A generic methodology for partitioning unorganised 3D point clouds for robotic vision

Nicolas Loménie

A generic methodology for partitioning unorganised 3D point clouds for robotic vision

Nicolas Loménie
Systems of Intelligent Perception lab (SIP - CRIP5), University Paris V, Paris, France
E-mail: lomn@math-info.univ-paris5.fr

Abstract : Range image segmentation has many applications in computer vision areas such as computer graphics and robotic vision. A generic methodology for 3D point set analysis in which planar structures play an important role is defined. It consists mainly of a specific K-means algorithm which is able to process different shapes in cluster. At the same time, within geometric and topologic considerations, a set of application-driven heuristics is designed. This helps to find out the right number of structures in point sets in order to give a good visualization and representation of a large scale environment without *a priori* models. Our aim is to propose a simple and generic frame for 3D scene understanding. Tests were realised on different types of environment data: natural and man-made. This research project has been realized with EADS (French Air Space Society). **Keywords:** Fuzzy clustering, 3D reconstruction and scene analysis, range image segmentation, environment modeling, stereovision.

1 Introduction

Given a point set and an application, how should we process the point set in order to complete the application? The answer to this question is obviously of great interest in the field of computer vision and pattern recognition. Anyone who attempts to solve the problem refers himself to a spectific application and so to a specific kind of cloud of points. For instance, a now famous contest [3] consists in comparing the results of the segmentation of range images into planar regions. In this case, 3D point sets are structured on a regular 2D grid. Unfortunately, in stereoscopic vision problems of reconstruction, unorganised, inhomogeneous in density 3D point sets are obtained while laser range systems provide very dense 3D point sets [10]. For each of these applications, authors have developped a specific algorithm to process point sets. In [6], the 4-connexity information is at the basis of the algorithms, making them unable to process unorganised point sets. In [1], the diversity in shape, density and homogeneity of clusters to be formed is not involved in the algorithms, making them unable to process stereoscopic data. In [4], the range image is integrated with the reflectance map. Besides, the only systematic study of the performance of range image segmention algorithms [3] uses only planar patches.

In this paper, we will define a generic methodology in order to make partition in any 3D unorganised point set where planar structures are important. For this purpose, the next section describes a very robust clustering method introduced by [5]. But we propose to make it unsupervised in a more generic way. The fourth section focuses on the design of a set of heuristics and the third section brings these tools together in a generic methodology. Finally, section five illustrates the efficiency of the proposed methodology on different data sets, from structured to unstructured environments.

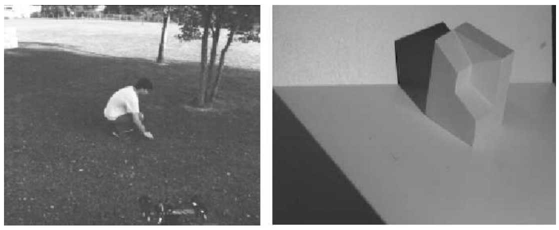


Figure 1: Different environments to be dealt with.

First of all, we propose to show the kind of environments we want to analyse, just to insist on the diversity of their nature. Figure 1[1] shows a natural environment which is captured by a stereoscopic system and a polyhedrical environment which is captured by a laser range system. However, for making simpler, we focus on a specific robotic outdoor navigation goal. Natural images were acquired by the LAMA robot belonging to LAAS-French CNRS lab.

[1] 3D point set images are best viewed in color, as many details are not as clear in grayscale.

2 Clustering algorithm as inference motor

Given a point set in $\Re^d$, the problem consists in self-paritionning it without learning. In this perspective, K-means algorithms can do automatic classification, based on the minimisation of an objective function U, a fuzzy K-partition, and V a set of K prototypes:

$$J = \sum_{j=1}^{N} \sum_{i=1}^{K} (u_{ij})^m d^2(X_j, V_i); K \leq N, \tag{1}$$

where X_j is the coordinate vector of the j^{th} point, V_i is the centroide of the i^{th} cluster, u_{ij} is the membership coefficient of X_j to V_i, $d(X_j, V_i)$ is a distance between X_j and V_i, N is the number of points and K the number of clusters, and $m(1 \leq m)$ controls the fuzzy degree of the final partition. The fuzzy partition is performed by an iterative optimisation of equation 1.

