3D Perceptual Organization of Laser Altimetry Data

Impyeong Lee, Toni Schenk
Dept. of Civil & Environmental Engineering and Geodetic Science
The Ohio State University
USA
lee.1517@osu.edu, schenk.2@osu.edu

Commision III, Working Group 3

KEY WORDS: Perceptual Organization, Segmentation, Merging, Grouping, LIDAR, Laser Altimetry, Points, 3D

ABSTRACT

Perceptual organization is to group sensory primitives arising from a common underlying cause by imposing structural organization on sensory data. It has been emphasized as a robust intermediate-level grouping process toward object recognition and reconstruction since it imparts robustness and computational efficiency to the perceptual process. Sarkar and Boyer (1993) proposed a classificatory structure for perceptual organization and clarified what should be done under each class. Despite intensive research on 2D data, 3D perceptual organization is still in its infancy, however. Increasing research efforts are needed to understand 3D data from various range sensors such as laser altimetry systems. Therefore, the purpose of this research is to develop a robust approach for constructing 3D perceptual organization from irregularly distributed 3D points acquired from laser altimetry systems. The scope of perceptual organization presented in this paper is limited to signal, primitive and structural levels. At the signal level, we organize raw 3D points into spatially coherent surface patches. Then, at the primitive level, we merge the patches into co-parametric surfaces and detect breaklines and occlusions. Finally, at the structural level, we derive useful surface combinations such as polyhedral structures. The approach has been successfully applied to real laser altimetry data. The organized output is on a much more abstract level than the raw data and makes information explicit. Thus, it serves as a valuable input to higher order perceptual processes, including the generation and validation of hypotheses in object recognition tasks.

1 INTRODUCTION

Perception is the process by which organisms interpret and organize sensory stimulus to produce a meaningful description of the world. Especially, the ability to impose organization on sensory data in human perception started to be emphasized by Gestalt psychologists from the early 20th century (Koffka, 1935; Khler, 1929). It has been recognized as a crucial component that makes human perception powerful and volatile. Hence, many systems in computer vision organize primitive sensor data into perceptually meaningful groups before advancing to higher-level processing such as object recognition. These grouping processes are known as perceptual organization, formally defined as a process that groups sensory primitives arising from a common underlying cause by imposing structural organization on sensory data (Sarkar and Boyer, 1993).

Since the roles of perceptual organization were emphasized as a robust intermediate-level vision process by Witkin and Tenenbaum (1983) and Lowe (1985), many researchers have demonstrated the importance of perceptual organization in the tasks under many levels and domains of computer vision, e.g., figure-ground discrimination (Herault and Horaud, 1993), motion-based grouping (Allmen and Dyer, 1993; Chang and Aggarwal, 1997), object recognition (Havaldar et al., 1996; Modayur and Shapiro, 1996; Nelson and Selinger, 1998; Zisserman et al., 1995), building detection (Henricsson, 1998; Lin et al., 1994) and change detection (Sarkar and Boyer, 1998).

Sarkar and Boyer (1993) propose a classificatory structure of perceptual organization based on the dimension over which organization is sought and the abstraction level of features to be grouped. The structure has two axes: one axis denotes 2D, 3D, 2D plus time and 3D plus time; and the other axis represents signal, primitive, structural and assembly level. For example, surface segmentation from laser altimetry data can be classified into 3D signal level perceptual organization. In addition, further grouping of the segmented surfaces falls under 3D primitive or structural level perceptual organization.

Even though they suggest what should be done under every class, the previous work has mainly concentrated on 2D organization, dealing with all the abstraction levels and emphasizing the structural level. In addition, some researchers also consider 2D plus time (Sarkar, 1995) but profound research is still required for successfully applying the organization to interpreting motion sequences. In 3D organization, most previous studies falls under signal level only, particularly focusing on range image segmentation. However, the need of perceptual organization in the various levels of 3D will significantly increase because 3D sensors become cheaper and more available. Hence, "perceptual organization in 3D" are emphasized as one of the most important research directions (Boyer and Sarkar, 1999).

The most important 3D sensors recently widely used in many photogrammetric applications are laser altimetry systems. A chain of photogrammetric processes traditionally starts from images. Inference of 3D information from images involves a matching process to find conjugated features from images (Schenk, 1999). However, matching is a highly intelligent process that cannot be easily archived by computers in spite of the astonishing development in artificial intelligence and computer vision during the last several decades. Therefore, the automation of the entire processes is still extremely challenging.

Laser altimetry data have been thus recently noticed as alternative or supportive to images, since laser altimetry systems produce 3D points by sampling directly physical surfaces. They provide a cloud of irregularly distributed raw surface points consisting of the x, y, z coordinates for each laser footprint without radiometric information. In addition, multiple echo data, radiometric values at the laser wave length, the waveforms of returned laser pulses are also available depending on the systems but not considered in this research.

Applications using the laser altimetry data are rapidly increasing, ranging from DEM (Digital Elevation Model) construction to urban modelling. Many post-processing algorithms are reviewed by Tao and Hu (2001). Most of the algorithms typically involve interpolating into regular grid data, separating the ground surface, detecting upper-ground objects such as buildings (Gamba and Houshmand, 2000; Maas and Vosselman, 1999), trees (Hyyppa et al., 2001) and other objects (Axelsson, 1999) and further identifying their changes (Murakami et al., 1999).

