IAPR Workshop on CV - Special Hardware and Industrial Applications OCT.12-14, 1988, Tokyo

Stable Estimation of a Topographic Primal Sketch for Range Image Interpretation

Pierre Boulanger
National Research Council of Canada
Laboratory for Intelligent Systems
Ottawa, Canada, K1A-0R8

Paul Cohen
Ecole Polytechnique de Montreal
Departement of Electrical Engineering
Montreal,Canada

Abstract

This paper describes how range images can be segmented using the sign of differential geometry operators such as the mean and Gaussian curvatures. This segmentation called topographic primal sketch is invariant to rigid body motion in space and is defined by eight fundamental surfaces. The first part of the paper presents two novel techniques where an initial estimate of the categories, full of inconsistent labelling due to noise, is transformed into a consistent one. One of the two methods is based on a label relaxation technique, where consistency is viewed as a local optimization problem, and the second is based on stochastic relaxation, where the local classification of a pixel is solved by statistical decisions. The advantages and disadvantages of each method are presented.

Introduction

The basis of the topographic primal sketch consists of segmenting range images into surface patches according to categories defined by differential geometry operators such as the Gaussian (K) and mean curvatures (H) (see [Haralick 83], [Besl 86a] and [Besl 86b]). From the sign of these invariant functions of directional derivatives, one can generate categories such as peak, pit, ridge, ravine, saddle, flat and hillside (see Table 1). From this initial classification, these categories can be grouped to obtain a rich, hierarchical, and structurally complete representation of the fundamental range image structure. This paper presents two novel techniques where an initial estimate of the categories full of inconsistent labelling due to noise is transformed into a consistent one. The first technique is based on a label relaxation algorithm where consistency is defined by geometrical considerations such as curvature continuity. The technique optimizes the global consistency by successively maximizing the local consistency in an immediate neighborhood. The second technique uses stochastic relaxation where the probability for a label is maximized using an optimization procedure. This paper describes the advantages and disadvantages of each method and also compares the quality of the segmentation produced.

Using relaxation labelling we will demonstrate on real range images produced by the National Research Council of Canada (NRCC) range finder [Rioux 84] that a stable estimation of a topographic primal sketch is possible.

	K		
H	+	0	-
-	7 Peak	4 Ridge	1 Saddle Ridge
0	8 (none)	5 Flat	2 Minimal Surface
+	9 Pit	6 Valley	3 Saddle Valley

Table 1. : Table of surface shapes and labels from Gaussian (K) and mean (H) curvature signs.

Numerical Estimation of K and H

A range image is a graph Z(x,y) of three-dimensional measurements at a fixed view point of a scene. In order to evaluate the Gaussian K(x,y) and mean H(x,y) at every point (x,y) one must compute the following equations:

$$K = \frac{(Z_{xx}Z_{yy} - Z_{xy}^2)}{(1 + Z_x^2 + Z_y^2)^2} \quad (1)$$

$$2H = \frac{Z_{xx}(Z_y^2 + 1) + Z_{yy}(Z_x^2 + 1) - 2Z_xZ_yZ_{xy}}{(1 + Z_x^2 + Z_y^2)^{\frac{3}{2}}} \quad (2)$$

One technique to compute these functions is to evaluate each derivative by a local quadric model expressed by:

$$Z(x_0, y_0) = a_0 + a_1(x - x_0) + a_2(y - y_0) + a_3(x - x_0)^2 + a_4(x - x_0)(y - y_0) + a_5(y - y_0)^2 \quad (3)$$

where the first and second order derivatives are proportional to the coefficients a_1, a_2, a_3, a_4 and a_5, that is:

$$a_1 = \left(\frac{\partial z}{\partial x}\right)_0 \ ; \ a_2 = \left(\frac{\partial z}{\partial y}\right)_0 \quad (4)$$

$$2a_3 = \left(\frac{\partial^2 z}{\partial x^2}\right)_0 \ ; \ 2a_5 = \left(\frac{\partial^2 z}{\partial y^2}\right)_0 \quad (5)$$

$$\left(\frac{\partial^2 z}{\partial x \partial y}\right)_0 = a_4 \quad (6)$$

In order to fit this local surface model to the actual range data one must use a mean-square technique that minimizes a L_2 norm inside a window centered at (x_0, y_0). This minimization process is similar to the solution of an over-determined system expressed by :

$$A = (C^T C)^{-1} C^T Z \quad (7)$$

where A$=(a_0, a_1, ..., a_5)$ are the coefficients, Z$=(Z_{11}, Z_{12}, ..., Z_{mm})$ are the Z values inside the window and C is the coordinate matrix.