This model is flexible enough to be able to deal with a great diversity of data, just by replacing the distance function or the expression of prototypes. For instance, when clusters of different shape and density are considered, we can introduce an "exponential" distance d_e, inspired by the estimation of the maximum of likelihood. This distance is involved in the computation of the *a posteriori* probability u_{ij} of selecting the i^{th} cluster knowing the j^{th} vector of coordinate X_j :

$$u_{ij} = \frac{\frac{1}{d_e^2(X_j, V_i)}}{\sum_{i=1}^{K} \frac{1}{d_e^2(X_j, V_i)}},$$

$$d_e^2(X_j, V_i) = \frac{[det(F_i)]^{1/2}}{P_i} exp[(X_j - V_i)^T F_i (X_j - V_i)/2],$$

$$where\ P_i = \frac{1}{N} \sum_{j=1}^{N} u_{ij}$$

$$and\ F_i = \frac{\sum_{j=1}^{N} u_{ij}(X_j - V_i)(X_j - V_i)^T}{\sum_{j=1}^{N} u_{ij}}$$

F_i being the fuzzy covariance matrix of the i^{th} cluster and P_i the *a priori* probability of selecting the i^{th} cluster.

This specific algorithm performs well in the cases of shape diversity (it is able to differentiate between linear and spherical clusters), density and size diversity, which are characteristic properties of stereoscopic data [8].

Besides, using a validity measure of partition called Average Partition Density (APD), this algorithm gives an idea of the optimal number of clusters in the point set:

$$APD(K) = \frac{1}{K} \sum_{j=1}^{K} \frac{S_j}{V_j}$$

where the fuzzy hypervolume V_j and the "sum of central moment" S_j are given by:

$$V_j = det(F_j)^{1/2}\ and\ S_j = \sum_{x_i \in X_j} u_{ij}$$

for each $X_j = \{x \in X : (x - v_j)^T \sum_j^{-1} (x - v_j) < 1\}$.

Thus, we focus on this specific clustering method called UFP-ONC algorithm (unsupervised fuzzy partition - optimal number of classes) described in [5]. This generic algorithm is fundamental to our methodology in which it acts as inference motors do in knowledge-based systems.

Yet, unsupervised clustering is a somewhat arbitrary concept in the field of data analysis. What is certain is that we rarely know what we are looking for. This is particularly the case in data mining applications. In the computer vision field, things are somewhat easier. Most of the time, we know rather or less what kind of structures we are looking for, even without any available *a priori* model. In autonomous robotics for instance, the vision task consists in building an obstacle map. In this specific application, the nature of the structures we are looking for is binary: either navigable areas like the ground or obstacle areas emerging from the ground. In the range segmentation contest described in [3], the nature of clusters is represented by planar patches. Hence, some basic knowledge of the nature of objects is sufficient to validate the clustering process. We propose to represent this basic knowledge by a set of heuristics. Henceforth an inference is a meaningful subset of points and the semantics is going to be captured by a set of topological, metric and geometrical heuristics.

3 Methodology

3.1 Axiom

Our methodology is based on an axiom.

Axiom 1 *For any point set in which planar structures play an important role, there is at least one K-partition derived by the $UFP-ONC$ algorithm exhibiting a structuration into areas and objects which is adapted to any specific computer vision goal.*