Instead of introducing another application oriented algorithm, we intend to establish a middle-level process which are less dependent on an application and hence sufficiently general for various applications. Using this process, we will derive a robust, explicit and computationally efficient description from raw data, which usually include many redundancies and outliers. As the most suitable approach to perform this, we propose 3D perceptual organization, which has been proved as a robust intermediate process for various tasks in computer vision.

We significantly benefit from the use of 3D perceptual organization as an intermediate step toward various applications of laser altimetry data. Main advantages are summarized as follows:

- Explicitness: perceptual organization provides more abstract and explicit description of raw data. For example, we do not need 50 points sampled from a planar roof of a building and prefer to have explicitly the boundary and the parameters of the plane.
- More information available: according to the principle of Gestalt laws, one plus one is not just two but much more than two. We can compute from a segmented surface various additional information, which is not meaningful to a point, such as point density, surface roughness, outlier ratio, area, orientation, surface normal and so on.
- Robustness: a grouping process contributes to identifying outliers since it tries to group mutually consistent entities and non-grouped entities usually corresponds to outliers. Hence, grouped entities are a robust description of the original data.
- Reduced complexity: perceptual organization significantly reduces the number of entities thanks to its explicitness and abstractness. For example, when we detect buildings from laser altimetry data, if we search from perpendicular surface combinations rather than from all surfaces (or extremely all points), the number of entities that we should check with a building hypothesis is much smaller.

In summary, the objective of this research is to present a framework that computes 3D perceptual organization from laser altimetry data. Here, the computed organization should be a robust, explicit and computationally efficient description of the original data so that they can be flexibly used as an input for various higher-level processes. The problem is more formally stated in the next subsection.

The problem statement is followed by three sections which describe the proposed approach, show the experimental results and conclude with discussion and future research, respectively.

1.1 Problem Statement

Given a set of irregularly distributed 3D surface points acquired from laser altimetry systems, compute perceptual organization at signal, primitive and structural levels. At the signal level, we organize the raw points into spatially coherent surface patches with their boundaries. Then, at the primitive level, we merge the patches into co-parametric surfaces, refine the boundaries and identify breaklines and occlusions. Finally, at the structural level, we derive useful surface combination such as parallel surfaces and continuous surfaces, identify the ground surface and generate hypothesized surfaces for the occluded areas. Table 1 summarizes the inputs and outputs at each organization level.

Table 1: Inputs and outputs at each organization level

Level	Inputs	Outputs
Signal	Points	Patches with their boundaries
Primitive	Patches	Surfaces with refined boundaries, breaklines and occlusions
Structural	Surfaces	Surface combinations with hypothesized surface

2 THE PROPOSED APPROACH

As acknowledged from the objective of this research presented in section 1, we have been focusing on developing an overall framework rather than inventing a new specific algorithm to constitute the framework. It is because the use of perceptual organization for the post-processing of laser altimetry data has been rare with our best knowledge and also the 3D perceptual organization is still in infancy although tremendous research has been performed for 2D data (Boyer and Sarkar, 1999). Thus, our fundamental strategy for this work is to review various research efforts in 2D data and extend them for 3D data, particulary in the domain of laser altimetry data, even though we developed inevitably new pieces in several cases.

Despite the lack of previous studies, we managed to introduce two representative studies. At the signal level, Ahuja and Tuceryan (1989) extracted perceptual organization from irregularly distributed 2D points called dot patterns. They classified dots into interior dots, border dots, curve dots, and isolated dots using their relationships with the neighborhood defined by Vornoi diagram. They also used a probabilistic relaxation process to produce a globally optimal result. At the primitive and structural level, Fisher (1989) grouped surfaces reconstructed from 2D images into 'surface clusters', that is, perceptually meaningful surface combinations. Since the surfaces are reconstructed from 2D images, they have many different aspects from those from 3D laser points. However, his research is a valuable basis for this work.

Based on the relevant studies including not only these representative studies but also many valuable research regarding 2D perceptual organization, range image segmentation and unsupervised point clustering, we establish a framework comprised of three grouping processes at the signal, primitive, and structural levels, where the inputs and outputs at each level are previously summarized in Table 1.

The process at each level includes preprocessing, main processing and postprocessing. Preprocessing performs the task supportive to grouping such as defining the adjacency among features and computing the attributes of features. Postprocessing complements grouping outcomes, for example, by filling gaps, determining boundaries, identifying breaklines and occlusions, adding hypothesized surfaces, and inferring more complex entities such as the ground surface.

The main processing at each level is designated as segmentation, merging, and grouping, respectively. Although we use three different terms so that they can be more appropriate for the features to be grouped at each level, all of them are actually classified to grouping processes.