The window size is critical for a good estimation of the curvatures. A larger window size produces a better signal-to-noise ratio for the estimation of K and H because we are computing an average in a larger neighborhood, but a larger window size will also reduce the locality of the K and H measurements resulting in a loss of small sized structures and produce erroneous estimation of the curvatures near discontinuities. A window size of 7 × 7 was used in most of our computations.

Production of the Topographic Primal Sketch

The topographic primal sketch is produced from the initial evaluation of the Gaussian and mean curvatures. It corresponds to a label image where each type of surface is coded between 1 and 9. One of the problems related to the production of this label map is the evaluation of two threshold values ϵ_k, ϵ_h corresponding to the zero values of K and H. These values are very critical because they correspond to an unstable region of the possible values of K and H. A technique for evaluating these thresholds is to measure with a 3D sensor a scene of a flat surface and then evaluate the K and H values with the same operator used to analyse the scene. After the evaluation of K and H for this surface, we produce a two-dimensional histogram where we can evaluate the distribution of the noise for K=0 and H=0. Using this distribution we can compute an optimal threshold based on a maximum likelihood separation between two classes. The threshold calculated by this technique can then be used for the classification of surfaces in a scene acquired by a laser scanner in the same configuration. One can see in Figure 1 a two-dimensional histogram of the K and H for a flat surface.

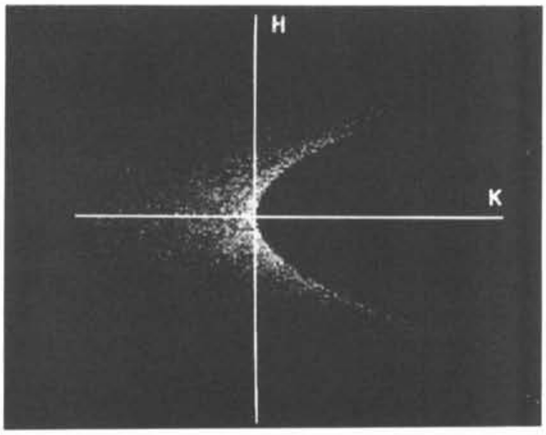

Figure 1: Bi-dimensional histogram of K and H. Horizontal axis is K and vertical axis is H.

Label Relaxation of the Topographic Primal Sketch

After the initial estimation of the topographic primal sketch a label relaxation process is applied to improve its consistency. Typically, noise in the range image produces labels that may be inconsistent with the region surrounding them.

Basically, relaxation labelling is an iterative procedure applied over a network of nodes. Associated with each node is a set of labels (in our problem numbers between (1) and (9)), and associated with each label is a measure of confidence or certainty. The degree of compatibility between a label and its neighborhood can be measured by what is known as the label's support. Relaxation labelling is the process of achieving ambiguity reduction by iteratively optimizing the local consistency. In the next sections we will described two new techniques to perform this operation.

Non Statistical Relaxation Labelling

First introduced by Rosenfeld et al. [1976], this relaxation labelling technique define compatibility between labels as a set of rules corresponding to assumption about the behavior of the world. Using his notation the degree of compatibility between a label and its neighborhood can be measured by what is known as the label's support function:

$$S_i(\lambda) = \sum_{\lambda'} R_{ij}(\lambda, \lambda') p(\lambda'), \quad (8)$$

which is a function of other label certainties $p(\lambda')$ in the neighborhood and their compatibility $R_{ij}(\lambda, \lambda')$ (pairwise) with the label being supported. The constraints between labels are represented by a matrix of compatibilities $R_{ij}(\lambda, \lambda')$ which corresponds, for the topographic primal sketch, to a continuity criterion on the curvature signs of K and H. In this notation $R_{ij}(\lambda, \lambda')$ denotes the compatibility between label λ' associated to node j and label λ associated to node i.

