

AN INTEGRATED APPROACH TO SEGMENTATION OF RANGE IMAGES OF INDUSTRIAL PARTS

Arun K. Sood and Ezzet Al-Hujazi

Dept. of Computer Science, George Mason University
4400 University Dr., Fairfax, VA 22030-4444**ABSTRACT**

The first step in a general 3-D vision system is the segmentation of a digital image into a number of regions that correspond to physical scene surfaces. To compensate for the difficulties associated with edge and region based segmentation methods, we present an integrated approach for range image segmentation. In the first stage we detect jump in the range image. The second stage, involves detecting fold edges using the absolute value of the residual and surface normals. We have used a Bayesian approach for the region growing part of our algorithm. Markov Random Field is used to model the a priori knowledge information in the Bayesian formulation. A number of experimental results show that this approach is very effective in segmenting a wide variety of range images.

1. INTRODUCTION

Computer vision researchers have been investigating object recognition and scene understanding since the early 1970's. Most of the previous research has limited itself to intensity images. During the last ten years, digital range images have become more available. Range images are unique in that they directly measure the shape of 3-D objects. Range images can be obtained from a variety of active (like laser) and passive (like stereo) sensors.

The use of data driven early processing techniques of image data is important for a widely used vision systems. This requires that no restrictive high level models be assumed in early processing stages of image analysis. For image segmentation, the previous research work can be classified into two groups. Edge detection based methods look for the dissimilarity of neighboring pixel points to detect edges while region based methods look for the similarity of neighboring pixel points to segment the image.

There are a number of problems with edge based and region based methods. For range images, due to the difficulties in detecting fold edges, range image segmentation based solely on edge detection is not robust. For region based methods^{3,6}, the problems of determining good seed regions, defining an appropriate similarity criteria for regions pixels and performing the region growing in the proximity of fold edges are difficult to solve in the presence of noise and slowly varying edges. Due to these difficulties integration of these two techniques is important to achieve the requirements of a flexible robust vision system. In this paper we present an approach for integrating both edge based and region based methods for range image segmentation. The approach also combines a number of features to detect edges in range images. The algorithm is data driven to make it applicable to a wide range of applications.

The integrated segmentation approach proceeds in a sequential fashion. We first detect jump edges. The resulting jump edge map is then cleaned using a relaxation algorithm that is based on the Maximum a posteriori (MAP) estimator. Regions isolated by jump edges are excluded from further processing if they pass a goodness of fit test. For fold edge detection we realize that no single edge detection technique is sufficient to detect fold edges reliably. In Al-Hujazi and Sood² we have developed the mathematical basis for using the absolute value of the residual (AR) for fold edge detection. We have used AR in addition to normal to detect fold edges. The resulting fold edge map is cleaned using an approach similar to that in cleaning jump edge map.

The remainder of this paper is divided into three sections. In the next section we describe our approach to detecting edges in range images. Section 3 describes the region growing part of the algorithm. In section 4 we describe the algorithm steps. Experimental results on a number of range images are presented in section 5. Finally, some conclusions are presented in section 6.

2. DETECTING EDGES IN RANGE IMAGES

Edge detection techniques play an important role in range image analysis. There are three basic edge types in range images: jump, fold and smooth edges. Jump edges correspond to discontinuities in depth values. Fold edges correspond to surface creases where the surface normals are discontinuous. Smooth edges are characterized by continuity of surface normals but discontinuity of curvature. Most of the previous research work has been concentrated on detecting the first two types of edges. Detection of fold and smooth edges is very difficult because they do not correspond to large depth variation and consequently, tend to hide in noise. In the algorithm presented here we concentrate on detecting jump and fold edges only.

A variety of methods are available for jump edge detection in the literature. The problem of detecting jump edges in range images of industrial objects is simpler than that of detecting fold and smooth edges. In the algorithm presented here jump edges are detected as follows: For each pixel a plane is fitted to a 3x3 region. The fit error is then determined and if it is high the pixel is marked as a possible jump edge. The threshold used for jump edge detection is determined from the noise standard deviation.

