High-Level Surface Descriptions from Composite Range Images

Vítor Sequeira†, João G. M. Gonçalves
European Commission
Joint Research Centre, 21020 Ispra (VA), Italy

M. Isabel Ribeiro
Institute for Systems and Robotics
Instituto Superior Técnico, Lisboa, Portugal

Abstract

This paper presents a new scene analysis system that automatically reconstructs the 3D geometric model of real-world scenes from multiple range images. The reconstruction is achieved through a fully comprehensive procedure that includes range data acquisition, geometrical feature extraction, registration and integration of multiple views. A new hybrid algorithm for the segmentation of range images has been developed to overcome the specific problems specific of this kind of images. This algorithm combines edge and region detection approaches to build a high-level 3D description of a real scene described by a noisy range image. Another relevant issue is the registration/integration of range images from different viewpoints into a consistent 3D model. The integration part combines registered and partially overlapping data sets into a full high-level surface representation. Relevant experimental results from the complete 3D scene modelling are shown.

1 Introduction

Modelling 3D scenes from 2D image sequences has been a research topic for a long time as reported in the overview of the field presented by Aggarwal and Nandhakumar [1]. The reported techniques are however inherently underconstrained and require assumptions about the scene. An important drawback in intensity images is the lack of explicit depth information preventing a direct relationship between the image and the real world. On the contrary, active range image systems provide direct geometric information about the shape of visible surfaces. Consequently problems in 3D object description and recognition can be more easily solved using range images rather than intensity images.

Automatic shape acquisition based on range images has progressed significantly in the last few years. Most recent research work aims at modelling small parts, usually for industrial applications. The range acquisition system is normally based on structured light techniques with the objective of identifying and/or recognising a given object so that it can be counted, sorted or grasped by a manipulator arm.

The 3D scene analysis system presented in this paper is based on data acquired by a Laser Range Finder, aiming at modelling real-world complex scenes (distance range from 0.5 to 20 meters). Previous work on scene modelling for indoor mobile robotics applications was done using vision based methods such as depth from stereo or motion [2].

Scene modelling is at the end of a long chain of processing steps. This chain starts with range data acquisition followed by the extraction of geometric features from range images. Feature extraction is required to provide accurate localisation of edges, specially for those corresponding to surface discontinuities, and to guarantee that edges are correctly classified (e.g., jump, crease). If a complete volumetric model is to be achieved, range data acquired from multiple viewpoints is required to solve all the ambiguities caused by occlusions. This raises the problem of registration and integration of range images into a single model, since a complete model is more than the sum of individual models, and should integrate geometric features from single views.

The paper describes the automatic reconstruction of a 3D geometric model of an unknown environment, e.g indoor scenes, based on data acquired by a Laser Range Finder on board of a mobile robot. Such a description is most useful for visualisation, scene understanding, mobile robot navigation and verification purposes. The paper presents a new hybrid segmentation algorithm that combines edge detection and region growing approaches to build a high-level description of a scene given a noisy range image. Range data is segmented into their constituent parts in terms of bi-quadratic surfaces. Multiple range images from different viewpoints are acquired to resolve the ambiguities caused by occlusions and/or to introduce higher resolution images in particular regions of interest. Registration and integration are the two main issues in creating a single model from multiple range images. A hierarchical registration scheme working directly at the pixel level was implemented for higher accuracy [3]. The integration procedure combines registered and partially overlapping data sets into a final surface representation. The integration scheme is powerful in that the combination of views may be done at the level of the final representation.

The paper is structured along the computational steps that define the current procedure for building a 3D model scene from range data. Section 2 briefly reviews relevant results of range image acquisition and error sources of the acquisition system. Section 3 concentrates on the extraction of surface characteristics from 3D noisy data. It presents a range image segmentation approach that combines edge and region based techniques. Section 4 focuses on the problem of combining data from multiple views and merging the corresponding surface description into a consistent 3D model. Along sections 3 and 4, relevant experimental results are presented. Conclusions are drawn in section 5.