Definition of the Consistency Matrix $R_{ij}(\lambda, \lambda')$ and Label Support Function $p_i(\lambda)$

We will define consistency in our problem as a continuous variation of the sign of K and H, that is, K or H must first pass zero when they vary from (-) to (+) or from (+) to (-). One can see in Table 2 the values of the consistency matrix from one label to another. The maximum consistency corresponds to similar pairs of labels and is reduced to half for labels corresponding to transitions between (+) or (-) to zero. Transitions between (-) to (+) or (+) to (-) are considered to be totally inconsistent.

The label support is defined as follow

$$p_i(\lambda) = p_{\epsilon_k}(\lambda) p_e(\lambda) p_{\epsilon_h}(\lambda) p_h(\lambda) p_k(\lambda) \qquad (9)$$

where $p_{\epsilon_k}(\lambda)$ and $p_{\epsilon_h}(\lambda)$ correspond to a normalized distance from the threshold values ϵ_k and ϵ_h. They are expressed by the following equations:

$$p_{\epsilon_n}(\lambda) = 1 - e^{\frac{-(|n|-\epsilon_n)^2}{\sigma_{t_1}^2}} \qquad N - \epsilon_n \leq 0 \qquad (10)$$

$$p_{\epsilon_n}(\lambda) = 1 - e^{\frac{-(|n|-\epsilon_n)^2}{\sigma_{t_2}^2}} \qquad N - \epsilon_n > 0 \qquad (11)$$

where n is equal to k or h and σ_{t_1}, σ_{t_2} correspond to the degree of confidence on the thresholds. The value of these functions $\cong 1$ for K=0 and/or H=0 and zero for $K = \epsilon_k$ and/or $H = \epsilon_h$.

	1	2	3	4	5	6	7	8	9
1	1.0	0.5	-1.0	0.5	0.5	-1.0	-1.0	-1.0	-1.0
2	0.5	1.0	0.5	0.5	0.5	0.5	-1.0	-1.0	-1.0
3	-1.0	0.5	1.0	-1.0	0.5	0.5	-1.0	-1.0	-1.0
4	0.5	0.5	-1.0	1.0	0.5	-1.0	0.5	-1.0	-1.0
5	0.5	0.5	0.5	0.5	1.0	0.5	0.5	-1.0	0.5
6	-1.0	0.5	0.5	-1.0	0.5	1.0	-1.0	-1.0	0.5
7	-1.0	-1.0	-1.0	0.5	0.5	-1.0	1.0	-1.0	-1.0
8	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0	-1.0
9	-1.0	-1.0	-1.0	-1.0	0.5	0.5	-1.0	-1.0	1.0

Table 2. : Consistency matrix $R_{ij}(\lambda, \lambda')$.

The $p_e(\lambda)$ function corresponds to the degree of confidence on the evaluation of the curvature values. The function is based on the error estimate of the local quadric model used to evaluate the curvatures. If the local quadric model is a good fit then the confidence level is high (max. 1) if not the confidence level is reduced in accordance with the following equation:

$$p_e(\lambda) = e^{-\mu^2/\sigma_e^2} \qquad (12)$$

where

$$\mu^2 = \sum_{window} (f(x,y) - z(x,y))^2 \qquad (13)$$

which corresponds to the mean square error inside the evaluation window. Experimental results show that $\sigma_e \simeq 4\,\epsilon_z$ is a good approximation of this parameter.

The last label support functions $p_k(\lambda)$ and $p_h(\lambda)$ compute the degree of similarity of the curvature values inside an immediate neighborhood (typically a 5 by 5 window). They are expressed by

$$p_n = e^{-(n-\bar{n})^2/\sigma_n^2} \qquad (14)$$

where

$$\sigma_n = \sum_{window} (n - \bar{n})^2 \qquad (15)$$

which corresponds to the standard deviation inside the window.

Label Relaxation Process

Our relaxation process is similar to the one developed by Hummel and Zucker [1983] where the consistency optimization problem is defined in variational terms.

