LABEL RELAXATION TECHNIQUE APPLIED TO THE STABLE ESTIMATION OF A TOPOGRAPHIC PRIMAL SKETCH

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

ABSTRACT

The basis of the topographic primal sketch consist of segmenting range images into surface patches according to categories defined by differential geometry operators such as the Gaussian and mean curvatures. From the sign of these invariant functions of directional derivatives, one can generate categories such as peak, pit, ridge, ravine, saddle, flat and hillside. From this initial classification, we can group these categories to obtain a rich, hierarchical, and structurally complete representation of the fundamental range image structure. In the paper, we present a novel technique where an initial estimate of the categories full of inconsistent labelling du to noise is transformed into a consistent one by label relaxation technique. We also discuss the problem of numerical stability of the Gaussian and mean curvatures and study the effects of different operators on these estimates.

KEYWORDS: Vision, Label Relaxation, Topographic primal sketch, Range image, Segmentation

Introduction

Segmentation of range images is one of the most important step in a three-dimensional object recognition system. Unlike grey scale images, segmentation of range images has a direct relationship to the properties of object surfaces. Using differential geometry Haralick [Hara 83] and more recently Besl [Besl1 86] [Besl2 86] were able to describe local properties of a surface using invariant parameters such as the Gaussian (K) and mean (H) curvatures. Using the sign of the Gaussian and mean curvatures one can segment a range image into regions corresponding to one of the eight fundamental surfaces such as peak, pit, ridge, flat, valley, saddle ridge, minimal surface and valley (see Table 1). This segmentation produce a labeled image called by Haralick a topographic primal sketch. One of the problems related to the computation of this primal sketch is the production of a stable and consistent estimation of each region as a function of noise and shadow effects inherent in any 3D sensor [Rioux 84]. In this paper we will first describe how to compute an initial estimate of the primal sketch using a mean squared evaluation technique. We will also describe how to evaluate the two thresholds for which K and H are considered to be zero. After the initial estimation of the topographic primal sketch, we will demonstrate how it is possible to further improve the sketch by using certain label consistency rules. Using relaxation labelling we will demonstrate on real range images that a stable estimation of a topographic primal is possible.

H \ K	+	0	-
-	7 Peak	4 Ridge	1 Saddle Ridge
0	8 (none)	7 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.

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

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

One technique used to compute these functions is to evaluate each derivative by finite difference equations expressed by

$$Z_x = (Z(x_0+h, y_0) - Z(x_0-h, y_0))/2h \tag{3}$$

$$Z_y = (Z(x_0, y_0+h) - Z(x_0, y_0-h))/2h \tag{4}$$

$$Z_{xx} = (Z(x_0+h, y_0) + Z(x_0-h, y_0) - 2Z(x_0, y_0))/h^2 \tag{5}$$

$$Z_{yy} = (Z(x_0, y_0+h) + Z(x_0, y_0-h) - 2Z(x_0, y_0))/h^2 \tag{6}$$

$$Z_{xy} = \begin{aligned}&(Z(x_0+h, y_0+h) - Z(x_0+h, y_0-h) \\ &-Z(x_0-h, y_0+h) + Z(x_0-h, y_0-h))/4h^2\end{aligned} \tag{7}$$

where h is equal to the distance between two samplings in the x or y direction.

One of the major difficulties with these finite difference equations is their high sensitivity to noise. If ϵ_z is the standard deviation of the noise on the range image, one can calculate a pessimistic estimate of the noise on K(x,y) and H(x,y) by:

$$\eta_k = \frac{16\epsilon_z^2}{h^4\sqrt{1 + 2\frac{\epsilon_z^2}{h^2}}} \tag{8}$$

$$\eta_h = \frac{-4\epsilon_z[2h^2 + 4\epsilon_z^2]}{h^4(1 + 2\frac{\epsilon_z^2}{h^2})^{\frac{3}{2}}} \tag{9}$$

One can see from these equations that a small increase in the noise level on the range data will produce a significant increase in the noise level on K and H. For example a typical noise level of 0.1 mm on the range map of a flat surface sampled at every 1 mm produce a noise level of $\eta_k = 0.16$ mm^{-2} for K(x,y) and of $\eta_h = 0.78$ mm^{-1} for H(x,y). The evaluation of K and H by these simple equations is highly unstable numerically since a small fluctuation in the sampling rate produces a large variation in η_k and η_h.