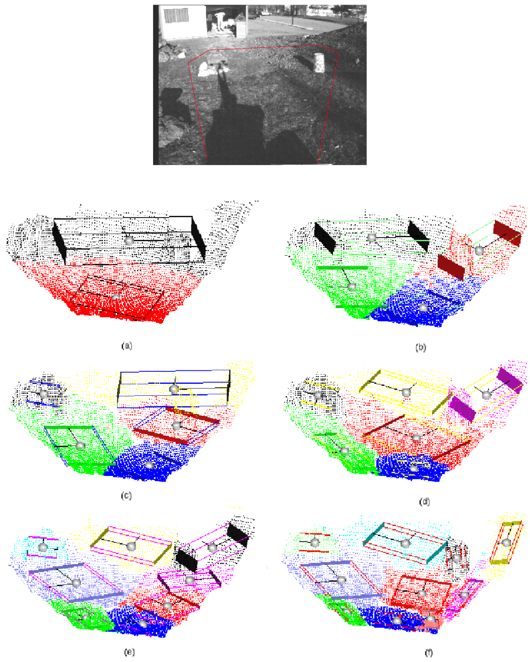
Let us illustrate this axiom in the context of autonomous robotic navigation goals. As a matter of fact, this observation has been verified on numerous such data scenes as outdoor natural environments. As a specific illustration, figure 2 shows that any K-partition for K between 2 and 10 provides a coherent interpretation of the scene. In association with each cluster, spheres represent the fuzzy center of gravity of the cluster and boxes represent the fuzzy eigen values and vectors of the correlation matrix of the point set corresponding to the cluster. The 2-partition exhibits a partitionning into foreground and background which differente each other by their point density. The 4-partition homogeneously divides the scene into four coherent areas, one of which, in the right background, gathers a white textured upset dust bin and the rising part of the ground. The 5-partition isolates for the first time the stoney area in the left background. Finally, the 10-partition exhibits every obstacle identified in the scene and correctly isolated from the navigable areas: the upset dust bin, the rising ground part and the stoney area. This series of K-partitions illustrates the "clever" behavior of the $UFP-ONC$ algorithm in this kind of data. As a matter of fact, for any K-partition, one obtains a coherent and useful interpretation of the scene in K clusters.

However, the $UFP-ONC$ algorithm doest not yet provide the optimal number of clusters to describe the scene. The whole strategy of interpretation is reliant on the determination of the optimal parameter K. With this in mind, we briefly overviewed the partition measure of validity proposed in the litterature, focusing on the Average Density Partition or $APD(K)$ defined in [5] which appeared to be the most robust for our experiments.

Unfortunately, the use of this sole measure is unable to guarantee an optimal interpretation of the scene for our robotic navigation purpose (the first local optimum of $APD(k)$ for k from one to ten is obtained for four clusters in the scene depicted in figure 2). It may provide an under segmentation result. More details are generally required. To find out this optimal number K of clusters, we need to use some more heuristic criteria based on the geometry and topology of the segmented structures.

3.2 Strategy

Depending on the application, some specific characteristics of the clusters are required to belong to an optimal description of the point set. For instance, in a famous contest launched by Herbert [3] , several research teams can compare their range data segmentation into planar regions results using a database made of polyhedrical objets. In this case, the required geometrical property for the clusters is planarity. In robotic stereoscopy, to avoid obstacles, we seek well-shaped and isolated clusters emerging from planar ground. Each application will need its own heuristics to perform the best partitionning according to some geometrical and topological requirements. Thus, the proposed methodology is based on the $UFP-ONC$ algorithm and the proposed axiom to build a generic range data partitionning process.



(a) (b) (c) (d) (e) (f)

Figure 2: Typical scene for outdoor robotic navigation. Red line delimitates the processed area and different K-partitions : (a) k=2 (b) k=4 (c) k=5 (d) k=6 (e) k=8 (f) k=10

First, we process point sets to exhibit an initial structuration by incrementally computing K-partition for K between 1 and 10, and we stop as soon as $APD(K)$ reaches a local maximum. From this first partitionning, in the case of clusters not verifying specific heuristics defined for the application, we continue the incremental partitionning process. This strategy is illustrated in the process chart of figure 3.

4 Heuristic set

The second fundamental element of the strategy consists in the creation of a set of heuristics to caracter-

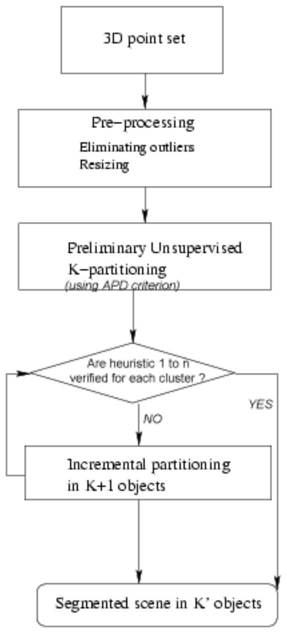


Figure 3: Process chart of the basic strategy

ize the validity of a cluster. Just as the declarative knowledge concept used in expert systems, this heuristic set is separate from the core of the system and can be easily modified and updated. This set may be a user accessible part to which some new application-oriented heuristic may be easily added.

