Finally, we present several results, to demonstrate the ability of our method.
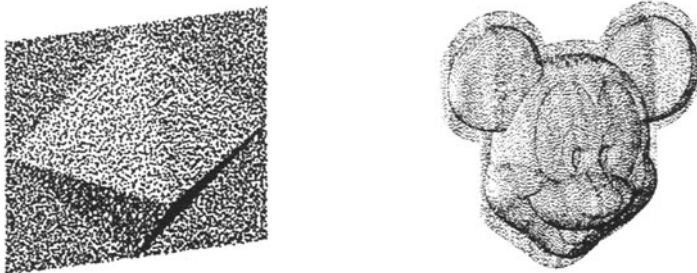
Fig. 1. 3D scattered data. Synthetical noisy set "pyramid" and real object "mouse" (Gracefully supplied by MATRADATAVISION). Points do not lie on a regular grid and density is not constant.

2 Problem Modelisation

Segmentation results in a partition of a points set. After this step, a point belongs to one and only one region. Then, segmentation can be tackled as a discrete labelling problem with one label for each region. As an equivalence relation, the equality of label between two points is a transitive relation. In practice, we use homogeneity criteria to gather some points under the same label. These criteria constitute a geometrical transcript of labels compatibility between two points. However, they can be non-transitive: Two points can have their labels compatible with a third point label, without being compatible with each other. The use of a non-transitive criterion to issue a transitive result can mask a label incompatibility. Intrinsic difference between the homogeneity of two points and the fact of belonging to a same region induces difficulties in transposing one notion into the other. These difficulties are inherent to the problem of information compression by discrete labelling . Thus, the original problem is turned into the determination of a labelling which complies with the numerical expression of homogeneity on a maximal number of neighbouring points couples. It is also a constraint satisfaction problem. Let us introduce:

- $X = \{x_1, \cdots, x_i, \cdots, x_n\}$: Set of points to be processed.
- Γ : Neighbouring graph. Γ is a non-oriented graph consisting of n nodes $1, \cdots, i, \cdots, n$. The presence of an arc between nodes i and j indicates closeness between x_i and x_j. In the model instantiation presented here, a triangulation provides the neighbouring relation.
- V_i : Set of indexes of x_i neighbours.
- l : Labelling function

$l : X \longmapsto I\!N$
$\quad x_i \longmapsto l(x_i)$

- Given $x_i \in X, x_j \in V_i$

$C_{ij} : I\!N \times I\!N \longmapsto I\!R$
$\quad a, b \qquad \longmapsto C_{ij}(a, b)$

$C_{ij}(a, b)$ measures the possibility for a and b to be the labels at points x_i and x_j, considering the two points homogeneity.

Given an l labelling, each arc (i, j) in Γ is assigned the value $C_{ij}(l(x_i), l(x_j))$. Desired labelling maximises the total cost of Γ. This total cost has the meaning of a consistency function. In this paper, our expression of C_{ij} brings in differential properties up to the first order of derivation only, but further derivation can be involved. Let n_i be the normal vector of the surface at point x_i. The condition $n_i.n_j > \epsilon$ where ϵ is one parameter of the system, expresses similarity between normal vectors at points x_i and x_j. It is the criterion we use to make our segmentation. In accordance with the classification of information, ϵ corresponds to a symbolic parameter of the semantic information.

$$C_{ij}(a, b) = (n_i.n_j - \epsilon) * \delta(a, b) \quad \text{where} \quad \delta(a, b) = \begin{cases} 1 \text{ if } a = b \\ -1 \text{ else } . \end{cases}$$

The determination of l is locally constrained by the maximisation of $C_{ij}(l(x_i), l(x_j))$. A negative sign for C_{ij} means a constraint violation. The adequacy of l with all

these constraints is modelled by a consistency function $F_X(l)$. This function integrates local adequacy of a labelling over all the points. We try to maximise this function. $F_X(l)$ permits to determine how much a labelling consistent and appropriate is, with our semantic interpretation of the shape of an object. This function is defined by:

$$F_X(l) = 1/2 \sum_{i=1}^{i=n} \sum_{j \in V_i} C_{ij}(l(x_i), l(x_j)) \ . \tag{1}$$