In order to solve this problem one must regularize it [Tiko 77] using a local surface model from which the first and second order derivatives can be computed. The technique consists of evaluating a local quadric model expressed by

$$\begin{aligned} Z(x_0, y_0) = \; & a_0 + a_1(x - x_0) + a_2(y - y_0) + a_3(x - x_0)^2 + \\ & a_4(x - x_0)(y - y_0) + a_5(y - y_0)^2 \end{aligned} \tag{10}$$

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

$$(\frac{\partial z}{\partial x})_0 = a_1 \tag{11}$$

$$(\frac{\partial z}{\partial y})_0 = a_2 \tag{12}$$

$$(\frac{\partial^2 z}{\partial x^2})_0 = 2a_3 \tag{13}$$

$$(\frac{\partial^2 z}{\partial x \partial y})_0 = a_4 \tag{14}$$

$$(\frac{\partial^2 z}{\partial y^2})_0 = 2a_5 \tag{15}$$

In order to fit this local surface model to the actual range data one must use a mean-square technique that minimizes a L2 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 :

$$A = (C^TC)^{-1}C^TZ \tag{16}$$

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. Using this particular norm, the relationship between the noise level of the range image and the mean and Gaussian curvatures is express by :

$$\eta_k = \frac{\delta a_3 \delta a_5 - (\delta a_2)^2}{\sqrt{1 + (\delta a_1)^2 + (\delta a_2)^2}} \tag{17}$$

$$\eta_h = \frac{\delta a_3(1 + (\delta a_2)^2) + \delta a_5(1 + (\delta a_1)^2) - 2\delta a_1 \delta a_2 \delta a_4}{(1 + (\delta a_1)^2 + (\delta a_2)^2)^{\frac{3}{2}}} \tag{18}$$

where

$$\delta a_1 = \frac{\sum_{window} x_n \epsilon_{zn}}{\sum_{window} x_n} \tag{19}$$

$$\delta a_2 = \frac{\sum_{window} y_n \epsilon_{zn}}{\sum_{window} y_n} \tag{20}$$

$$\delta a_3 = \frac{\sigma_x^2 \nu_y^4 - \nu_{xy}^2 \sigma_y^2}{\nu_y^4 \nu_x^4 - \nu_{xy}^4} \tag{21}$$

$$\delta a_4 = \frac{\sum_{window} x_n y_n \epsilon_{zn}}{\sum_{window} x_n^2 y_n^2} \tag{22}$$

$$\delta a_5 = \frac{\sigma_y^2 \nu_x^4 - \nu_{xy}^2 \sigma_x^2}{\nu_y^4 \nu_x^4 - \nu_{xy}^4} \tag{23}$$

where

$$\nu_x^4 = \sum_{window} n x_n^4 - (\sum_{window} x_n^2)^2 \tag{24}$$

$$\nu_y^4 = \sum_{window} n y_n^4 - (\sum_{window} y_n^2)^2 \tag{25}$$

$$\nu_{xy}^2 = \sum_{window} n x_n^2 y_n^2 - \sum_{window} x_n^2 \sum_{window} y_n^2 \tag{26}$$

$$\sigma_x^2 = \sum_{window} n x_n^2 \epsilon_{zn} - (\sum_{window} x_n^2)(\sum_{window} \epsilon_{zn}) \tag{27}$$

$$\sigma_x^2 = \sum_{window} n x_n^2 \epsilon_{zn} - (\sum_{window} x_n^2)(\sum_{window} \epsilon_{zn}). \tag{28}$$

As one can see from these equations, 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 neighbourhood, but a larger window size will also reduce the locality of the K and H measurements resulting in a loss of small sized structures. One can see in Figure 1a, 1b, 1c and 1d the effect of an increasing window size on the initial estimate of the topographic primal sketch of a simple scene composed of one sphere, a cylinder and polyhedrons.