We focus on outdoor scene analysis. In this context, obstacles and navigable areas must be identified. Navigable areas correspond to planar structures in the direction of the ground, and obstacles are structures emerging from these structures. Besides, obstacles are caracterized by a main direction which has to be not parallel to the ground direction (physical property). In fact, ground direction can be easily obtained as the closest large planar area to the captor obtained after the preliminary K-partitioning.

When clustering (as defined in the last section) is performed on 3D outdoor scene data, two biased behaviors of algorithm appear. As there is a "preference" for topological shape information in the distance function, it may happen that:

- obstacles emerging from the ground are gathered because they emerge together from a large planar structure;

- small obstacles like little rocks are "swallowed up" by a large planar structure.

4.1 Connexity Ambiguity heuristic

One of the biased behaviours of the UFP-ONC algorithm consists in gathering some obstacles (the black ones) which emerge from a wide ground cluster (the gray one) just because they are similar in term of shape (see Figure 4). We solve this problem by designing a CAh heuristic.

Figure 4: 2D schema of the first biased behaviour of the UFP-ONC algorithm: unfortunately, there are only two clusters (in black and gray) as opposed to three clusters for "ground truth".

First, we project the 3D point set corresponding to a cluster on a specific plane. In general, the projection plane corresponds to the best approximating plane of the point set (see subsection 4.4). Then, we compute a filtered mesh representing the projected point set (see Figures 6 and 7). This operation is based on Delaunay triangulation and α-shape theory [2]. In a precedent article [9], we exposed our design of morphological operators like α-erosion, α-dilatation, α-opening performing on such structure. These operators can filter shape represented by 2D unorganised point set (see figure 5).

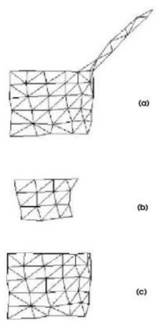


Figure 5: (b) Erosion and (c) Opening of the (a) original mesh

Then, if the number of significative connex components is greater than one, we decide that the cluster is ambiguous. Let us call this heuristic the CAh heuristic for Connexity Ambiguity heuristic. This heuristic captures the topological incoherence of a cluster.

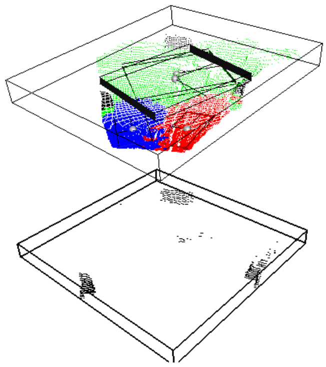


Figure 6: Above: 4-partition in which three little rocks with hyperellipsoidal shapes emerge from a ground plane and are gathered in a single cluster. Below: the ambiguous object has been isolated

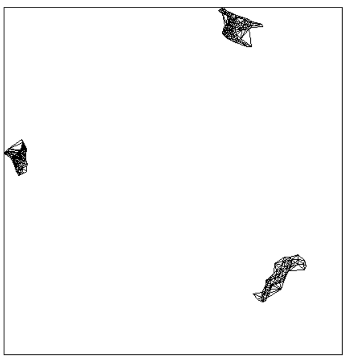


Figure 7: Connexity Ambiguity for black-colored object

4.2 Size Ambiguity heuristic

Let us define now a more specific heuristic for outdoor vision goal, which will be very useful for the robotic navigation purpose. The other biased behavior of the UFP-ONC algorithm consists in gathering a wide, dense planar ground cluster with non-dense, little obstacles (see Figure 8). This makes a cluster which is not planar but whose direction is very similar to ground direction. In this case, we do not classify the cluster as an obstacle because its main orientation is almost parallel to the ground orientation. So it remains an ambiguous cluster. Let us call this heuristic SAh for Size Ambiguity heuristic.

Figure 8: 2D schema of the second biased behaviour of the UFP-ONC algorithm: unlikely, there is only one cluster as opposed to two clusters for "ground truth"

4.3 Planar Ambiguity heuristic

The last important heuristic to add consists in a simple planarity test. This very useful heuristic is sufficient for the planar range segmentation contest of [3] as will be illustrated in the results section. Let us call it the PAh, Planar Ambiguity heuristic.

Note that these three heuristics are quite general and that they can be applied to most applications.