†PhD program supported by Programa Praxis XXI - JNICT - Portugal

2. Range Image Acquisition

In the present work, range images are captured using a time-of-flight Laser Range Finder (LRF) mounted on a computer controlled pan-and-tilt unit on board of a mobile robot (see Fig. 1).

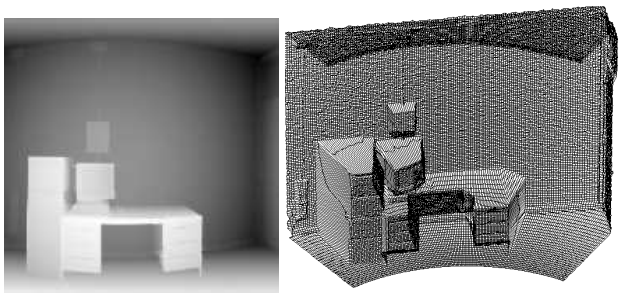
Figure 1: LRF on board of the mobile robot

The LRF uses an eye-safe laser beam with 4 cm diameter and a divergence of 0.16°. The measuring system has a distance range from 0.5 to 20 meters. Range images are acquired within a user-established solid angle and distances measured on a grid with equally spaced pan and tilt angles. A typical scan has 140x140 range measurements over a 60°x 60° field of view for a scanned area of 4x4 meters. Fig. 2 a) displays a grey level representation of a typical range image from an office scene in which lighter areas correspond to shorter distances. When displayed with a regular grid, range images suffer from deformation associated to the geometry of the scanning system (e.g., curved line between the wall and floor), and therefore a geometrical correction must be performed to obtain data in a format independent from the acquisition system. A 3D Cartesian view of the same range image with geometrically corrected data is shown in Fig. 2 b).

Several error sources are present on the range acquisition system. Some result from the characteristics of the LRF, namely measurement noise and limited measurement resolution (<2 cm). Others are inherent to the scanning system. A main limitation in the spatial and range resolution is due to the fact that the laser beam is not an ideal line. The distance information of each pixel is an average of range values over a 2-D area, the laser footprint, resulting from the intersection of the cylinder beam with the target surface. The size of the footprint also depends on the angle between the surface normal and the beam. A major effect that distorts the image at specific locations (occluded edges) is the mixed point problem, where the laser footprint crosses an edge between two objects [4]. In this case, the range measurement is a combination of distances to the two objects, having no physical meaning. This mixed point effect (see left and top of the file cabinet in Fig. 2 b) creates some artefacts rendering edge detection more difficult.

3. Hybrid Approach to Range Image Segmentation

The LRF system provides an image containing the explicit 3D surface information of the scene being scanned. This 3D point-based surface information must be further processed by a segmentation procedure to obtain a more abstract surface representation. The segmentation process is a data reduction



(a) (b)

Figure 2: Range image from an office scene: a) Grey level display b) 3D perspective view

mechanism converting sensorial measurements into some pre-defined primitive elements representing the available knowledge. To construct a geometric model, at least one geometric entity (e.g., surfaces, edges and vertices) must be extracted. Range image segmentation techniques described in the literature, may be classified under two major categories: a) edge or surface discontinuity based, and b) surface or region based. Edge segmentation based techniques [5] attempt the detection of edges or surface discontinuities in the range image. Region based techniques [6,8] detect regions corresponding to continuous surfaces. Most region-based techniques detect only planar surfaces, and since their goal is to detect regions rather than edges, resulting region boundaries do not necessarily correspond to the projection of surface discontinuities. Furthermore, region-based methods often do not classify region boundaries, and cannot detect curvature curves extreme. Results obtained with edge-based methods are usually more closely related to real edges than those obtained with region-based methods. Detected edges are however often fragmented, no matter which method is used. Edge linking and polygonal approximation are additional processing steps required to organise fragmented edges into a structured description corresponding to relevant scene features.