The grouping process consists of three components which should be deliberately selected mainly based on the features to be grouped and the groups to be sought. The components are "grouping cues", "testable feature subsets", and "cue integration method" (called grouping mechanism here), as described by Berengolts and Lindenbaum (2001). Grouping cues are the information that indicates whether two or more entities arise from an object, such as proximity, connectedness, continuity, similarity, parallelism, symmetry, common region and closure (Sarkar and Boyer, 1994b). Testable feature subsets indicate a subset of features, inside of which we examine the validity of the grouping cues. The size of the subset can be determined by considering the meaningful range of the cues and the computational complexity. Grouping mechanism is the means by which we produce globally optimized grouping of the entire set by integrating the grouping cues locally computed inside the testable feature subsets. The method is frequently implemented as an optimization process that minimizes a cost function.

One should also determine how to represent the perceptual information being processed during the grouping process. Representation using a graph structure where nodes indicate the entities to be grouped and arcs describe perceptual information between the entities is a promising choice as indicated by many other researchers (Zahn, 1971; Geman et al., 1990; Herault and Horaud, 1993; Matula, 1997; Shaashua and Ullman, 1988; Shapiro and Haralick, 1979; Wu and Leahy, 1993).

2.1 Signal Level

At the signal level, we group raw 3D surface points into surface patches (or point clusters). The process is summarized in Figure 1.

Preprocessing

- Establish adjacency among points.
- Compute the attributes of points in robust ways.
- Refine the adjacency based on the computed attributes.
- Identify isolated points based on connected component analysis.

Segmentation

- Select grouping cues, testable feature subsets, grouping mechanism.
- Perform grouping to generate patches.

Postprocessing

- Fill patches using the isolated points.
- Determine the boundaries of the patches.

Figure 1: Signal level process

Adjacency: neighborhood of a point Adjacency defining the neighborhood of a point is required for a grouping process, which often access the neighborhood to compute multi-feature grouping cues and check them with grouping criteria. Neighborhood in a set of irregularly distributed points is not obvious, however. Various neighborhood concepts are well reviewed early by Ahuja and Tuceryan (1989) and recently by Chaudhuri (1996). We describe several examples here. The best one among them is determined according to the size and distribution of a data set.

The Delaunay triangulation is a acceptable choice since the region of influence of a point is determined by the Voronoi diagram and the adjacency of the regions is then expressed by the edges of the Delaunay triangulation. Some researchers use 2D Delaunay triangulation considering only horizontal coordinates (Maas and Vosselman, 1999) while others also use 3D Delaunay triangulation (Lee and Schenk, 2001). The computational requirement constructing the Delaunay triangulation is the hindrance to its use in a large data set.

Simple approaches such as selecting the k nearest points and the points inside a sphere of a fixed radius are thus useful in such sets. In addition to proximity considered by these approaches, the distribution of neighboring points is important. With the idea that neighborhood should be not only as near but also as symmetric as possible, Chaudhuri (1996) propose 'nearest centroid (or median) neighborhood'.

Attributes of points Attributes of points are the fundamental information which we can use to examine grouping hypotheses. The most primitive attribute of a point is its 3D coordinates. We frequently require more complex information such as point density, roughness, surface normal, plane parameters, outlier ratio and other properties which can be defined with a set of points rather than a single point. Such complex attributes are meaningful since the laser footprint of a point is actually not a zero-dimensional point but a two-dimensional elliptical area.

Although they can be considered as the attributes of a point, they cannot be defined or computed from only a point. Therefore, we compute them from a small patch defined around a point. The patch should be large enough to include the points from which the attribute can be computed. For example, at least three points are required for determining surface normal. Contrarily, the patch should be small enough to represent a point and its local area. In addition, the patch should consist of the neighboring points which locate not only near to its representing point but also symmetrically around the point. The nearest centroid (or median) neighborhood can be thus a strong candidate.

To compute an attribute of a patch, we establish a system of the equations, which formalizes the contribution of the interior points to the attribute. For example, if we compute plane parameters from a patch, each equation shows a plane equation substituted by the three coordinates of a point with the noises associated with each coordinate.

The computation of the attributes from this established system (usually overdetermined) should be equipped with a robust estimation approach rather than the least mean squares approach, because every small patch can include outliers, which result in an significantly different attribute comparing to the others. For example, Least Median Square Error (LMedS) estimation is a promising alternative since it allows up to 50 % outlier ratio in theory (Koster and Spann, 2000).

In addition, we measure the tendency of a point to be an outlier from every computation of the attribute. After the computation of all the attributes, we conservatively classify some points into outliers by synthesizing the outlier tendency measured from each computation.

Refined adjacency The points classified as outliers with significant evidence should not maintain its adjacency to at least the representing point of the patch. Accordingly, we refine the adjacency established before so that the outliers cannot be linked to the inliers.

Isolated points Based on the refined adjacency, we classify isolated points by the connected component analysis. This analysis produces groups of connected points based on the adjacency. Some of the groups may include very small number of points, which can be labelled isolated points.

Grouping cues among points Another aspect of grouping is the selection of grouping cues. At the signal level, proximity, similarity and continuity are typically considered. It is natural to group points which locate near to each other and show similar attributes. Grouping cues can be defined on at least two points. The more entities considered, the stronger cues can be realized. For example, if we intend to group the points expected to be on the same plane, we can check the similarity of the fitted plane parameters of the points. If we have more points involved, the similarity is stronger evidence for grouping. However, such multi-feature cues cannot be explicitly represented in a graph structure, since it can only include as arcs the bi-feature cues. To overcome this, Amir and Lindenbaum (1998) propose a procedure that enhances the strength of the bi-feature cues based on the multi-feature cues founded. They increase the strength of a bi-feature cue of two entities if multi-feature cues around the entities support the bi-feature cues and decrease them otherwise.

