Intrinsic Filtering of Range Images Using a Physically Based Noise Model

Pierre Boulanger[1], Olli Jokinen[2], J-Angelo Beraldin[3]
[1]Department of Computing Science
University of Alberta
Edmonton, Alberta, Canada, T6G 2E8
[2]Institute of Photogrammetry and Remote Sensing
Helsinki University of Technology
P.O. Box 1200, FIN-02015 HUT, Finland
[3]Institute for Information Technology
National Research Council Canada
Ottawa, Ontario, K1A 0R6

Abstract

This paper presents a new multi-scales range data filtering technique which produces a scale-space filtering analogous to Gaussian filtering but has several interesting properties such as viewpoint invariance and automatic edge preservation. One of the main contribution of this paper is that it takes into account a physical model of the sensor to ensure optimum filtering of the signal. Using this filter, new algorithms can be developed to detect at multi-scale depth and orientation discontinuities or segment robustly range data based on the sign of Gaussian and mean curvatures.

1 Introduction

More often than not, the computer vision community has processed range images with techniques developed in the context of intensity images, in spite of the fact that range data and images are a sampled set of measurements corresponding to a 3–D surface, observed from a particular viewpoint. In that sense, they are fundamentally different from intensity images. Haralick et al. [13] and later Besl and Jain [4, 5] were among the first to consider range images as true geometric information, using differential geometry to compute intrinsic properties of surfaces such as the Gaussian and mean curvatures.

To properly detect geometric features in range images one must be able to represent them at various scales in order to compute variation of surface properties of increasing globality. In the past, several authors have proposed different methods to solve this problem. One of these methods consists of convolving the image signal with Gaussian kernels of increasing size (scale) and then analyzing the evolution of signal features along the scale dimension. These stack of images as a function of increasing inner scale was coined as a linear "Scale Space" by Witkin [Witkin 1983] and Koenderink [Koenderink 1984]

Instead of using a few discrete scales, Witkin [31] has proposed the use of a continuum of scales and has studied the properties of this scale space for 1-D signals. In the case of more complex signals, the discretization of the formulation led to nontrivial heuristics to establish the correspondence between scales.

Koenderink [15] realized that the generating equation of linear scale-space can be expressed by a linear diffusion equation:

$$\frac{\partial L}{\partial s} = \vec{\nabla} \cdot \vec{\nabla} L = \triangle L = L_{xx} + L_{yy} \qquad (1)$$

for the 2–D case. This equation states that the derivative to scale s equals the divergence of the gradient of the luminance function L, which is the Lapacian, the sum of the second derivatives. The blurring is considered as a diffusion of the intensity over time where time is the scale parameter. The standard Gaussian kernel is the Green's function of the diffusion equation. The derivation of the diffusion equation for scale-space by many other researchers among which we mention:

- Analysis of Causality [15]: coarser scales can only be the causal result of what happened in the finer scales;

- Maximum principle [14]: any increase of the inner scale the maximum luminance of the coarser scale

is always lower than the maximum intensity at the finer scale, the maximum is always larger;

- No new extrema are formed at larger scales [18]

This method was applied to range data analysis with some success by Asada and Brady [2] for 2-D curves and then by Ponce and Brady [24] and Fan et al. [12].

Adaptive smoothing, instead of Gaussian smoothing, has been proposed to ease the interpretation of scale-space representations. The general idea behind adaptive smoothing is to apply variable convolution kernels, whose supports and shapes vary with the local properties of the signal to be smoothed. An overview of adaptive smoothing techniques can be found in Mastin [20] and more recently in Romery [25]. One of the first investigations of adaptive smoothing is done by Lev et al. [17] where an iterative weighted averaging method is proposed. The weight of each point in the convolution kernel is determined locally by calculating the difference between the pixel value at the center of the kernel and the pixel value of the neighborhood inside the kernel. A similar method is suggested by Wang et al. [32], in which the weights of the coefficients are defined as the normalized inverse of the signal gradient magnitude.

Other methods [9, 24, 30] perform adaptive smoothing based on variational methods. In these methods smoothing across previously detected discontinuities is prevented by the use of variable operators composed of computational molecules. One major drawback of these methods is that they imply a prior knowledge of the position of the signal discontinuities.