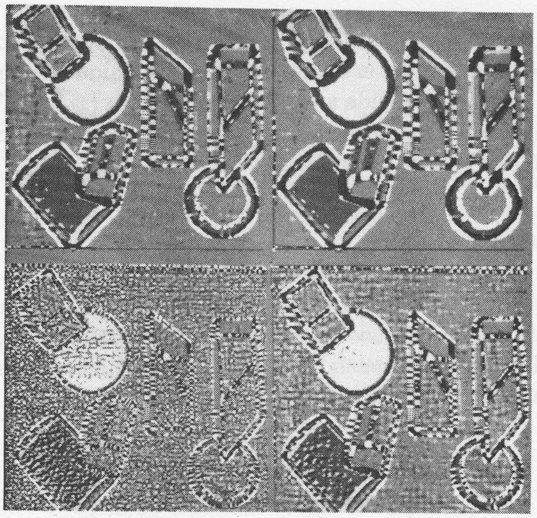

Figure 1: Initial primal sketch as a function of window size. (a) lower left, 5*5, (b) lower right, 7*7, (c) upper left, 9*9, (d) upper right, 11*11.

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 2 a two-dimensional histogram of the K and H for a flat surface.

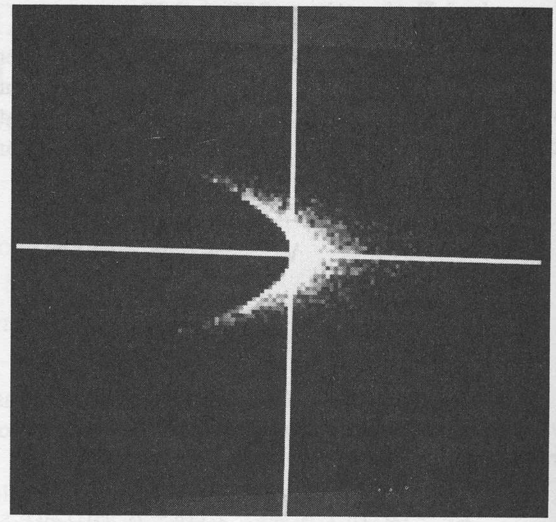

Figure 2: Bi-dimensional histogram of K and H. Horizontal axis is K and vertical axis is H.

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 it.

Basically, relaxation labelling is an iterative precedure 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 neighbourhood can be measured by what is known as the label's support

$$S(\lambda,\lambda') = \sum_{i=1}^{n} R_{ij}(\lambda,\lambda')p(\lambda) \tag{29}$$

, which is a function of other label certainties in the neighbourhood $p(\lambda)$ and their compatibility $R_{ij}(\lambda,\lambda')$ (pair-wise) 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. Relaxation labelling is the process of achiving global consistency by iteratively optimizing the local consistency.

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

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. Support for the label $p(\lambda)$ corresponds to a normalized distance from the threshold values ϵ_k and ϵ_h. For example, $p(\lambda) = 1$ for K=0 and/or H=0 and $p(\lambda) = 0$ for K=ϵ_k and/or H= ϵ_h.