Testable feature subsets If we compute the grouping cues from all the points in a set, we would be confronted with the combinatorial explosion. Furthermore, some cues such as proximity are meaningless for two distant points. Hence, we have to specify a certain range named testable feature set, only the entities inside which we consider to compute these grouping cues. The range is deliberately determined by considering the validity of the cues and the computational complexity. At the signal level organization, a point and its connected points in terms of the refined adjacency are considered as the testable feature set.

Grouping mechanism Another component to be determined for a grouping process is grouping mechanism. The mechanisms range from optimized processes such as simulated annealing and probabilistic relaxation, usually involving heavy computation, to the connected component clustering of only linear time complexity.

According to the performance of the connected component clustering assessed by Berengolts and Lindenbaum (2001), it can be suitable for many practical applications requiring less computation complexity and medium quality of grouping. Iterative growing is also a strong candidate, which is similar to connected component clustering since it also follows the connection among the entities. Its uniqueness comes from performing iteratively testing a new point, including (or discarding) the point, and updating the attributes of a growing group (Lee and Schenk, 2001). In addition, some researchers use scalar or vector voting (Sarkar and Boyer, 1994a; Guy and Medioni, 1997; Tang and Medioni, 1998; Lee and Medioni, 1999) and graph spectral partition (Sarkar and Soundararajan, 2000).

Although the selection of a mechanism depends on the allowable computation complexity and the application to be sought, we are not willing to choose a complex time-consuming process in general. It is because grouping as a middle-level process should reduce the overall complexity of a whole process. Consequently, iterative growing is a reasonable choice.

Segmented patches: point clusters Based on the grouping cues, the testable feature sets, and the grouping mechanism considerately selected, the segmentation process is applied to a set of points. The segmented outputs are spatially coherent surface patches represented by the interior points and the surface parameters shared by them.

Filling The segmented patches may have small holes. Some of them are filled if they are matched to their close isolated points. Similar approach is used by Boyer et al. (1994).

Boundaries Additional useful description of the segmented patches are their boundaries which can be computed as a post-processing from the distribution of the interior points. It is a intricate problem that many researchers have attacked because the interior points are irregularly distributed on a 2D surface locating in 3D space.

An option is the convex hull (Berg, 2000), which represents the outlines of the minimum convex area covering all the interior points. It is good for representing an overall coverage of a patch but weak for describing the actual boundaries which may include concave shapes and holes. To accommodate concave shape and holes in a certain degree, Richards and Mullins (1977) presented the space filling hull defined as the union of the discs, each of which is associated with a point. The radius of the disc is a parameter to be determined and can be selected as the half distance between a point and its nearest point (Toussaint, 1988). Edelsbrunner et al. (1983) proposed α-shape as a general description of the boundaries. It is a family of graphs, each of which is a subset of Delaunay triangulation (Berg, 2000). The parameter α varies ∞ to 0 and controls the level of details. It conceptually corresponds to the radius of the space filling hull. It is extended to weighted α-shape for accommodating the variable density of points (Edelsbrunner, 1992). By synthesizing the ideas of the space filling hulls and α shapes, Melkemi and Djebali (2000) and Melkemi and Djebali (2001) propose more sophisticated description such as "A-shapes" and "weighted A-shapes", respectively. While most of these description are subsets of the Delaunay triangulation or a regular triangulation, Chaudhuri et al. (1997) propose "r-shape" based on a regular grid defined over the points with the intervals prudently selected by considering the point density.

Preprocessing

- Establish adjacency among patches.
- Compute the attributes of patches.

Merging

- Select merging cues, testable feature subsets, merging mechanism.
- Perform merging to generate co-parametric surfaces.

Postprocessing

- Refine the boundaries of the surfaces.
- Identify breaklines and occlusions.

Figure 2: Primitive level process

Preprocessing

- Establish adjacency among surfaces.
- Compute the attributes of surfaces.

Grouping

- Select grouping cues, testable feature subsets, grouping mechanism.
- Perform grouping to generate useful surface combinations (parallel and connected surfaces)

Postprocessing

- Identify the ground surface.
- Add hypothesized surface on occluded areas.
- Extract polyhedral structures.

Figure 3: Structural level process

2.2 Primitive Level

At the primitive level, we merge the patches segmented at the signal level into co-parametric surfaces, refine the boundaries of the merged surfaces and identify the breaklines and occluded areas, as shown in Figure 2

Adjacency: neighborhood of a patch The adjacency established among points at the signal level is extended for defining the adjacency between edges and between patches. If at least a point of an edge is adjacent to a point of another edge, we define that the two edges are adjacent each other. The boundaries of a patch computed at the signal level are a set of the edges, each of which links two boundary points. If at least a boundary edge of a patch is adjacent to a boundary edge of another patch, we define that the two patches are adjacent each other. Furthermore, we define the ratio between the total length of adjacent boundary edges and the total length of all boundary edges as a measure indicating the degree of adjacency between patches.