A different approach, proposed by Saint-Marc and Richetin [27], uses a directional mask in the direction of the lowest curvature in the case of highly curved areas and a standard square kernel in other regions. Other researchers such as Perona and Malik [22] have proposed the modelling of the image signal at all scales as the solution to the heat propagation equation in an anisotropic medium. This method seems to produce good results but requires extensive iterative calculations. Using a similar method, Saint-Marc et al. [28] have generalized and improved the method described by Perona by using repeated convolutions of a simple 3×3 operator where the coefficients of the window are modulated by a measure of the signal discontinuity at each point. This measure is defined as the magnitude of the gradient computed in a 3×3 window. Other work using similar approach can be classified in a number of distinct categories:

- Vector value diffusion [34], where the simulations influence of differential properties is viewed as a set of coupled PDE's

- Tensor value diffusion [33], where the conductance is made-up of the second moment matrix

$$D = \left\{ \begin{array}{cc} L_x^2 & L_x L_y \\ L_y L_x & L_y^2 \end{array} \right\}.$$

  The eigenvectors of this matrix indicate the main directions of the local structure. The consequence is that the diffusion is enhanced in the principle ridge direction;

- Affine and projective scale-space [21]: In this method the differential properties are expressed in the invariant arc-length.

- Other method use axiomatic morphological scale-spaces: Alvarez showed in a set of classical papers [1] an axiomatic approach to come to the equation $L_s = L_{vv}$ which he coined the "fundamental equation of image processing"

- Other method [11] use a diffusion process based on a robust curvature flow operator that achieves a smoothing of range data that is independent of parameterization and orientation.

One of the major problems in applying all of these methods to range data processing is that they are not invariant with respect to the viewpoint and do not take into account the fundamental difference in properties of the range image noise. Since range images are sampled versions of an object surface at a particular viewpoint, the actual sampling on the surface of the object is far from being isometric. Indeed, the 3–D distance between two consecutive surface samples depends upon the local curvatures of the scene and its orientation with respect to the sensor. As a consequence, linear methods, which use fixed weights in the convolution kernel, are not suitable because the effect on a specific 3–D surface will vary with this surface attitude in space. Also, contrary to intensity images the noise on range data is dependent on sensor parameters and on the orientation of the sensors relative to the surface. What is needed is a convolution kernel whose coefficients are locally established as a function of intrinsic surface properties and a local noise model. The second problem with these methods is that they do not take into account the very different nature of the noise properties of the signal.

In the method proposed herein, a distance measurement between a reference point $\mathbf{p}$ and a neighbor point $\mathbf{q}$ on a parametric surface $\mathbf{s}(u, v)$ is defined. This measurement (*the intrinsic surface distance*) is the length of the minimum trajectory $\mathbf{l}(t)$ on the surface joining these two points.

A new nonlinear operator that was first introduced in Boulanger and Cohen [6] is generalized

to be applied on a parametric surface $\mathbf{s}(u,v) = (x(u,v), y(u,v), z(u,v)^T$. This operator is defined as

$$G_\sigma * \mathbf{s}(u,v) = \mathbf{N}_\sigma{}^{-1} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{d_S^2(\mathbf{p},\mathbf{q})}{2\sigma^2}\right) \mathbf{\Sigma}^{-1}(\tau,\xi)\mathbf{s}(\tau,\xi)d\tau d\xi \tag{2}$$

where d_S represent the surface distance between two points on the 3–D surface, of parametric coordinates $\mathbf{p} = \mathbf{s}(u,v)$ and $\mathbf{q} = \mathbf{s}(u-\tau, v-\xi)$ respectively, and N_σ is the normalization factor equal to

$$\mathbf{N}_\sigma = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{d_S^2(\mathbf{p},\mathbf{q})}{2\sigma^2}\right) \mathbf{\Sigma}^{-1}(\tau,\xi)d\tau d\xi \tag{3}$$

The function $\mathbf{\Sigma}^{-1}(\tau,\xi)$ is the covariance matrix of the measured point determined by a physical model of the sensor. The parameter σ controls the size of the operator.

The following properties make such an operator attractive:

- Its effect on a 3-D surface is independent of the viewpoint
- It behaves as an anisotropic filter that performs adaptive smoothing on the signal but does not require prior determination of the surface discontinuities or curvature.
- It is an optimal filter based on a physical model of the sensor.

The following sections introduce a new model of a real range sensor noise and how to compute an optimal filter from this noise model. We also define what is an intrinsic distance d_S in the continuous case and then present an algorithm to compute a discrete approximation.