Our study of the use of AR¹ shows that this feature is very effective in detecting edges but it has some difficulties at the proximity of a corner or near interacting edges. Fold edges can also be detected by observing the surface normals variation. We have combined AR with surface normal to detect fold edges. We have observed that the use of this multiple approach to detect fold edges is effective and produces better results than would have been obtained using each one of them separately.

3. BAYESIAN APPROACH TO REGION GROWING

In this section we present a Bayesian approach for the region growing part of the algorithm. The Bayesian optimal estimator approach requires the development of: 1) Model to encode apriori knowledge. 2) Stochastic models for the observation. 3) MAP estimator. 4) Algorithms for the computation of these estimates. Next, we describe requirements 1 through 3, the algorithm for computing the estimate will be described in the next section.

3.1 MODELING THE APRIORI KNOWLEDGE

In the work presented here, we have used Markov Random Field (MRF) on a lattice to model our apriori knowledge^{4,5}. MRF is a direct extension of a Markov process to higher dimensions and originated in the work of Ising⁴. Most useful for our purpose is the definition of a discrete MRF, a generalization of the concept of a Markov chain. A discrete MRF on a finite lattice is defined as a collection of random variables, corresponding to the sites of the lattice, whose probability distribution is such that the conditional probability of a given variable having a particular value given the rest of the variables, is identical to the conditional probability given the values of the field at a small set of sites.

A major difficulty in applying MRF formulation is in the definition of a valid conditional distribution. An alternative way of defining a MRF is based on the Markov-Gibbs equivalence established by the Hammersley-Clifford theorem. Before stating the theorem we need the following definitions:

Definition: Let G be a neighborhood system defined over a lattice S , we define a "clique" c as either a single site, or a set of sites of the lattice, such that all the sites that belong to c are the neighborhood of each other. Fig. 1 shows the cliques associated with G^2 neighborhood system. The labels in Fig. 1 correspond to the respective clique potentials.

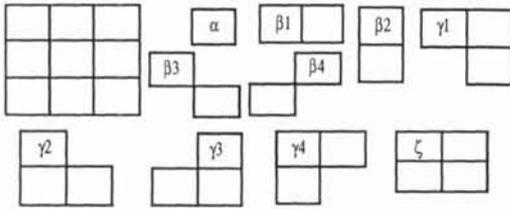


Fig. 1. G^2 and its associated clique types.

Definition: A random field F defined on S has Gibbs Distribution (GD) or equivalently is a Gibbs Random Field (GRF) with respect to G if and only if its joint distribution is of the form

$$P(F=f) = \frac{1}{Z_0} e^{-\frac{1}{T} U_0(f)}$$

where Z_0 is a partition function, $U_0(f)$ is the energy function, $V_c(f)$ is the potential associated with clique c , and T is a parameter that corresponds to temperature in a physical system.

Hammersley-Clifford Theorem:- Let G be a neighborhood system on a finite lattice S . A random field F is a MRF with respect to G if and only if its joint distribution is a GD with cliques associated with G .

The MRF concept is used to model the apriori knowledge about the spatial interaction among the image pixels within neighborhoods that are small enough for practical purposes. Modeling the apriori knowledge using MRF-GRF equivalence requires specifying the clique potentials associated with the neighborhood system. The MRF approach allows us to integrate a number of features by defining appropriate clique potentials for each image feature. In the work presented here a second order neighborhood system is chosen.

For refining the jump edge map a coupled depth and line model is used. The cliques potentials for the line process (l) of the coupled model are as follows: $\alpha=10$, $\beta=-2$, $\gamma=6$, $\zeta=8$. The same values are used for all the β 's and γ 's and all other configurations have zero potential value. In the coupled model presented here the dual line process lattice is assumed to coincide with the original image. The choice of parameters effectively discourages both the formation of thick edges and the presence of sharp turns.

For the depth process we consider the variation in the depth values in each clique, and we model the potentials of the continuous depth process as follows:

$$V_c(f, l) = | \text{Average of pixels depth values in the clique for which the line process indicates no edge} |$$

Encoding the apriori knowledge for fold edges follows the same approach as that of jump edges. The depth model is not useful for fold edge detection since a fold edge does not introduce large depth variation. We use a coupled line, normal and AR model to encode the apriori knowledge for fold edge processing. A second order neighborhood system is used for the coupled model.