Attributes of patches The attributes of patches are various, ranging from those already described at the signal level such as point density and roughness to those associated with the shape of the boundaries such as the orientation and the aspect ratio. The most useful attributes for the merging process are the surface parameters indicating the shape of the patch (for example, plane parameters) and their associated fitting errors.

Merging cues among patches The merging cues we selected are proximity and similarity so that we can merge two patches which are sufficiently near and shows the similar attributes, that is, the surface parameters and their fitting errors.

Testable feature sets Testable feature sets are simply defined based on the adjacency among patches. We intend to check every pair of adjacent patches with a merging hypothesis.

Merging mechanism We use iterative growing approach similar to the approach that Koster and Spann (2000) utilize for range image segmentation. It iteratively proceeds until no adjacent patches meet merging criteria. At each iteration, it investigates every pair of adjacent patches and computes a measure indicating how well each pair satisfies the merging criteria. For example, if we use a statistical test such as F-test to check a merging hypothesis, the p-value is the measure. Then, merging starts from the pair with the largest measure in a descending order of the measure. This contributes to producing a globally optimal merged set.

Merged surfaces The iterative growing approach equipped with the merging criteria based on adjacency and similarity produces a set of merged co-parametric surfaces. Their preliminary boundaries are computed using the same method as at the signal level. Based on the boundaries, we refine the adjacency among them.

Refined boundaries Refined boundaries are interesting outcomes supporting the Gestalt argument that one plus one is much more than two. Since no laser pulse can be reflected from only 0 or 1-dimensional entities in general (extraordinarily, from electrical power transmission lines, see (Axelsson, 1999)), we never directly extract from a set of laser points such information as the boundary of a roof. However, we can accurately infer the boundaries from two adjacent surfaces produced under the Gestalt principles. The boundaries of all adjacent surfaces are refined using their intersections.

Breaklines and occlusions The boundary edges of a surface which are not adjacent to other surface are identified as breaklines. Every breakline invokes a hypothesis for occlusion. The empty space between the breaklines and the nearest boundary of other surfaces are thus hypothesized as occlusions.

2.3 Structural Level

At the structural level, we group the merged co-parametric surfaces organized at the primitive level into useful surface combinations. Furthermore, we identify the ground surface, add hypothesized surfaces over occlusions and extract polyhedral structures. The process is summarized in Figure 3.

Adjacency: Neighborhood of a surface The adjacency among surfaces is also defined using the adjacency among their boundaries, which is already computed during the post-processing stage at the primitive level. In addition to the 3D adjacency, we establish in a similar way the 2D adjacency, that is, the adjacency defined by considering only the horizontal locations.

Attributes of surfaces All the same attributes as defined at the primitive level are meaningful at this level. Roughness, surface normal and many parameters describing the shape of the boundary are particularly useful.

Grouping cues among surfaces Several cues can be selected from various choices such as proximity, connectedness, continuity, similarity, parallelism, symmetry, common region and closure depending on the properties of the group to be sought. For example, proximity, connectedness and continuity are particularly useful to group the surfaces into the ground surface.

Testable feature subsets While we use just the adjacent features at the signal and primitive level, we should define the testable subsets by deliberately considering the selected attributes and cues. For example, the range where we check parallelism should be proportional to the area of the surface rather than just constant.

Grouping mechanism The mechanism can be also selected among the various ones described at the signal level. The connected component analysis is very useful in many applications such as identifying the ground surface and polyhedral structures.

Connected surfaces They are easily found by the connected component analysis on the 3D adjacency graph.

Parallel surfaces We define for every surface a testable feature subset, the size of which is proportional to the area of each surface. We examine the parallelism based on the similarity of the surface normal with every entity inside this subset. Based on the result, we construct a parallelism graph where an arc indicates the parallelism of the two surface linked through the arc. Parallel surfaces are thus identified by the connected component analysis on the parallelism graph.

The ground surface We examine continuity with the same testable feature subset as used for parallel surfaces and construct a continuity graph. The largest connected components of the graph is detected as the ground surface.

Hypothesized surfaces For each occlusion detected at the primitive level, we add a hypothesized surface and update the adjacency. For example, a roof of a building is adjacent to the ground surface not in 3D but in 2D. This inconsistency triggers a hypothesis of a vertical surface between them.

Polyhedral structures We perform the connected component analysis on the graph constructed by subtracting the detected ground surface and the very rough surfaces from the adjacency graph. Each connected surfaces corresponds a polyhedral structure.

3 THE EXPERIMENTAL RESULTS

The proposed approach was applied to constructing perceptual organization from a real data set. As a test area we selected a sub-site of the Ocean City test site. A more detailed description of this test site is presented by Csatho et al. (1998). The sub-site includes 4633 points with a point density of 1.2 points$/m^2$.

The data set, acquired by an airborne laser altimetry system, covers a small portion of an urban area in Ocean City. As Figure 4 illustrates, the sub-site contains a large building with complex roof structures.