Maximisation of such a function is a combinatorial problem which is NP complete. In its current form, the problem formalisation imposes every point label to be homogeneous with those of the neighbours. Is it necessary to test consistency on the whole neighbourhood? Obviously, it is prohibitive to take so many constraints into account. What is more, they do not have the same level of importance. Intuitively, the existence of particular points, having a topologically privileged position to transmit their labels to other points seems attractive. We are looking for the points that are most qualified to diffuse their labels, keeping in mind both contrary requirements of singularity detection on one hand, and minimisation of the total number of regions on the other. We want to get rid of the current isotropy of our formalisation.

3 Anisotropic Model of Diffusion

The idea of an anisotropic processing has already been tackled in some fields, such as grey level images restoration. The filters used in a multiscale process entail an undifferentiated integration in every direction. The redundancy of information involved by the consideration of different images corresponding to the same image observed at different scales, can be dynamically expressed by the heat equation. Multiscale process is equivalent to the proposition of an isotropic evolution model of the image in function of a simplification scale parameter. The general idea of Perona & Malik [14] and Alvarez, Lions, Morel [2] consists in the transformation of this model to make diffusion anisotropic. Close to a discontinuity, diffusion must be reduced, or even blocked. In the solutions they propose, diffusion is inhibited in the gradient direction, with the add of a viscosity bias in the equation of diffusion.

These ideas correspond to the consideration of directional contextual constraints in diffusion processes. The introduction of anisotropy in our process is directly inspired by the evolution models mentioned above. It consists in label propagation along low curvature directions and the possibility of label blocking in high curvature directions. As far as we are concerned, we do not have a regular grid structure adapted for differential filters decomposition. Therefore, it is difficult to determine preferential directions exactly. What is more, finer information of second derivatives is not involved in the level of constraints we consider at present. So, we estimate directions of propagation from the sole information of first derivatives.

4 Directed Normal Vector Estimation

At one point, the estimation of the normal vector is carried out using the results of principal component analysis. Let M_i be the covariance matrix of points $\{x_j, j \in V_i\}$ in the neighbourhood of x_i. The eigenvector associated with the smallest eigenvalue of M_i gives the less discriminant direction of the points set. It corresponds to the normal direction. For each normal vector, two orientations are possible. It is necessary to coordinate orientations over the points set. The algorithm proposed by Hoppe et al [11] permits a surface based normal orientation. They consider a particular spanning tree of Γ and they propagate a normal vector orientation from an initial point over this structure. In following chapters, we draw our inspiration from this geometrical use of a minimum spanning tree.

5 Minimal and Maximal Escarpment Trees Introduction

Once a global consistency of inferred normal vectors is achieved, we can interpret the points set as a differential manifold. At every point of a regular enough oriented surface, an infinity of curvatures can be computed. Each of them corresponds to a direction in the tangent plane [6]. Umbilic points excepted, two privileged directions and two privileged values are enough to describe all the curvatures. These are principal directions, each of them corresponding to one extremal curvature value. By integration of each principal direction field, we obtain lines of curvature. At every point, these curves are tangent with one of the principal directions. Lines of low and high curvature are respectively lines of low and high (energy of) deformation. We have said that the set of points is represented by a graph Γ, expressing spatial proximity notion. Considering all the spanning trees included in Γ, we want to introduce an extremality notion, based on a similar notion of deformation. In their study, Hoppe et al [11] use the same idea of an optimal spanning tree. The standard of attainment they use to define this tree does not cover information of normal vector orientation. This tree is a good tool to achieve a global orientation of the points set, but it is not the best candidate for a deformation minimisation. The optimal trees we introduce involve the information of orientation:

- We denote minimal escarpment tree, a minimal spanning tree of the neighbouring graph Γ, each edge of which being assigned the cost $-n_i.n_j$.
- We denote maximal escarpment tree, a minimal spanning tree of Γ, each edge of which being assigned the cost $n_i.n_j$ (see Fig2).