The roles of edge and region based techniques can be seen as complementary, as they require the detection and identification of both surface discontinuities and continuous surfaces. This suggests the use of a segmentation method in which each approach compensates for the drawbacks of the other. The following sub-sections describe a hybrid algorithm combining edge and region detection approaches for range image segmentation. The algorithm outputs both a segmentation map and an edge map. In addition, the edge map conveys C^0 (jump) and C^1 (crease) discontinuity information, and the segmentation map provides information on adjoining surfaces that resulted in the formation of edges and defines boundary conditions. This information will be most useful for occlusion detection. Fig. 3 is an overview of the implemented algorithm.

3.1 Edge-Based Initial Segmentation

Methods for detecting edges without computing surface curvatures are divided in two categories:

i) detecting extreme or zero-crossings of surface derivatives,
ii) geometric model fitting.

Surface derivative methods [7], suffer from edge displacement because an implicit smoothing is used before the computation of derivatives. The basic idea of model fitting

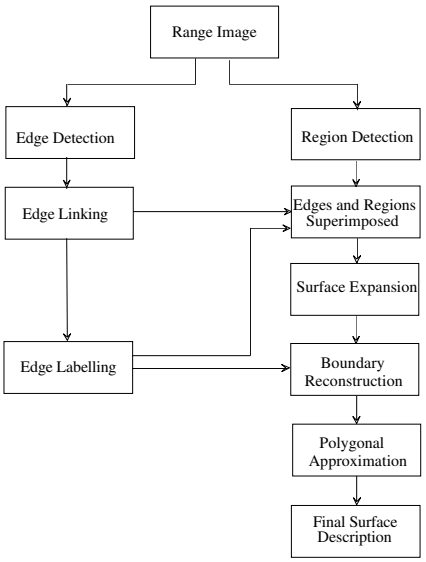


Figure 3: Flowchart of the hybrid segmentation algorithm.

methods [5], is to fit pre-defined edge models directly to data without pre-smoothing, and then make a decision to either accept or reject the presence of such models. These methods explicitly give a mechanism for classifying the detected edges.

For the current application the edges are detected on the 3D space based on geometric model fitting, because good localisation is a requirement for feature extraction in geometric modelling, specially for the edges corresponding to surface discontinuities. The discontinuity detection for 1D sampled points can be summarised as follows:

i) Fit a curve to each of the left and right hand neighbourhoods of every adjacent pair of points,
ii) Eliminate those points whose neighbourhoods contain a discontinuity by performing non-minimum suppression on the values of the sum of the fit errors,
iii) Apply an F-distribution test to the fitted curves of each remaining point to decide whether it is a discontinuity.

Step ii) prevents the fitting (smoothing) from crossing a potential discontinuity.

Given N observations of an nth order polynomial corrupted by white Gaussian noise u_i:

$$y_i = b_0 + b_1 x_i + b_2 x_i^2 + \cdots + b_n x_i^n + u_i, \quad i = 1, \ldots, N, \quad (1)$$

the maximum likelihood estimates of the coefficients, can be computed using the least-squares method. Having estimated the coefficients of left and right polynomials and having eliminated all false points, it is determined whether b_{l0} is significantly different from b_{r0} (a C^0 (jump) discontinuity) and whether b_{l1} is significantly different from b_{r1} (a C^1 (crease) discontinuity), by testing the following two hypothesis:

$$H_1: \quad b_{l0} = b_{r0} \text{ and } b_{l1} = b_{r1};$$
$$H_2: \quad b_{l0} = b_{r0}.$$

If H_1 is accepted, both coefficients are equal and the function is smooth. If H_1 is rejected, then H_2 is tested. If H_2 is accepted there is a C^1 discontinuity. If H_2 is also rejected, there is a C^0 discontinuity. Hypothesis H_1 is tested by performing a least square fitting to the constrained model:

$$\begin{bmatrix} Y_l \\ Y_r \end{bmatrix} = \begin{bmatrix} 1 & X_l & X_l^2 & 0 & \cdots & X_l^n & 0 \\ 1 & X_r & 0 & X_r^2 & \cdots & 0 & X_r^n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_{l2} \\ b_{r2} \\ \vdots \\ b_{ln} \\ b_{rn} \end{bmatrix} + \begin{bmatrix} U_l \\ U_r \end{bmatrix} \quad (2)$$