In our problem a window of size 3 × 3 centered at each node i, j is analysed so that the consistency functional expressed by

$$S_i(\lambda) = \sum_{\lambda'} R_{ij}(\lambda, \lambda') p_j(\lambda') \qquad (16)$$

is optimized. That is, in a neighborhood of 3 × 3 find the best label at the center of this window that optimize the label support.

The relaxation process is (a) to increase the label support if other unit labels that have high weighted labels are compatible with λ at the center of the window and (b) decrease the label supports if other highly weighted labels are incompatible with λ. At each iteration the label support is updated by:

$$p_i^{n-1}(\lambda) = \frac{p_i^n(\lambda)[1 + S_i^n(\lambda)]}{\sum_\lambda p_i^n(\lambda)[1 + S_i^n(\lambda)]} \qquad (17)$$

A vectorized version of this algorithm was implemented on an array processor. One can vectorize the problem by computing in one operation all the label support for an incremental sequence of labels and in another operation compute the sum of the label support for each one of them. Finally, a third operation searches the label with the maximum support.

One can see in Figure 2a the initial estimate of the topographic primal sketch of a range image composed of a sphere, a cylinder and polyhedrons. Figure 2b, 2c, and 2d illustrate the evolution of the primal sketch as a function of the relaxation labelling process.

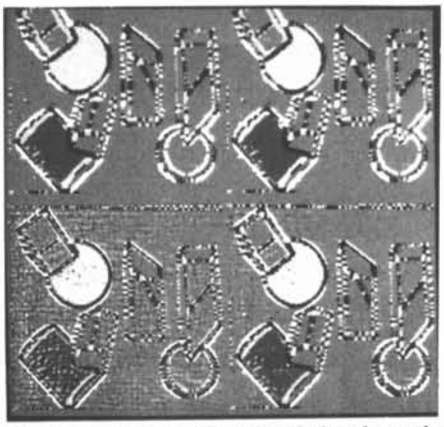

Figure 2: Relaxation of the primal sketch vs the number of iterations (non statistical method). (a) lower left (0) iteration, (b) lower right (1) iteration, (c) upper left (5) iterations, (d) upper right (10) iterations.

Typically, convergence is obtained after four or five iterations. On a 7 Mflops array processor the computation speed is 10 sec per iteration. An increase of the convergence rate is possible if we would use a larger window for the label optimization, but the consistency matrix would be more complicated since we are not optimizing with immediate neighbors.

Statistical Relaxation Labeling

In this section we will describe a new relaxation labelling method based on a generalized version of the algorithm described by Therrien et al. [1986].

This algorithm uses a two dimensional stochastic linear model to improve the local consistency of the primal sketch. In this algorithm, local consistency is achieved by maximizing the support for a label using a simple majority rule. Unlike the previous algorithm, where the update equation for the label support is linear, this algorithm use a non linear statistical decision making process . The process is also sequential, meaning that the decision of the classification of a pixel is affected by decisions on the previous one.

Stochastic Relaxation Process

Let E be a set of labels and F be the set of all possible pairs of labels defined as:

$$E = \{\lambda_1, \lambda_2, ..., \lambda_L\} \quad (18)$$

$$F = \{(\lambda_m, \lambda_n)|\ \lambda_m \in E,\ \lambda_n \in E,\ \lambda_m \neq \lambda_n\} \quad (19)$$

Given a pair of labels (λ_m, λ_n), one can define the support for a label as a conditional probability problem.

Let

$$p(\lambda_m|V) = \frac{1}{D}\exp\{\alpha C_{0,0} + \beta(C_{1,1} + C_{-1,-1} + C_{-1,1} + C_{1,-1}) + \gamma(C_{0,1} + C_{1,0} + C_{-1,0} + C_{0,-1})\} \quad (20)$$

be the probability that the central pixel has label λ_m as a function of its neighborhood V and

$$p(\lambda_n|V) = \frac{1}{D}\exp\{\alpha C_{0,0} + \beta(C_{1,1} + C_{-1,-1} + C_{-1,1} + C_{1,-1}) + \gamma(C_{0,1} + C_{1,0} + C_{-1,0} + C_{0,-1})\} \quad (21)$$

be the probability that the central pixel has label λ_n.