The patches segmented at the signal level are visualized with the boundaries in Figure 5. The adjacency is defined by a sphere of a fixed radius (2.5 m) and then refined by eliminating the links between the outliers and inliers. The similarity of the plane parameters and the roughness are used as the grouping criteria. Each patch thus indicates a plane with certain roughness. The boundaries are then computed using the α-shape algorithm with $\alpha = 2.5$.

The merged surfaces with the preliminary and the refined boundaries are shown in Figure 6. The similarity of plane parameters and roughness with less strict threshold is used as merging criteria. The criteria are checked with the F-test and the resulting p-value is used for a measure indicating the tendency of merging. Merging starts from the patches with the highest measure. The merging process iteratively repeats until no adjacent patches can satisfy the criteria. The boundaries are also refined based on the intersections between surfaces. The breaklines and the occlusions are also identified.

The ground surface and polyhedral structure organized at the structural level are shown in Figure 7. The ground surface are polyhedral structures are identified based on the continuity graph and the 2D adjacency graph, respectively.

4 CONCLUSIONS

We recognized the need of an intermediate process common toward various applications using laser altimetry data. As the common process, we proposed generating a robust, abstract and explicit description from the raw data, called perceptual organization. For the process, we established a framework comprised of three organization processes at the signal, primitive and structural levels, represented as segmentation, merging and grouping, respectively. Furthermore, we elaborated the diverse components constituting the framework, inspired by the previous work on perceptual organization in various levels and domains.

The experimental results based on real data illustrate the outcomes expected at each level, demonstrate the good performance of the proposed approach and emphasize the need of perceptual organization as an intermediate process. A complete quantitative and computational analysis using various synthetic and real data will be performed for a reliable assessment about the performance. Furthermore, we will demonstrate the effectiveness of the proposed perceptual organization to higher-level processing by applying them to building reconstruction from urban data.

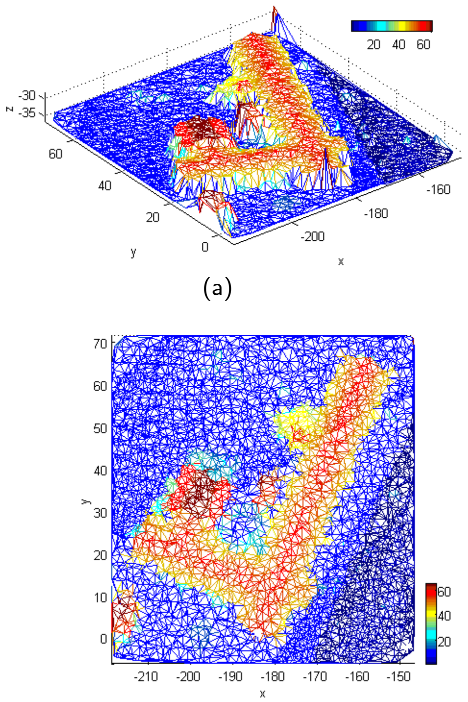
REFERENCES

Ahuja, N., Tuceryan, M., 1989. Extraction of early perceptual structure in dot patterns: integrating region, boundary, and component gestalt. Computer Vision, Graphics, and Image Processing 48 (3), 304–356.

Allmen, M., Dyer, C. R., 1993. Computing spatiotemporal relations for dynamic perceptual organization. CVGIP: Image Understanding 58 (3), 338–351.

Amir, A., Lindenbaum, M., 1998. A generic grouping algorithm and its quantitative analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (2), 168–185.

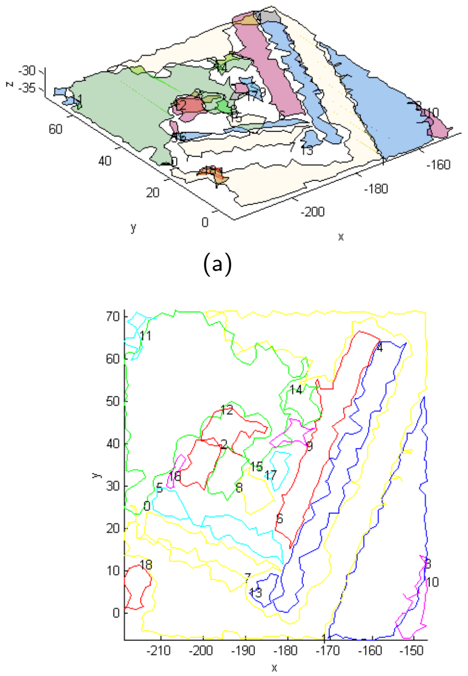
Axelsson, P. E., 1999. Processing of laser scanner data - algorithms and applications. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 138–147.