	1	2	3	4	5	6	7	8	9
1	2	1	0	1	1	0	0	0	0
2	1	2	1	1	1	1	0	0	0
3	0	1	2	0	1	1	0	0	0
4	1	1	0	2	1	0	1	0	0
5	1	1	1	1	2	1	1	0	1
6	0	1	1	0	1	2	0	0	1
7	0	0	0	1	1	0	2	0	0
8	0	0	0	0	0	0	0	2	0
9	0	0	0	0	1	1	0	0	2

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

Label Relaxation Process

Our relaxation process is similar to the one developed by Hummel and Zucker [Humm 83] 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 functionel expressed by

$$S(\lambda, \lambda') = \sum_{window} R_{ij}(\lambda, \lambda')p(\lambda) \qquad (30)$$

is optimized. That is, in a neighbourhood of 3 * 3 find the best label at the center of this window that optimizes the label support.

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 3a the initial estimate of the topographic primal sketch of a range image illustrated at Figure 4. Figure 3b, 3c, and 3d illustrate the evolution of the primal sketch as a function of the relaxation labelling process.

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 convergence rate is possible if we could use a larger window for the label optimization, but consistency matrix would be more complicated since we are not optimizing with immediate neighbours.

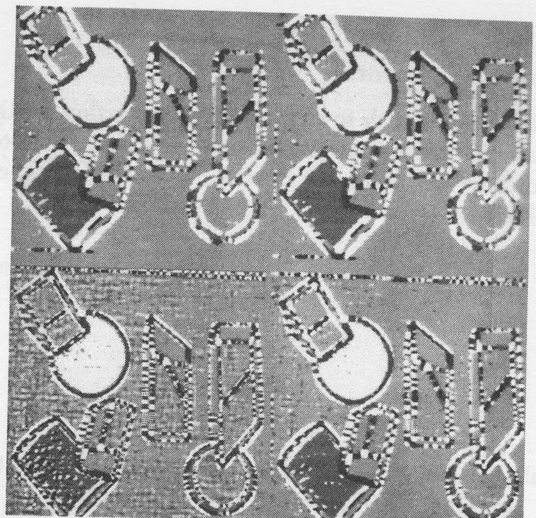

Figure 3: Relaxation of the primal sketch vs the number of iterations. (a) lower left (0) iteration, (b) lower right (1) iteration, (c) upper left (5) iterations, (d) upper right (10) iterations.

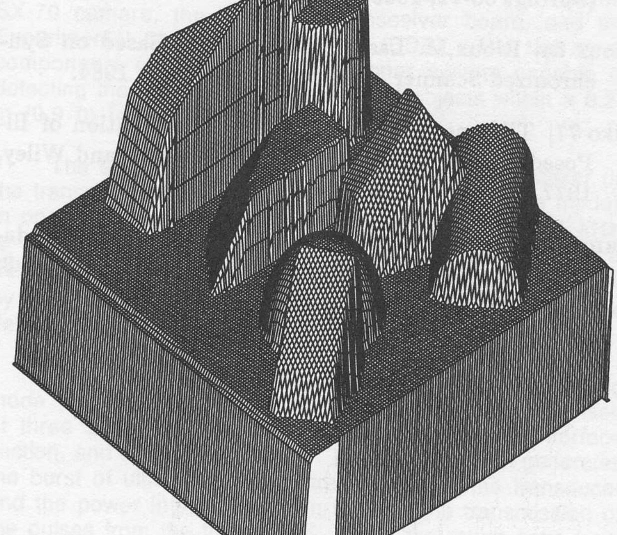

Figure 4: Isometric view of the range image used in the segmentation.

Conclusion

We have demonstrated that 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 improvement 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 K and H are numerically unstable which means that a regularized version of this operator is necessary.

Acknowledgements

The author wish to express his appreciation for the image acquisition to Marc Rioux and Luc Cournoyer from the Laboratory of Intelligent Systems . Their assistance was invaluable.

References

[Besl1 86] Besl. P.J. Surfaces in Early Range Image Understanding, PhD Thesis. University of Michigan, 1986.

[Besl2 86] 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.

[Hara 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.

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

[Tiko 77] Tikhonov, A.N., Arsenin, V.Y., Solution of Ill-Posed Problems, Washington,DC:Winston and Wiley, 1977.

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