4.4 Parameter setting

All these parameters are computed from the fuzzy correlation matrix F_i of the point set corresponding to the cluster i, that is the eigen values $\mu_1 \leq \mu_2 \leq \mu_3$ and vectors of this matrix, plus the fuzzy center of gravity C of the cluster.

heuristic	parameters
Best approximating plane (BAP)	Plane(C,$\vec{\mu_2}, \vec{\mu_3}$)
Ground direction	The closest to the captor best approximating plane
PAh	μ_1
SAh	Angle($BAP, \overrightarrow{ground}$)

5 Results

We test the proposed methodology on different sorts of data:

- natural range image database from LAAS-CNRS laboratory at Toulouse, France, acquired by a stereoscopic system on the LAMA robot (100 scenes);

- contest range image database from different universities, and describing arrangments of polyhedrical objets (30 scenes).

5.1 Natural range image database

We illustrate in figure 9 the way our methodology processes such 3D point sets with :

heuristic	parameters
PAh	$\mu_1 \leq 5cm$
SAh	Angle($\overrightarrow{\mu_2\mu_3}, \overrightarrow{ground}) \leq \pi/8$

The proposed methodology gives coherent results for the entire scene set, allowing the robot to avoid the obstacles during autonomous navigation, and also the 3D reconstruction of the scenes in terms of obstacle shape and navigable area.

We are also working on large scale environment exploration. Then, the point is to merge the different partitioned views of a sequence of 3D point clouds to globally reconstruct the environment as an obstacle map. We worked on a robot which moved straight forward and obtained an initial 3D mesh reconstruction of the different obstacles encountered on its path, as illustrated in Figure 10.

5.2 Contest range image database

Figure 11 illustrates the way our methodology processes such 3D point sets. The PAh heuristic only was used with a $1cm$ threshold for planar tests.

The evaluation of the proposed methodology is also performed following the methodology of Hoover et al. [3]. The evaluation methods involves 30 structured light scanner images (ABW set) which were segmented using CAD models of the objects in order to obtain *ground truth* (GT) images. A tool tries to objectively compare the *machine segmentation* (MS) to the corresponding GT image using a set of defined performance metrics for instances of correctly segmented, missed and noise regions, over