(a)

(b)

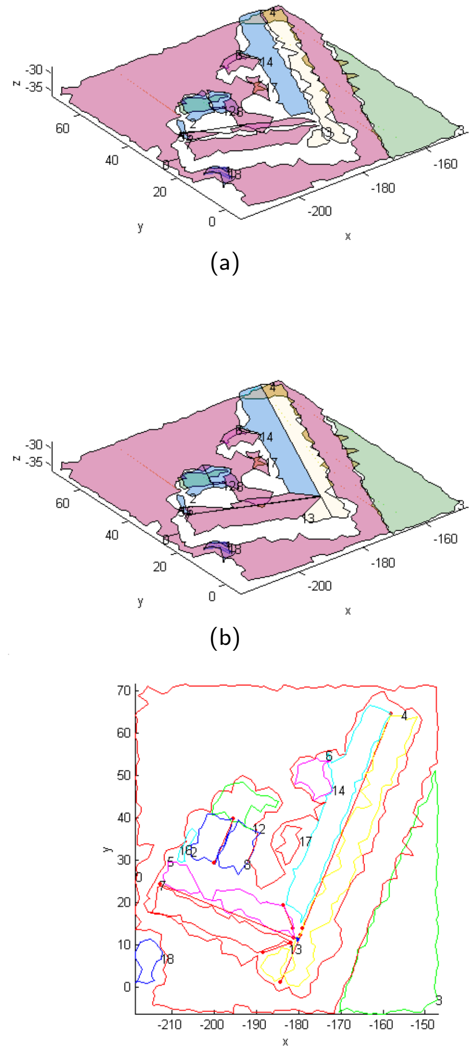
Figure 4: The test data is visualized as the mesh graph based on the Delaunay triangulation in (a) 3D and (b) 2D view.



(a)

(b)

Figure 5: The segmented patches organized at the signal level are visualized with the boundaries in (a) 3D and (b) 2D view.



(a)

(b)

(c)

Figure 6: The merged surfaces organized at the primitive level are visualized with the preliminary boundaries in (a) and with the refined boundaries (b). The preliminary boundaries, the intersections (red lines and dots) and the corners (blue dots) are also indicated in (c).

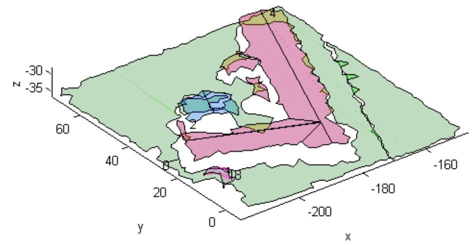


Figure 7: The ground surfaces (green) and the polyhedral structure organized at the structural level are visualized. The same color indicates the same group.

Berengolts, A., Lindenbaum, M., 2001. On the performance of connected components grouping. International Journal of Computer Vision 41 (3), 195–216.

Berg, M. d., 2000. Computational geometry : algorithms and applications, 2nd Edition. Springer, Berlin ; New York.

Boyer, K. L., Mirza, M. J., Ganguly, G., 1994. The robust sequential estimator: a general approach and its application to surface organization in range data. IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (10), 987–1001.

Boyer, K. L., Sarkar, S., 1999. Perceptual organization in computer vision: status, challenges, and potential. Computer Vision and Image Understanding 76 (1), 1–5.

Chang, Y.-L., Aggarwal, J. K., 1997. Line correspondences from cooperating spatial and temporal grouping processes for a sequence of images. Computer Vision and Image Understanding 67 (2), 186–201.

Chaudhuri, A. R., Chaudhuri, B. B., Parui, S. K., 1997. A novel approach to computation of the shape of a dot pattern and extraction of its perceptual border. Computer Vision and Image Understanding 68 (3), 257–75.

Chaudhuri, B. B., 1996. A new definition of neighborhood of a point in multi- dimensional space. Pattern Recognition Letters 17 (1), 11–17.

Csatho, B., Krabill, W., Lucas, J., Schenk, T., 1998. A multisensor data set of an urban and coastal scene. International Archives of Photogrammetry and Remote Sensing 32 (3/2), 26–31.

Edelsbrunner, H., 1992. Weighted alpha shape. Tech. Rep. UIUCDCS-R-92-1760, University of Illinois at Urbana-Champaign.

Edelsbrunner, H., Kirkpatrick, D. G., Seidel, R., 1983. On the shape of a set of points in the plane. IEEE Transactions on Information Theory 29 (4), 551–559.

Fisher, R. B., 1989. From surfaces to objects : computer vision and three dimensional scene analysis. J. Wiley, Chichester ; New York.

Gamba, P., Houshmand, B., 2000. Digital surface models and building extraction: a comparison of ifsar and lidar data. IEEE Transactions on Geoscience and Remote Sensing 38-2 (4), 1959–1968.

Geman, D., Geman, S., Graffigne, C., Dong, P., 1990. Boundary detection by constrained optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7), 609–628.

Guy, G., Medioni, G., 1997. Inference of surfaces, 3d curves, and junctions from sparse, noisy, 3d data. IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (11), 1265–1277.

Havaldar, P., Medioni, G., Stein, F., 1996. Perceptual grouping for generic recognition. International Journal of Computer Vision 20 (1-2).

Henricsson, O., 1998. The role of color attributes and similarity grouping in 3d building reconstruction. Computer Vision and Image Understanding 72 (2), 163–184.

Herault, L., Horaud, R., 1993. Figure-ground discrimination: a combinatorial optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (9), 899–914.

Hyyppa, J., Kelle, O., Lehikoinen, M., Inkinen, M., 2001. A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. IEEE Transactions on Geoscience and Remote Sensing 39 (5), 969–975.