Compared with the other spanning trees, the minimal escarpment tree constitutes a path avoiding geometrical singularities. Locally, this path sometimes tallies with lines of low curvature. It is the least sinuous spanning tree considering the only information of oriented normal vectors. In a similar way, the maximal escarpment tree has the properties of a path which tends to go across singularities. Locally, it corresponds to some extent with lines of high curvature.

We have underlined that the choice of an anisotropic process results in the introduction of new directional constraints. Low curvature lines seem to embody the best directions for a diffusion process. Still, all these lines must be covered to include the totality of the object. It is prohibitive. The minimal escarpment tree makes it possible to bypass this problem. It presents both advantages of covering the points set as a whole and not entailing the computation of second order differential properties.

In return, the only use of this tree for label diffusion is not enough. Skirting of singularities can be brought about, and two regions separated by a singularity can be assigned the same label. To avoid this problem, label diffusion is carried out along the maximal escarpment tree. In a way similar to the directional inhibition of diffusion in image restoration models, the purpose of this path is not to enable propagation through a first order singularity and to prevent the skirting of this singularity.

Fig. 2. Minimal and maximal escarpment trees over the "mouse" ear. The visible differences between both structures are located on edges.

6 Diffusion Algorithm and Results

At the beginning of the segmentation process, the points do not have any prior label. Labelling is obtained using an iterative procedure, the stopping of which is caused by the stabilisation of the number of regions. Every iteration corresponds to a diffusion or an inhibition of label along the maximal escarpment tree.

- Propagation of label is allowed each time the condition $n_i.n_j > \epsilon$ is true. ϵ is the threshold of the cosinus value of the angle between two normal vectors, above which the points can be considered as homogeneous.

- If the propagation of label is inhibited, there is a change of label. The next label is chosen in the minimal escarpment tree. It is the label of the most homogeneous neighbour in this tree. If there is no candidate point, either because they do not have a label yet, or because no point is homogeneous enough, then a new label is created.

By iteration of this procedure, labels are modified at each step and they are replaced by labels more consistent with the notion of curvature over the neighbourhood. The number of labels diminishes bit by bit until stabilisation.

At the end of this step, we can consider two kinds of boundaries :

- The boundaries which underline a real non-homogeneity between two regions.
- Artificial boundaries which are created at the leaves of the maximal escarpment tree by lack of propagation beyond these leaves (see Fig3).

A second step is necessary to merge the regions on both sides of these artificial boundaries. This second phase is operated with a merging process along the minimal escarpment tree. It reflects a will to merge along the lines of low curvature. This merging phase is operated twice, with two different values for ϵ. The first value of ϵ is a high, strict value to prevent the regions from filtering. At the end of this phase, many tiny regions remain in the proximity of real boundaries. A second merging phase with a lower value of ϵ makes it possible to include these tiny regions in larger regions, we call it the merging phase with hysteresis.
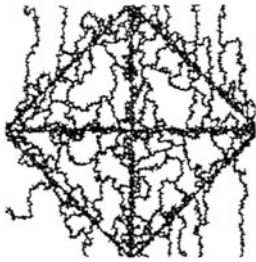
(a) Points set "pyramid". Result of the first diffusion step along the max. esc. tree.
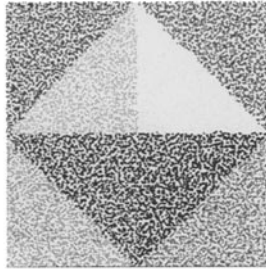
(b) Points set "pyramid" final segmentation. We observe an adequation between the result and original model ground truth.
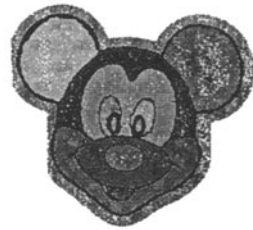
(c) Points set "mouse" segmentation. The thickness of boundaries around eyes and ears encloses an over segmentation.