2 Range Sensors and Their Noise Model

The basic geometrical principle of optical triangulation is shown in Figure 1. A light beam generated by the laser is deflected by a mirror and scanned on the object. The returning light is then focussed by a lens onto a position sensitive photo-detector and measured. By simple trigonometry, the X and Z coordinates of the illuminated point on the object are calculated.

In conventional triangulation, a compromise among field of view, precision of the 3-D measurement, and shadow effect must be considered. A synchronized geometry provide a way to alleviate these tradeoffs. Rioux [26] introduced a synchronized scanning scheme, with a large field of view with small triangulation angle can be obtained without sacrificing the precision. By the fact that the range sensor have smaller triangulation angles shadow effects is reduced significantly. One can see in Figure 2 an illustration of NRC auto-synchronized approach for single axis system. Let p

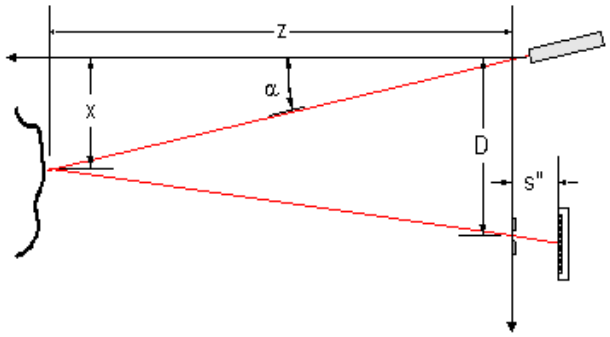

Figure 1: Basic geometrical principle of optical triangulation.

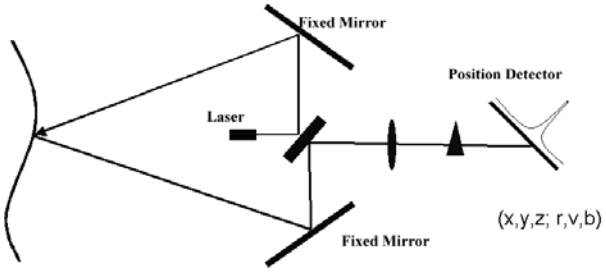

Figure 2: Schematic of NRC auto-synchronized range sensor geometry: single axis system.

be the measured position of a monochrome laser spot onto the detector, let θ and ϕ be the deflection angles of the laser beam in the dual axis range sensor. Then according to the convention set, the equation of a point (X,Y,Z) in the camera field of view as a function of the measured parameters (p,θ,ϕ) are expressed by the following calibration functions:

$$X(p,\theta,\phi) = X_s(p,\theta) \tag{4}$$

$$Y(p,\theta,\phi) = (Z_s(p,\theta) - h_y)\sin(2\phi) - h_z\cos(2\phi) + h_y \tag{5}$$

$$Z(p,\theta,\phi) = (Z_s(p,\theta)) - h_y)\cos(2\phi) - h_z\sin(2\phi) + h_z \tag{6}$$

The coordinate pair $X_s(p,\theta)$, and $Z_s(p,\theta)$ is the position of a point that would be measured with a range

camera having a single scan axis, i.e.:

$$X_s(p,\theta) = X_{-\infty}(\theta) + P_\infty \frac{X_0(\theta) - X_{-\infty}(\theta)}{P_\infty - p} \quad (7)$$

$$Z_s(p,\theta) = Z_{-\infty}(\theta) + P_\infty \frac{Z_0(\theta) - Z_{-\infty}(\theta)}{P_\infty - p} \quad (8)$$

For a given optical angle θ of the x-axis, $Z_{-\infty}(\theta)$ is the location of the vanishing point on the projection axis for $p \to -\infty$, and $Z_0(\theta)$ is the location corresponding to $p = 0$. Similar definitions apply to $X_{-\infty}(\theta)$ and $X_0(\theta)$. The parameter P_∞ is the location of the vanishing point on the detection axis, i.e.,

$$P_\infty = f_o \frac{\sin(\gamma)}{\cos(\beta - \gamma)} = \frac{(f_o - f)}{\sin(\beta)} \quad (9)$$

where f is the focal length of the collecting lens, f_o is the effective distance of the position detector to the imaging lens, β is the tilt angle of the position detector determined by the Scheimpflug condition, and γ is the triangulation angle corresponding to $p = 0$. A total of 25 parameters are used to describe a dual axis range sensor. These parameters include also distortion parameters not discussed here.