In these equations $C_{x,y}$ represents the support between the central pixel λ and its neighbor. For example, if the two possible labellings of the central pixel are λ_m, λ_n, then we can compute the support for one particular interpretation as follow: $C_{x,y} = 1$ if the neighboring label $\lambda' = \lambda_n$, $C_{x,y} = -1$ if $\lambda' = \lambda_m$ and $C_{x,y} = 0$ if $\lambda' \neq \lambda_m$ and $\lambda' \neq \lambda_n$.

After the computation of the support functions $p(\lambda_m|V)$ and $p(\lambda_n|V)$ we can give credit to the hypothesis that the central pixel is λ_m if $p(\lambda_m|V) > p(\lambda_n|V)$.

After the comparison pair-wise of all the possible labelling we then compute the following credit function $T(\lambda_i)$:

$$T(\lambda_i) = \frac{Number\ of\ Credits\ Given\ to\ Label\ i}{Cardinal\ of\ the\ set\ F} \quad (22)$$

We then assign to the central pixel the label corresponding to the one with the maximum credit function.

The algorithm computes for one iteration all the most probable labels and then uses this result recursively until no labels are changed.

One can see in Figure 3 the evolution of the label map as a function of the number of iterations. One of the most interesting properties of the algorithm is the fast convergence rate compared to the non statistical algorithm (About two times the speed). For this technique the choice of the appropriate window size is crucial for the preservation of small regions. The best choice in our experimental results is 3 × 3 for the window size and $\alpha = \beta = \gamma = 1$ for the label support functions parameters.

Figure 3: Relaxation of the primal sketch vs the number of iterations (statistical method). (a) lower left (0) iteration, (b) lower right (1) iteration, (c) upper left (5) iterations, (d) upper right (10) iterations.

Conclusion

We have demonstrated that the label relaxation process can improve significantly the quality of the topographic primal sketch. The simple rule of curvature sign continuity has produced good results but improvements such as a larger window size and some rules on long range curvature consistency may increase the convergence rate and the quality of the sketch.

We have also demonstrated that stochastic relaxation using majority rules can also produce good results. One of the significant advantages of this method is its fast convergence rate over the non-statistical one. But, since this algorithm is serial, it would be impossible to improve its speed on future parallel machines.

These algorithms, based on different philosophies of computing the label consistency problem produced similar results. But the non-statistical one has rules such as curvature continuity criterion and label support function based on the actual values of the curvatures which make it less sensitive to noise and more importantly less sensitive to the initial labelling of the primal sketch.

Further study of both algorithms is necessary to determine which one is the most efficient to produce a good consistent topographic primal sketch.

Acknowledgements

The authors wish to express theirs appreciation for the image acquisition to Marc Rioux and Luc Cournoyer from the Laboratory of Intelligent Systems . Their assistance was invaluable. We would also like to tank M. Hai Nguyen Hong for his excellent contribution in the software development.

References

[Besl 86a] Besl. P.J. Surfaces in Early Range Image Understanding, Phd Thesis. University of Mariland, 1986

[Besl 86b] Besl, P.J. and Jain, R.C. Invariant Surface Characteristics for Three-dimensional Object Recognition in Range Images, Computer Vision, Graphics, Image Processing 33,(1) (January), 33-80, 1986.

[Haralick 83] Haralick,R.M., Watson,L.T., and Laffey,T.J. The Topographic Primal Sketch,Int. J. Robotics Res. 2, (1), (Spring) 50-72, 1983.

[Hummel 83] Hummel,R.A., Zucker,S.W., On the Foundation of Relaxation Labeling Processes, PAMI-5, No 3, May, 1983.

[Rioux 84] Rioux,M. Laser Range Finder Based on Synchronized Scanners , Appl. Opt. 23, 3837, 1984.

[Rosenfeld et al 76] Rosenfeld, A., Hummel, R. and Zucker, S.W. Scene Labeling by Relaxation Operations, IEEE Transactions on System Man and Cybernetics, Vol. 6, pp. 420-433, 1976.

[Therrien et al. 86] Therrien, C.W., Quatieri, T.F., Dudgeon, D.E. Satistical Model-Based Algorithms for Image Analysis, Proceedings of the IEEE, Vol. 74, No. 4, April 1986.