The line process for fold edge processing is similar to that of jump edges. The potential assignments for this process are: $\alpha = 15$, $\beta = -2$, $\gamma = 10$, $\zeta = 20$. The same values are used for all the β 's and γ 's and all other configurations have zero potential value. The second model used for fold edge detection is that of the surface normals. The clique potentials for the normal process are defined as follows:

$$V_{C_S}(f, l) = | \text{Average of the angles between the surface normals in the clique for which the line process indicates no edge} |$$

Fold edges are located by detecting a maxima in AR in a direction perpendicular to the edge direction. For the AR model, we chose the non-zero clique potentials as follows:

$\beta = -10$ If both points in the clique are edge points and AR is maximum in a direction perpendicular to the edge.

3.2. MODELING THE OBSERVATION

For the purpose of image segmentation, the image S can be modeled as the union of M regions R_j . Let us assume that the observations g correspond to samples of the surface f taken at a set of sites S . We assume that the observations are corrupted by a zero mean white additive Gaussian noise process with standard deviation of σ_n . This leads to the definition of the conditional distribution¹

$$P(G=g | F=f) = \prod_{j=1}^M \prod_{i \in R_j} \frac{1}{\sqrt{2\pi} \sigma_n} \exp - \frac{(f_i^{R_j} - g_i)^2}{2\sigma_n^2}$$

3.3. MAXIMUM APOSTERIOR ESTIMATOR

The Bayesian approach to the solution of reconstruction problems has been adopted by several researchers^{4,5}. In most cases, the criterion for selecting the optimal estimate has been the MAP estimator. Estimation for the scene that maximize the a posteriori distribution is derived as follows: The posterior distribution is given by:

$$P(F=f | G=g) = \frac{P(F=f) P(G=g | F=f)}{P(G=g)}$$

The MAP estimator can be shown¹ to be equivalent to the minimization of:

$$U_p(f, l | g) = \sum_C V_C(f, l) + \sum_{C_l} V_{C_l}(l) + \frac{1}{2\sigma^2} \sum_{j \in 1}^M \sum_{i \in R_j} (f_i^{R_j} - g_i)^2 \quad (1)$$

for a jump edge and

$$U_p(f, r, n, l | g) = \sum_{C_r} V_{C_r} + \sum_{C_n} V_{C_n} + \sum_{C_l} V_{C_l} + \frac{1}{2\sigma^2} \sum_{j \in 1}^M \sum_{i \in R_j} (f_i^{R_j} - g_i)^2 \quad (2)$$

for a fold edge.

4. ALGORITHM DESCRIPTION

The formulation presented in the previous sections is now used in our integrated segmentation approach. The application of edge detection techniques to segment images can lead to three basic types of errors: 1) The false detection of edges (T₁). 2) The false rejection of edges (T₂). 3) Localization error (T₃). T₁ errors generally result in oversegmenting the image. The oversegmentation can be recovered by merging adjacent regions. Errors of T₂ are more severe in range images due to the presence of fold edges and result in undersegmenting the image. T₃ errors are also difficult to correct in range images because depth variation, close to a fold edge, is small.

Fig. 2 shows the steps in the algorithm. The range image is first smoothed using a median filter then jump edges are detected. The applications of the jump edge detection technique presented in section 2 will result, in general, in errors of T₁. Using a 3x3 window, jump edges which are, at the most, three pixels thick will be formed. The process might introduce errors of T₂ also, but these errors can be corrected in the fold edge detection step. Due to the large depth variation for jump edges, T₃ errors are negligible.

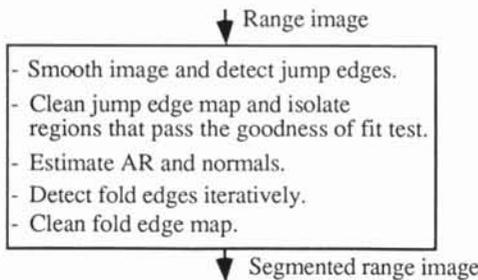


Fig. 2. Steps in the algorithm.