If the sum of the fit errors resulting from this constrained model is ε_a^2, and ε_s^2 is the sum of the fit error resulting from the unconstrained fit of the left and right polynomials, then the

$$\frac{\left(\varepsilon_a^2 - \varepsilon_s^2\right)/2}{\varepsilon_s^2/\left(2N - 2(n+1)\right)} \quad (3)$$

statistic follows the $F_{(2,2N-2(n+1))}$ distribution. Thus H_1 is rejected if this statistic is larger than a given confidence level of the distribution. The hypothesis H_2 is then tested with a similar approach.

To detect discontinuities in a range image, this procedure is generalised to 3D space applying the described procedure to each scan line (both horizontal and vertical) in 3D Cartesian coordinates. With this technique, detected edge points are labelled as jump or crease edges. For crease edge points, it is further decided if they are convex or concave by comparing their fitted curves at the left and right-hand neighbourhoods.

The previous procedure considers that a detected jump edge belongs to the limit of the surface which is closer to the sensor. A relationship between the referred jump edge and the far way surface from the sensor is established, aiming at providing boundary limits for both adjacent surfaces along the jump edge. These boundary limits are used in the combination of range images from multiple views for planning the next view. After being extracted, discontinuity edge points are linked into edges with the same label on the basis of m-connectivity or mixed connectivity. This linking is performed in such a way that range values of resulting edges are continuous in depth.

3.2 Region-Based Initial Segmentation

The surface segmentation module uses a technique proposed by Leonardis et al [8] in which range images are segmented into piecewise continuous surfaces. Data aggregation is done via model recovery. The process of recovering the primitives from data is performed by following the iterative regression approach described by Besl [6], although selecting different initial estimates (seed regions). Besides, this method considers the recovered models as hypothesis which then compete to be selected in the final description. The models used for representation are variable-order bivariate polynomials that are linearly parameterizable in the Euclidean space as,

$$f(r, \mathbf{a}, \mathbf{x}) = \sum_{0 \le i+j \le r} a_{ij} x^i y^j \quad (4)$$

where the vector **a** is defined in a parameter space with dimension r restricted to $0 \le r \le 2$. Thus the model admits planar and biquadratic surfaces.

Planar patches (zero or first-order) are represented as

$$f(x, y) = a_{10} x + a_{01} y + a_{00} \quad (5)$$

while curved patches (second-order), have the general form

$$f(x,y) = a_{20}x^2 + 2a_{11}xy + a_{02}y^2 + 2a_{10}x + 2a_{01}y + a_{00}. \quad (6)$$

The curved surfaces belong to the class of non-central quadratics. The sign of the following invariant quantity, *J*,

$$J = \begin{vmatrix} a_{20} & a_{11} \\ a_{11} & a_{02} \end{vmatrix} = a_{20}a_{02} - a_{11}^2, \quad (7)$$

with respect to translation and rotation in the image plane determines the surface type.

The region segmentation process comes to a global description of the image through an iterative process. The process starts by placing a grid pattern of seed surface patches across the smoothed image. For each seed patch a bi-variate polynomial surface is fit and the fit error compared with an error threshold. Bi-variate polynomial surface fitting is accomplished through general linear least squares regression.

The overall efficiency of the method is achieved by combining model recovery and model selection in an iterative way. The selection of the surface patches on which to continue processing is based on maximazing the objective function:

$$F(m_i) = \mathrm{K}_1 n_i - \mathrm{K}_2 \xi_i - \mathrm{K}_3 N_i \quad (8)$$

where n_i is the number of points in the ith model (benefit), ξ_i is the global least square error (cost), and N_i is the order of the polynomial selected (cost). K1, K2, K3 are weights which can be adjusted to give more preference to a particular description. The complexity of the procedure is proportional to the number of currently valid models multiplied by the time needed to evaluate the optimisation matrix.