One of the fundamental differences between a range image and an intensity image is the fact that one can no longer assume that the noise present in the signal is a Gaussian with a constant variance. Even though the sensor parameters (p, θ, ϕ) are not statistically correlated, because of the calibration function, the resulting measurements X, Y, Z are. This implies that the noise on the X, Y and Z axis can be represented by a joint anisotropic Gaussian model expressed by:

$$p_r(\theta,\phi) = \frac{1}{N} \exp -\frac{1}{2}(\mathbf{r}(\theta,\phi) - \bar{\mathbf{r}}(\theta,\phi))^T \mathbf{\Sigma}^{-1}(\theta,\phi)(\mathbf{r}(\theta,\phi) - \bar{\mathbf{r}}(\theta,\phi)) \quad (10)$$

Where N is a normalization factor equal to: $\det(\mathbf{\Sigma}(\theta,\phi))^{\frac{1}{2}}(2\pi)^{\frac{3}{2}}$ and $\mathbf{\Sigma}(\theta,\phi)$ is equal to the variance-covariance matrix for each measurements and is defined by:

$$\mathbf{\Sigma}(\theta,\phi) = \begin{pmatrix} \sigma_{xx}(\theta,\phi) & \sigma_{xy}(\theta,\phi) & \sigma_{xz}(\theta,\phi) \\ \sigma_{yx}(\theta,\phi) & \sigma_{yy}(\theta,\phi) & \sigma_{yz}(\theta,\phi) \\ \sigma_{zx}(\theta,\phi) & \sigma_{zy}(\theta,\phi) & \sigma_{zz}(\theta,\phi) \end{pmatrix} \quad (11)$$

and is equal to $E((\mathbf{r}(u,v) - \bar{\mathbf{r}}(u,v))(\mathbf{r}(u,v) - \bar{\mathbf{r}}(u,v))^T)$ where $\bar{\mathbf{r}}(u,v) = E(\mathbf{r}(u,v))$. Assuming that the calibration functions $x = X(p,\theta,\phi)$, $y = Y(p,\theta,\phi)$, and $z = Z(p,\theta,\phi)$ have no sudden jumps in their corresponding domains then:

$$\sigma_{xx} \simeq \left(\frac{\partial X}{\partial p}\right)^2 \sigma_p^2 + \left(\frac{\partial X}{\partial \theta}\right)^2 \sigma_\theta^2 + \left(\frac{\partial X}{\partial \phi}\right)^2 \sigma_\phi^2 \quad (12)$$

$$\sigma_{yy} \simeq \left(\frac{\partial Y}{\partial p}\right)^2 \sigma_p^2 + \left(\frac{\partial Y}{\partial \theta}\right)^2 \sigma_\theta^2 + \left(\frac{\partial Y}{\partial \phi}\right)^2 \sigma_\phi^2 \quad (13)$$

$$\sigma_{zz} \simeq \left(\frac{\partial Z}{\partial p}\right)^2 \sigma_p^2 + \left(\frac{\partial Z}{\partial \theta}\right)^2 \sigma_\theta^2 + \left(\frac{\partial Z}{\partial \phi}\right)^2 \sigma_\phi^2 \quad (14)$$

$$\sigma_{xz} \simeq \left(\frac{\partial X}{\partial p}\frac{\partial Z}{\partial p}\right) \sigma_p^2 + \left(\frac{\partial X}{\partial \theta}\frac{\partial Z}{\partial \theta}\right) \sigma_\theta^2 + \left(\frac{\partial X}{\partial \phi}\frac{\partial Z}{\partial \phi}\right) \sigma_\phi^2 \quad (15)$$

$$\sigma_{yz} \simeq \left(\frac{\partial Y}{\partial p}\frac{\partial Z}{\partial p}\right) \sigma_p^2 + \left(\frac{\partial Y}{\partial \theta}\frac{\partial Z}{\partial \theta}\right) \sigma_\theta^2 + \left(\frac{\partial Y}{\partial \phi}\frac{\partial Z}{\partial \phi}\right) \sigma_\phi^2 \quad (16)$$

$$\sigma_{xy} \simeq \left(\frac{\partial X}{\partial p}\frac{\partial Y}{\partial p}\right) \sigma_p^2 + \left(\frac{\partial X}{\partial \theta}\frac{\partial Y}{\partial \theta}\right) \sigma_\theta^2 + \left(\frac{\partial X}{\partial \phi}\frac{\partial Y}{\partial \phi}\right) \sigma_\phi^2 \quad (17)$$

One can see in Figure 3 the variation of the noise on the variance of the coordinate Z as a function of distance.