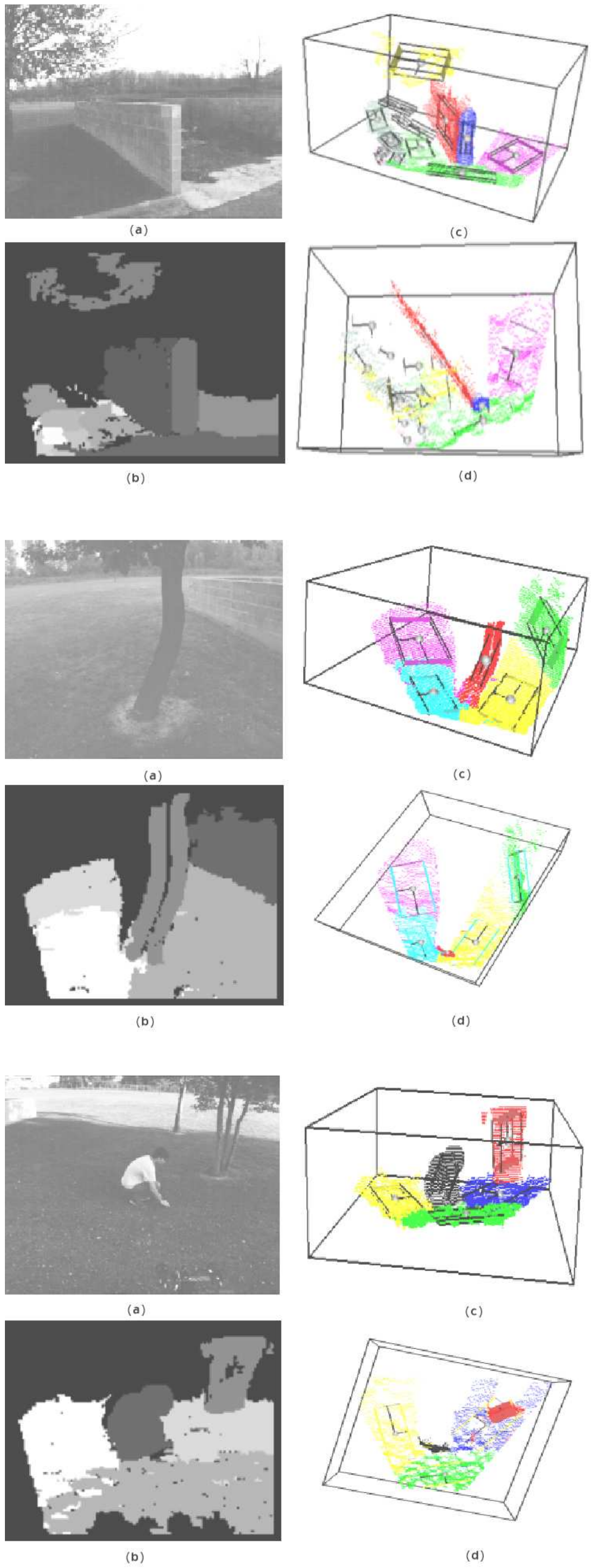


Figure 9: Segmenting natural range image.

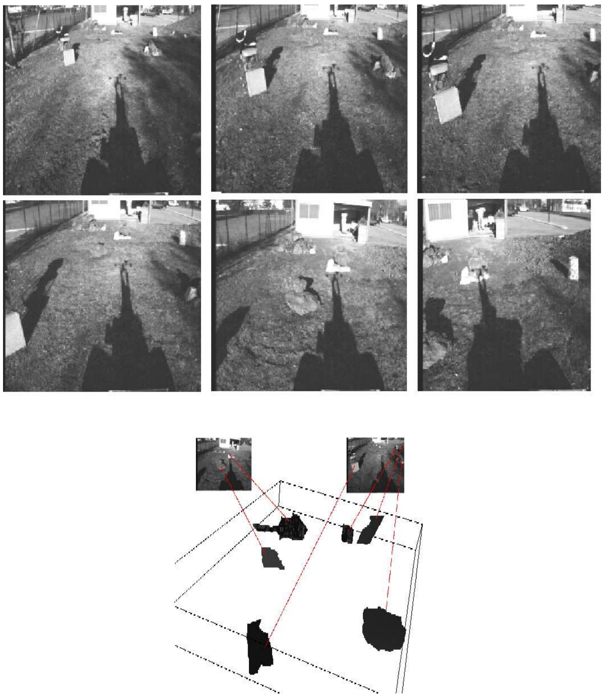


Figure 10: Sequence of scenes captured by LAMA robot moving straight forward and obstacle map wih 3D-meshed reconstructed structures.

and undersegmentation, and accuracy of the recovered geometry. For the sake of comparaison, we implement some post-processing: for instance, our algorithm automatically merges ground planes, which is a better performance in terms of interpretation, but it would yield undersegmentation results in the Hoover frame.

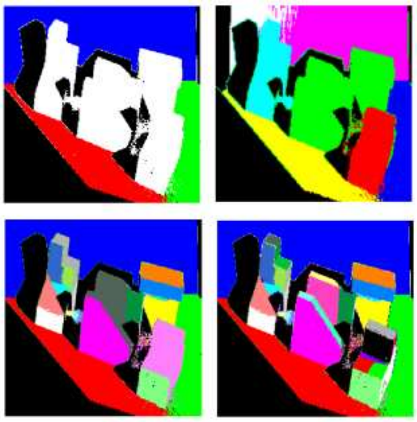


Figure 11: Different steps of incremental partitioning of a polyhedrical range image.

Numerous research teams have attempted to compare their segmentation algorithm using this database. Following [7], we propose some quantitative results as illustrated in Table 1. Numbers are averages over the set of images : on an average image, there are 15.2 GT regions.

research group	GT regions	correct detection	angle diff. (std. ev.)
USF	15.2	12.7	1.6 (0.8)
WSU	15.2	9.7	1.6 (0.7)
UB	15.2	12.8	1.3 (0.8)
UE	15.2	13.4	1.6 (0.9)
UBham	15.2	13.4	1.6 (0.9)
UP5	15.2	12.2	1.7 (0.9)

research group	over-segment.	under-segment.	missed	noise
USF	0.2	0.1	2.1	1.2
WSU	0.5	0.2	4.5	2.2
UB	0.5	0.1	1.7	2.1
UE	0.4	0.2	1.1	0.8
UBham	0.4	0.3	0.8	1.1
UP5	0.3	0.1	2.6	2.2