3.3 Final Segmentation

The final segmentation/representation is based on a two steps approach for the integration of the surface and edge representations:

i) Superimposing the 3D edge onto the 3D region map,

ii) Expansion of surface regions.

It should be noted that the integration of the edge and surface maps is done in 3D Cartesian coordinates. The results obtained by the initial segmentation method using the model *z=f(x,y)* are used to derive the surface properties of the final view independent representation.

In the initial region-based segmentation, the region map is a 3D labelled image in which each point has a region number as its label. The 3D edge map, in which each edge point has a label depending whether it is a jump or a crease edge, is then superimposed onto the region map. The resulting surface-edge map is composed of surface points with region numbers as their labels and edge points with jump or crease labels.

Each surface region in the previously obtained surface-edge 3D map is expanded and/or contracted to generate a final 3D region map. The following steps are used in the expansion procedure: First the contour lines of each surface are extracted using a scan line algorithm. Then, each contour line is expanded aiming at minimising the Euclidean distance between the contour points and the edge points. This expansion is constrained to the surface parameters, i.e., it follows the surface tangent at the contour points, and may not cross the boundaries of other surface regions in the 3D surface map. As a result, small surface regions on discontinuities are eliminated because they will collapse to the corresponding edge. The final boundary points are then ordered and a 3D curve aproximation is performed on the range values of the extracted boundary. The algorithm for 3D curve approximation recursively refine an initial polygonal approximation of the curves where the aproximation error is larger than a given threshold. Such a refinement is performed by adding those points where the approximation error is local maxima in the previous approximation.

3.4 Experimental Results

Real images of indoor scenes used in the experiments were obtained with the time-of-flight LRF described in section 2. The images are originally provided as a 2D array of pixels (polar coordinates) with a range resolution of about 2 cm and floating-point resolution. The scanning system used to acquire images is based on a spherical coordinate system. Image pixels map real world points using two angles and a distance *(i,j,r(i,j))* Thus a geometrical correction transformation of data to the system's Cartesian coordinates *(x,y,z)* is performed prior to processing. Fig. 4 present two examples of surface extraction of 3D real world scenes. The algorithm processed real raw data without any preprocessing and thresholds were kept constant for all the examples presented in the paper. A range image with geometrically corrected data is shown in (a). Edge detection results are shown in (b). They were obtained for a fitting polynomial of degree 1 on 4 point left and right neighbourhoods, and using a 0.99 statistical confidence level of the F-distribution, on the residual errors. It can be seen that both jump and crease edges were well detected. However, some of the jump edges are very noisy (see top of the cupboard in (b)) due to the mixed point effect discussed in section 2. Initial 2 1/2 D surface segmentation in terms of bi-quadratic surfaces is represented in (c). Final 3D boundary reconstruction and full view independent 3D surface representation obtained by the hybrid algorithm are displayed in (d) and (e) respectively. Different display perspectives were used to clarify the relevant aspects of each figure. The final surface representation contains explicit information about the boundaries of the segmented data. The good localisation of the edge detection scheme enhances the accuracy of the constructed boundaries. The white areas on the final (e) representation are caused by the range shadows (occlusions) or show up in areas with ambiguous data where no useful information is available. To resolve the ambiguities caused by occlusions, images from different viewpoints are required (see section 4). Small scene details (e.g., thermostat on the wall in the first example of Fig. 4) are considered as noise and are not present on the final representation, because it is not possible to simultaneously extract large features for scene modelling and small object shapes. It should be stressed that reconstructed models must extract meaningful information from noisy sensor data while making the weakest possible assumptions about observed shapes. Nonetheless, it is possible to work on the residual image, detecting differences between the model and the raw image, to find non modelled areas and reprocess them either by changing the parameters for small object shapes or by re-acquiring those areas with higher resolution.