Jump edge map is then processed using the jump edge relaxation algorithm. The relaxation algorithm minimizes Equation 1 without the observation part. The connected components in the resulting jump edge map is found using a 4-connected component algorithm. A fourth order polynomial is then fitted to the resulting connected components and the regions are verified using a goodness of fit test. The goodness of fit test we employed compares the residual of the fit to that of the noise. Regions that pass the goodness of fit test are removed from further processing. In most cases, this step successfully isolates the background region and the regions surrounded by jump edges only. This reduces the computation needed in the following steps.

The AR values and surface normals are estimated for all image pixels. Both AR and normal require thresholds to detect fold edges. Instead of using thresholds that might work under some conditions and fail in others we have used an iterative process to detect all possible fold edges in the image. The iterative process starts with a high threshold value for both features. The threshold values are then reduced until all possible fold edges are detected. The detection of fold edges will result, in general, in T₁ and T₃ errors. It is essential in our algorithm that all errors of T₂ be eliminated. This is achieved by verifying regions obtained using the iterative process. The verification ensures that surface patches obtained belong to only one object region. This is achieved through the use of the goodness of fit test. If a particular region is not verified by the goodness of fit test, more fold edges are detected by lowering the threshold values for both AR and normals. The process is repeated until all regions are verified.

The resulting edge map is then processed using the fold edge relaxation algorithm. The goal of this step is to eliminate all T₁ errors and to reduce the localization errors (T₃ errors). Localizing fold edges are difficult because they do not correspond to a large depth variation. The use of normal and AR (Equation 2) attempt to reduce the localization errors.

The relaxation algorithm minimizes Equation 2 to obtain the final segmentation. The strategy we have followed is divided into two steps as follows: 1) In step one Equation 2 is minimized until no further changes are possible. In most cases, the output after this step has a good localized edges and all possible merges between adjacent surfaces are performed. 2) In step two, the image is processed using only the observation part of Equation 2. This in effect will reduce the MAP estimate to region growing based on fit error criteria. In our experiments surface normals have the most effect on the segmentation output.

5. EXPERIMENTAL RESULTS

The algorithm presented above has been applied successfully to a wide variety of range images. The algorithms performance on 3 range images is discussed in this section.

Two joint cylinders range image results: Fig. 3 shows a noisy range image of a part consisting of two cylinders of different diameters. As can be seen the algorithm segmented the image correctly. Some points in the background region are left unclassified. This is because the background region has a slope variation at the upper left and lower right hand corners of the image.

Cube range image results: The segmentation results of a cube range image with holes drilled through them is shown in Fig. 4. This image provides an example of range image with a combination of flat and cylindrical

surfaces. Gaussian noise of mean zero and standard deviation of 2.0 were added to the image. The algorithm segmented the image correctly with the exception of two small background components which are left unclassified. This is because these regions are small compared to the minimum size required by the fitting algorithm.

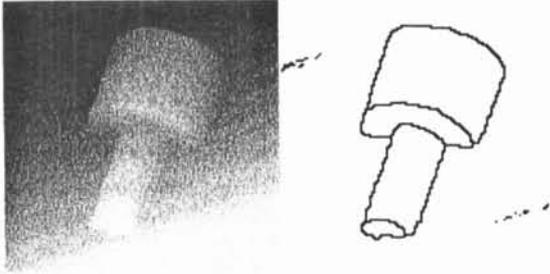


Fig. 3. Segmentation results on a two cylinders image.

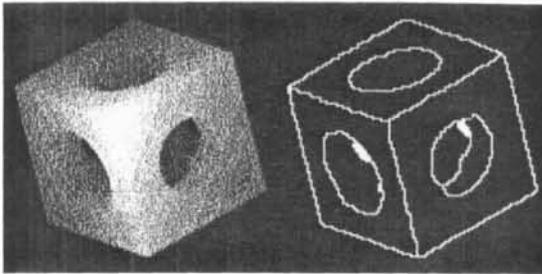


Fig. 4. Segmentation results on a cube image.