Table 1: Quantitative comparison with other segmentation results

These tables show the good quantitative performances of our algorithm (UP5 results) even though they do not surpass the correct detection rates of UB and Ubham. However, the other algorithms were specifically designed for this contest: they provide segmentation results in the case of range images which are structured depth information. As a matter of fact, segmentation is performed on depth or disparity maps which are classical 2D intensity images with neighbourhood relationships. Our algorithm performs as well on such organised data as on unorganised geometric point sets. It performs partitionning much more than segmentation. In this sense, it outperforms the other algorithms by the weak hypothesis on data, and so by its genericity. Besides, the Hoover tool is not as objective as it is said to be. Actually, the relative loss of performance comes mostly from the missed or noise regions which are subtracted from the score of correct detection. This means that whatever the size of the missed region, it counts plus one in the score of missed detection and minus one for the score of correct detection. In fact, we observe that most of our missed regions are little regions - that our fuzzy modelling considers as outliers and - whose detection can be controlled by internal parameters but whose importance is actually not so obvious.

6 Conclusion

Basically, we sought to answer a difficult methodological problem in computer vision: how to develop generic tools for point set partitionning in the field of computer vision. The word generic here means able to easily adapt itself not only to new data with new characteristics but also to new goals with new specifications.

This generic aspect is performed in the context of knowledge-based systems as illustrated in figure 12. The inference motor acts independently of the knowledge module. A specific algorithm called $UPF - ONC$ was used as "cluster maker". In parallel, a set of heuristics is fed to incorporate either basic knowledge on expected cluster characteristics depending on the specific application, or more complex knowledge models if available.

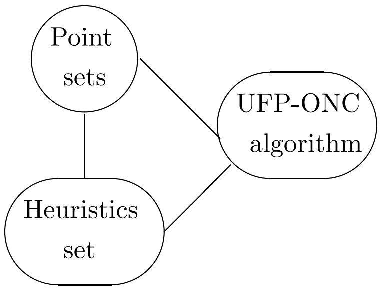


Figure 12: Knowledge-based system paradigm for point set analysis

First experiments show qualitatively that this flexibility makes it possible to deal with any kind of unorganised point sets for any application. They also yield good quantitative results after comparison with results coming from an international range segmenter contest.

Further experiments and real implementation on a robot for obstacle avoidance are planned. A recognition module of obstacles based on the point set analysis is also under consideration.

Basically, we think that it is a promising step towards modeling large scale environments.

Acknowledgements : This work is supported by the firm EADS (French Air-space Society). The author would like to thank Nicole Cambou and Simon Lacroix for their support.

References

[1] R. Chaine and S. Bouakaz. Analyse surfacique de donnes 3-d non structures : une approche basee sur les graphes. In *RFIA*, pages 37–46, 2000.

[2] H. Edelsbrunner and E.P. Mcke. Three-dimensionnal alpha-shapes. *ACM Transactions on Graphics*, 13(1):43–72, 1994.

[3] A. Hoover et al. An experimental comparison of range image segmentation algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(7):673–689, july 1996.

[4] Z. Tu F. Han and S-C. Zhu. A stochastic algorithm for 3d scene segmentation and reconstruction. In *European Conf. on Computer Vision*, pages 502–516, Copenhagen, Denmark, May 2002.

[5] I. Gath and A.B. Geva. Unsupervised optimal fuzzy clustering. *PAMI*, 11(7):773–781, july 1989.

[6] X.Y. Jiang and H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. *Machine Vision and Application*, 7(2):115–122, 1994.

[7] Klaus Kster and Michael Spann. Mir: An approach to robust clustering - application to range image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(5):430–443, may 2000.

[8] N. Loménie, L. Gallo, N. Cambou, and G. Stamon. Structuration plane d'un nuage de points 3d non structuré et détection des zones d'obstacle. In *Vision Interface*, Trois-Rivieres, Canada, may 1999.

[9] N. Loménie, Laurent Gallo, N. Cambou, and G. Stamon. Morphological operators on representations based on delaunay triangulation. In *Proc. International Conference on Pattern Recognition*, pages 556–559, Barcelona, Spain, september 2000.

[10] I. Stamos and M. Leordeanu. Automated feature-based range registration of of urban scenes of large scale. In *Int. Conf. on Computer Vision and Pattern Recognition*, pages 555–561, Madison, WI, June 2003.