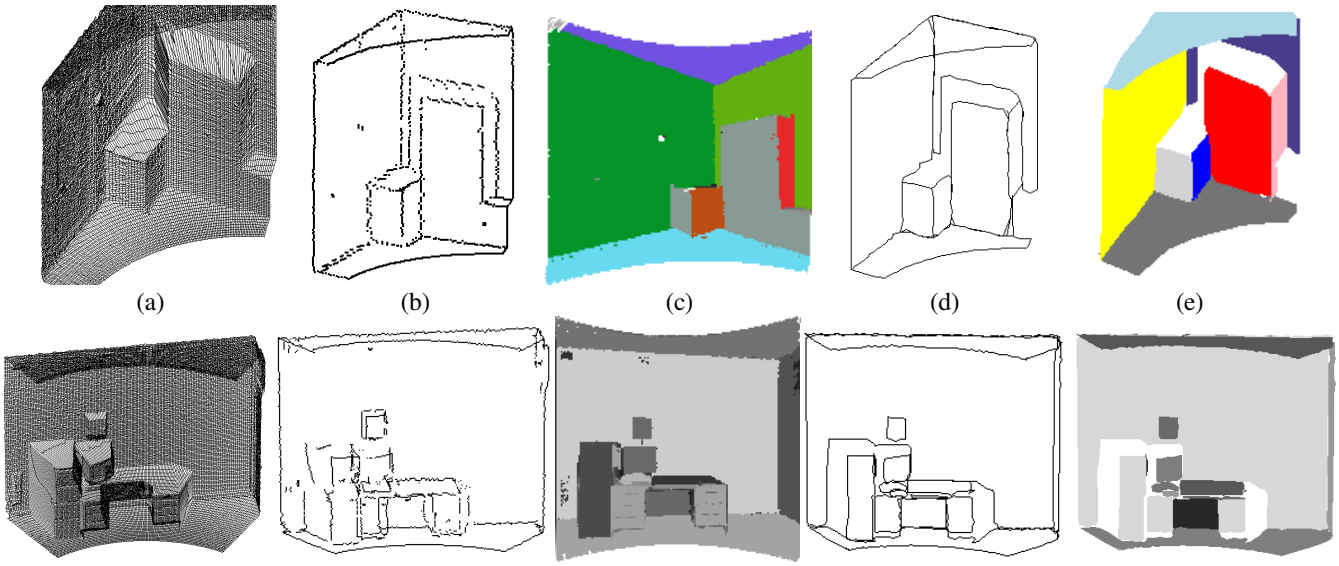
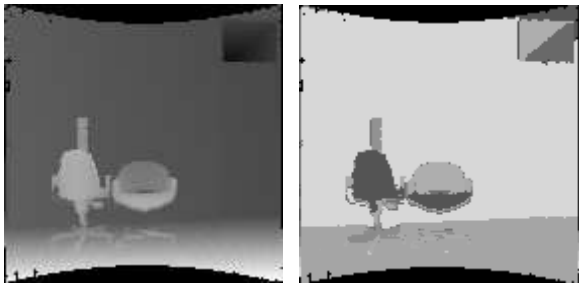


Fig. 4: Surface representation from a single range view

Fig. 5 refers to a single view of a scene with non plane surfaces (two chairs in different orientations and a vertical cylinder on the top of the left chair). The range image after geometrical correction is displayed in Fig. 5a. The surface representation in Fig. 5b clearly distinguishes different curved patches (concave and convex).



(a) 3D range image (b) surface description

Figure 5:Surface Representation from a single range view: curved patches.

4 Combining Multiple Range Images

Combining range images taken from multiple viewpoints into a single representation is essential for building a complete 3D model of a scene. One range image is only a portion of the information present in a 3D working domain. To resolve the ambiguities caused by occlusions, images from different viewpoints are required. Occlusions are automatically detected based on surface discontinuities (jump edges).