Industrial part range image results: Fig. 5 shows a noisy industrial part range image with Gaussian noise of standard deviation of 2.0. The final edge map shows that the algorithm segmented the image correctly. The iterative region identification algorithm resulted in oversegmenting the region inside the circular part of the object into two regions. The fold edge processing procedure correctly merged these two regions to obtain the final results.

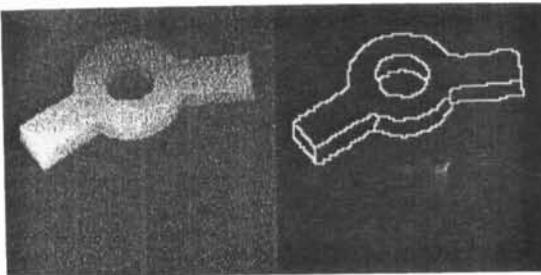


Fig. 5. Segmentation results on an industrial part image.

To assess the algorithm computation complexity we have measured the CPU time for three time consuming stages of the algorithm. Table 1 shows the CPU time (in seconds on a VAX 8800) at the three stages of the algorithm for the objects shown in Fig. 3 (T1) and the object in Fig. 5 (T2). For the planar fit, considerable time saving can be achieved by using convolution window to determine the plane parameters using parallel hardware. The iterative region identification stage in the algorithm is also time consuming for some of the images tested. It can be seen from the table, that the computation will be reduced substantially if the object contains a number of regions surrounded by jump edges. This illustrates the advantage of the two stage process (processing jump edges then fold edges) used in the algorithm. Fold edge processing is another stage where a lot of computation is needed. If parallel hardware

is available, considerable saving can be achieved since minimizing Equation 2 is only done in a small area around the processed pixel. Thus the algorithm can be implemented using parallel architecture if the neighboring pixels are not updated simultaneously.

	T1	T2
Planar fit	47.11	48.51
Iterative region identification	58.16	22.13
Fold edge processing	99.06	64.94

Table 1 The CPU time (seconds) for two objects at three stages of the algorithm. T1 is for the object in Fig. 3, and T2 is for the object in Fig. 5.

6. CONCLUSIONS

Due to the difficulties associated with edge based and region based segmentation methods, we have explored the idea of integrating these two approaches for range data segmentation. In addition to integrating edge and region based techniques, we have used a number of image cues to detect discontinuities in range images. The experiments show that our algorithm produces good results for the images tested.

The algorithm is attractive from the computational and robustness point of view. We have observed that the use of the sequential approach to detect edges is computationally attractive since a large number of regions is isolated early in the segmentation process. A number of experiments shows that the use of the Bayesian approach for the region growing part of the algorithm produces better results than that of a region growing algorithm based on the fit error.

ACKNOWLEDGMENTS

This work was supported in part by the office of Naval Research under contract N00014-89-K-0186.

REFERENCES

1. E. Al-Hujazi and A. K. Sood, "Integrating Edge detection and Region growing for Range Image Segmentation," in *Proc. SPIE Conf. 1381 Intell. Robots and Compu. Vision IX*, Boston, MA, Nov. 1989.
2. E. Al-Hujazi and A. K. Sood, "Range Image Segmentation with Applications to Robot Bin-Picking Using Vacuum Gripper," *IEEE Sys. man cyberne.*, SMC-20, No. 6, Nov./Dec. 1990.
3. P. J. Besl, and R. C. Jain, "Segmentation Through Variable-Order Surface Fitting," *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-10, No.2, Mar. 1988, pp. 167-192.
4. H. Derin, and H. Elliott, "Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields," *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-9, No.1, Jan. 1987, pp. 39-55.
5. S. Geman, and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-6, No.6, Nov. 1984, pp. 721-741.
6. G. G. Pieroni, "A Multiresolution Approach for Segmenting Surfaces," in *Issues on Machine Vision*, G. G. Pieroni (Editor), Springer-Verlag, 1989.