Fig. 3. Segmentation results

7 Conclusion and Prospects

This paper proposes an intrinsic segmentation of 3D scattered data based on an anisotropic label diffusion. Although the process is not directly based on differential equations, it can be considered, to a certain extent, as a discrete transposition of ideas recently developed for grey level images restoration. We point out that an analogy can be made with variational problems. The introduced structures allow the system to be constrained properly. Particularly, they constitute a natural way to realise duality between regions and boundaries. We intend to evaluate the power of our method with the methodology proposed by Hoover et al [10]. At present we already have satisfactory results. The emergence of some regions can be observed, the consistency of which is compatible

with tricky notions of homogeneity and similarity of the human vision. Yet, we find out the presence of many regions, the dimensions of which is too small to give them a topological sense. This over segmentation is partly due to the fact that the neighborhood is based on a Delaunay triangulation of the projected points set. A 3D surface triangulation can provide better results.

References

1. ABDELMALEK, N.N.: Algebraic error analysis for surface curvatures and segmentation of 3D range images. Pattern Recognition **23** (1990) 807–817
2. ALVAREZ, L., LIONS, P.L., MOREL, J.M.: Image selective smoothing and edge detection by non linear diffusion. SIAM J Numer Anal **29** (1992)
3. ARMAN, F.,AGGARWAL, J.K.: Model-based object recognition in dense-range images - A review. ACM Computing surveys **25** (1993)
4. BERTHOD, M.: L'amélioration d'étiquetage: Une approche pour l'utilisation du contexte en reconnaissance des formes. PhD thesis, Université Pierre et Marie Curie Paris VI (1980)
5. BESL, P.J., JAIN, R.C.: Segmentation Through Variable-Order Surface Fitting. IEEE Transactions on pattern analyse and machine intelligence 1ω (1988)
6. DO CARMO: Differential Geometry of Curves and Surfaces. Prentice Hall, Englewood, N.J. (1976)
7. FAUGERAS, O., HEBERT, M.: The representation, recognition and locating of 3D objects. Int. J. Robotics Res **5** (1986) 27–52
8. FAUGERAS, O.: Three dimensional computer vision, a geometric view point. MIT Press (1993)
9. HOFFMAN, R., JAIN, A.K.: Segmentation and classification of range images. IEEE Trans PAMI **9** (1987)
10. HOOVER, A., JEAN-BAPTISTE, G., JIANG, X., FLYNN, P., BUNKE, H., GOLDGOF, D., BOWYER, K., EGGERT, D., FITZGIBBON, A., FISHER, R.: An Experimental Comparison of Range Image Segmentation Algorithms. IEEE Trans PAMI July (1996) 1–17
11. HOPPE, H., DEROSE, T., DUCHAMP, T., MCDONALD, J., STUETZLE, W.: Surface Reconstruction from Unorganized Points. Computer Graphics (1992)
12. MOHAND, R., NEVATIA, R.: Using Perceptual Organization to Extract 3-D Structures. IEEE Tr PAMI **11** (1989)
13. PAL, N.,PAL, S.: A review on image segmentation techniques. Pattern Recognition **26** (1993)
14. PERONA, P., MALIK, J.: Scale space and edge detection using anisotropic diffusion. IEEE Comput. Soc. Workshop on Comput. Vision (1987)
15. SANDER, P.T., ZUCKER, S.W.: Inferring Surface Trace and Differential Structure from 3D Images. IEEE Transactions on pattern analyse and machine intelligencence **12** (1990)
16. SANDER, P.T., ZUCKER, S.W.: Singularities of principal direction fields from 3D Images. IEEE Transactions on pattern analysis and machine intelligence **14** (1992)
17. THIRION, J.P.: The extremal mesh and the understanding of 3D surfaces. Technical Report INRIA (1993)
18. ZUCKER, S.W.: Survey region growing; Childhood and adolescence. CVGIP **5** (1976) 382–399