Registration and integration are the two main issues in creating a single model from multiple range images. Registration refers to computing the rigid transformation that brings the points of one range image into alignment with the overlapping parts of a second range image. An iterative registration algorithm was implemented working directly on range data to find a 3D transformation which minimises a distance measured between two geometric shapes being registered. Details about the algorithm, that estimates the least squares rotation matrix and the translation vector using the dual number quaternions [9], can be found in [3]. Integration is the process of creating a single representation (model) of the sample points from several range images.

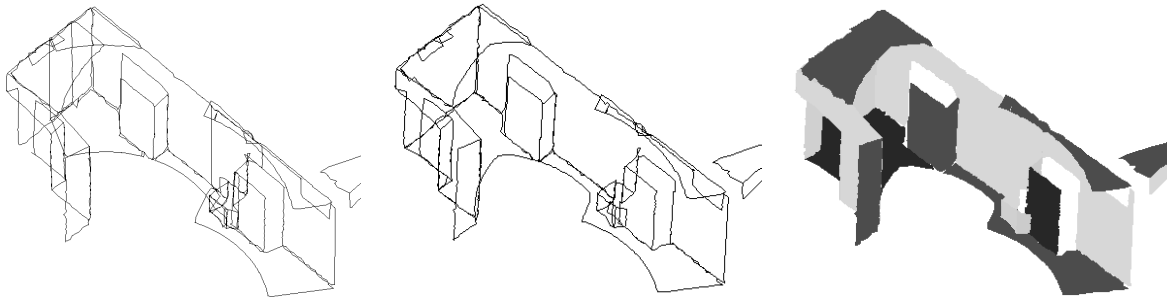
4.1 Integration of Views

The integration part combines registered and partially overlapping data sets. Integration of different views can be conducted at various abstraction levels, for example at pixel level, at surface patch level or using a contour propagation technique [3]. The problem with pixel level fusion is that the image structure may be lost once the fusion is performed, preventing the use of efficient geometric techniques for extracting surface characteristics. To overcome this problem, the integration method used in the present work is made at surface level, by refitting a surface model to the fused surface points of the overlapping surfaces. At this level the image structure is no longer required. The integration of surface descriptions can be summarised as follows:

1. Construct a surface description for each view using the hybrid algorithm.
2. Find overlapping surfaces. The surface descriptions of the different views were previously transformed to a common coordinate frame using the 3D rigid transformation obtained from the registration process.
3. Merge points from the partially overlapping surfaces.
4. Refit a surface model to the merged surface points.
5. Fusion of edge points. Expansion of the merged surface to the edge points.
6. Boundary reconstruction.

After the registration of neighbouring views (partially overlapping), a surface description using the hybrid algorithm is constructed for each view. The surface descriptions of the different views are then transformed to a common coordinate frame. For constructing the model resulting from the combination of the multiple views, an average of the samples that are in the overlapping surfaces of the views is performed. The region fusing procedure enables the elimination of occluded areas (or areas with noisy and/or ambiguous data in the raw range image).

A complete 3D surface model resulting from the integration of four range images of a laboratory scene (scanned area of 10x4 m, see Fig. 1) is presented in Fig. 7.



(a) boundary before merging (b) boundary after merging (c) final surface description

Figure 7: High-level surface description from composite range images

5 Discussion

Range data should be always associated to the intrinsic geometry of the scanning mechanism: range images will be different if scanning mechanisms with different geometry are used. A geometrical correction phase is needed prior to processing to obtain data in a format independent from the acquisition system. The geometrical corrected representation of the range image in 3D Cartesian coordinates exhibits a number of problems due to the conversion of a regularly sampled image to a different reference plane. Since some kind of image warping is performed, the distribution of data points is not uniform and, as result, conventional image processing algorithms cannot be applied directly to the range image with geometrically corrected data.