Khler, W., 1929. Gestalt psychology. H. Liveright, New York.

Koffka, K., 1935. Principles of gestalt psychology, lst Edition. International library of psychology, philosophy and scientific method. Harcourt Brace and Company, New York,.

Koster, K., Spann, M., 2000. Mir: an approach to robust clustering-application to range image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (5), 430–444.

Lee, I., Schenk, T., 2001. Autonomous extraction of planar surfaces from airborne laser scanning data. In: Proc. ASPRS conference. St. Louis, MO, USA.

Lee, M.-S., Medioni, G., 1999. Grouping ., -, to , theta , into regions, curves, and junctions. Computer Vision and Image Understanding 76 (1), 54–69.

Lin, C., Huertas, A., Nevatia, R., 1994. Detection of buildings using perceptual grouping and shadows. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos, CA, USA, pp. 62–69.

Lowe, D. G., 1985. Perceptual organization and visual recognition. Kluwer Academic Publishers, Hingham, MA, U.S.A.

Maas, H. G., Vosselman, G., 1999. Two algorithms for extracting building models from raw laser altimetry data. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 153–163.

Matula, D., 1997. Graph theoretic techniques for cluster analysis algorithms. In: Classification and clustering. Academic Press, pp. 95–129.

Melkemi, M., Djebali, M., 2000. Computing the shape of a planar points set. Pattern Recognition 33 (9), 1423–36.

Melkemi, M., Djebali, M., 2001. Elliptic diagrams: application to patterns detection from a finite set of points. Pattern Recognition Letters 22 (8), 835–44.

Modayur, B. R., Shapiro, L. G., 1996. 3d matching using statistically significant groupings. In: Proc. International Conference on Pattern Recognition. Los Alamitos, CA, USA, pp. 238–242 vol.231.

Murakami, H., Nakagawa, K., Hasegawa, H., Shibata, T., Iwanami, E., 1999. Change detection of buildings using an airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing 54 (2-3), 148–152.

Nelson, R. C., Selinger, A., 1998. A cubist approach to object recognition. In: International Conference on Computer Vision. New Delhi, India, pp. 614–621.

Richards, F. M., Mullins, L. J., 1977. Areas, volumes, packing, and protein structure. In: Annual review of biophysics and bioengineering, vol.6. Palo Alto, CA, USA, pp. 151–76.

Sarkar, S., 1995. Tracking 2d structures using perceptual organizational principles. In: Proc. International Symposium on Computer Vision. Los Alamitos, CA, USA, pp. 283–288.

Sarkar, S., Boyer, K. L., 1993. Perceptual organization in computer vision: a review and a proposal for a classificatory structure. IEEE Transactions on Systems, Man and Cybernetics 23 (2), 382–399.

Sarkar, S., Boyer, K. L., 1994a. A computational structure for preattentive perceptual organization: graphical enumeration and voting methods. IEEE Transactions on Systems, Man and Cybernetics 24 (2), 246–267.

Sarkar, S., Boyer, K. L., 1994b. Computing perceptual organization in computer vision. Series in machine perception and artificial intelligence ; v. 12. World Scientific, New Jersey, USA.

Sarkar, S., Boyer, K. L., 1998. Quantitative measures of change based on feature organization: eigenvalues and eigenvectors. Computer Vision and Image Understanding 71 (1), 110–136.

Sarkar, S., Soundararajan, P., 2000. Supervised learning of large perceptual organization: graph spectral partitioning and learning automata. IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (5), 504–525.

Schenk, T. F., 1999. Digital photogrammetry. TerraScience, Laurelville, OH, USA.

Shaashua, A., Ullman, S., 1988. Structural saliency: the detection of globally salient structures using a locally connected network. In: Proc. International Conference on Computer Vision. Washington, DC, USA, pp. 321–327.

Shapiro, L. G., Haralick, R. M., 1979. Decomposition of two-dimensional shapes by graph-theoretic clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence (1), 10–20.

Tang, C.-K., Medioni, G., 1998. Inference of integrated surface, curve and junction descriptions from sparse 3d data. IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (11), 1206–1223.

Tao, C. V., Hu, Y., 2001. A review of post-processing algorithms for airborne lidar data. In: Proc. ASPRS conference. St. Louis, MO, USA.

Toussaint, G. T., 1988. A graph-theoretical primal sketch. In: Toussaint, G. T. (Ed.), Computational Morphology. North-Holland, Amsterdam, pp. 229–260.

Witkin, A., Tenenbaum, J., 1983. On the role of structure in vision. In: Beck, J., Hope, B., Rosenfeld, A. (Eds.), Human and Machine Vision. Academic Press, New York, pp. 481–543.

Wu, Z., Leahy, R., 1993. An optimal graph-theoretic approach to data clustering - theory and its application to image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (11), 1101–1113.

Zahn, C. T., 1971. Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Transactions on Computers (1), 64–82.

Zisserman, A., Mundy, J., Forsyth, D., Liu, J., Pillow, N., Rothwell, C., Utcke, S., 1995. Class-based grouping in perspective images. In: Proc. International Conference on Computer Vision. Los Alamitos, CA, USA, pp. 183–188.