Edges detected by the edge detector method possess good localisation properties. This helps in steering the region growing process towards accurate border partitioning. In addition, the incorporation of the region growing process eliminates internal microedges and provides for missing border reconstruction due to its ability to detect weak edges. The combination of both methods gives stable surfaces, eliminates internal microedges and provides for missing border reconstruction leading to a more reliable representation.

The final model admits planar and biquadratic surfaces. For the type of environment under study (indoor environments) most objects can be represented by planar patches bounded by linear edges, no higher order models being required. Besides, higher order surfaces are computationally expensive, are often unstable and, even if they are considered, there is no guarantee that they will always fit to the data.

The resulting scene description can be directly converted to a CAD file format (e.g., DXF) so that can be manipulated or displayed using conventional CAD programs.

No a priori information about the scene is used for the segmentation process. Future work will integrate other sensory information such as the reflectance image to make the overall process more robust and reliable, and to have more realistic models by rendering the geometric models with natural texture

6 Conclusions

A 3D scene geometric modelling approach using multiple range images has been described in this paper. The system uses real range images acquired by a Laser Range Finder onboard a mobile robot.

The paper presents all the processing steps needed for the description of the environment such as range image acquisition, geometric processing and procedures for constructing surface descriptions. Another important component is the information integration procedure which constructs a full surface description from different range views.

A new hybrid segmentation algorithm that combines edge detection and region growing techniques was presented. The major novelty of the approach lies in an effective combination of different algorithms to achieve the final surface representation. Local and global properties such as surface type, edge type and surface parameters are part of this representation. Experiments on a large number of real range images indicate that this approach gives better results than traditional edge or region detection algorithms, producing good models in the presence of noise. A complete 3D model reconstruction from multiple range views of a polyhedral environment (office scene) is presented. The use of bi-quadratic patches is demonstrated through an example with concave and convex shapes in the environment.

The 3D scene analysis system presented here was primarily designed for CAD modelling and design verification of buildings and plants although it can be used in robot navigation and scene understanding or as an input to virtual reality systems.

References

[1] J. K. Aggarwal, N. Nandhakumar - On the Computation of Motion from Sequences of Images-A Review, Proc. of IEEE, vol. 76, n. 8, pp. 917-935, August 1988.

[2] X. Lebegue, J. K. Aggarwal - Automatic Creation of Architectural CAD Models, Proc.IEEE Second CAD-Based Vision Workshop, pp. 82-89, February 1994.

[3] V. Sequeira, J.G.M. Goncalves, M.I. Ribeiro - 3D Modelling of In-Door Scenes using Laser Range Sensing, Proc. 4th IAPR Intern. Worksh. Machine Vision Apps., Japan, pp.315-318, Dec. 1994.

[4] V. Sequeira, J.G.M. Gonçalves, M.I. Ribeiro - 3D Environment Modelling Using Laser Range Sensing, Proc. of Int. Symposium on Intelligent Robotic Systems'94, France, pp. 238-245, July 1994.

[5] Y. G. Leclerc, S. W. Zucker - The Local Structure of Image Discontinuities in One Dimension, IEEE Trans. Patt. Analysis and Machine Intell., vol. 9, n. 3, pp. 341-355, 1987.

[6] P. J. Besl, R. C. Jain - Segmentation Through Variable-Order Surface Fitting, IEEE Trans. Patt. Analysis and Machine Intell., vol. 10, n. 2, pp. 167-192, March 1988.

[7] J. Canny - A Computational Approach to Edge Detection, IEEE Trans. on Patt. Analysis and Machine Intell., vol. 8, n. 6, pp. 679-698, November 1986.

[8] A. Leonardis, A. Gupta, R. Bajcsy - Segmentation as the Search for the Best Description of the Image in Terms of Primitives, Proc. 3rd Int. Conf. on Computer Vision, pp. 121-125, 1990.

[9] M. W. Walker, L. Shao, R. A. Volz - Estimating 3-D Location Parameters Using Dual Number Quaternions, CVGIP:Image Understanding, vol. 54, n. 3, pp. 358-367, 1991.