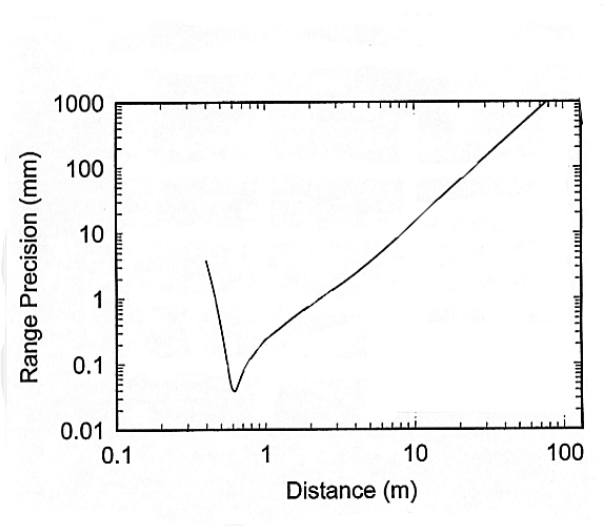

Figure 3: Variation of the noise on the variance of the coordinate Z as a function of distance.

2.1 Model for Pointing Error

It was shown experimentally in [23] that the precision of the range sensor also depends on the incident angle, i.e., the angle between the laser beam and surface normal. This error can be described as a pointing error which varies from point to point. In the following, we propose a mathematical model for the pointing error and present test results which support our model.

Let x', y', z' denote a rotating coordinate system where the origin is located in the laser projector $\mathbf{r}_0$ and the z'-axis points into the direction of the incident beam. The location of the laser projector can be solved using a model-based fitting approach described in [3] for the single scan axis case. The dual axis case requires

further studies so that in the following, the analysis is restricted to the single axis case where the scanning is performed by moving the object mechanically in the y-direction.

Let $\Delta\mathbf{r}_i$ be the error in a measurement point $\mathbf{r}_i$ propagated from $\Delta p_i, \Delta\theta_i, \Delta y_i$ through the calibration of the system. We have $\mathbf{r}'_i = R_y(\theta_i + \gamma/2)(\mathbf{r}_i - \mathbf{r}_0 - [0, y_i, 0]^T)$ where R_y denotes a rotation around the y-axis. The object surface is approximated as planar in the neighborhood of each measurement point. Figure 4 illustrates an intersection of the projection of the incident beam onto the surface in the plane determined by the z'-axis and the surface normal. It is proposed that the uncertainty in the position of the spot on the surface is proportional to the length of line AC. This provides an error estimate in the direction perpendicular to the surface normal. In the rotating coordinate system, the uncertainty is increased in the direction of the z'-axis by an amount proportional to the length of line BC. We have $|BC| = |AB| \tan\delta$, where δ is the incident angle and the x'- and y'-coordinates of A and B are solved from the equation system

$$(x'/\Delta x'_i)^2 + (y'/\Delta y'_i)^2 = 1, \qquad y' = (n'_{y,i}/n'_{x,i})x'. \tag{18}$$

After some manipulation, the error in z'_i, taking into account the incident angle, is given by

$$\Delta z'_{i,\text{new}} = \Delta z'_i + \frac{2\alpha((n'_{x,i})^2 + (n'_{y,i})^2)\Delta x'_i \Delta y'_i}{n'_{z,i}\sqrt{(n'_{x,i})^2(\Delta y'_i)^2 + (n'_{y,i})^2(\Delta x'_i)^2}}, \tag{19}$$

where α is a nonnegative parameter to be adjusted during calibration. The errors in x'_i and y'_i are considered to remain the same. The error in the sensor coordinate system is then given by $\Delta\mathbf{r}_{i,\text{new}} = R_y(-\theta_i - \gamma/2)\Delta\mathbf{r}'_{i,\text{new}}$ and the covariance matrix is obtained as

$$E(\Delta\mathbf{r}_{i,\text{new}}(\Delta\mathbf{r}_{i,\text{new}})^T). \tag{20}$$

If we assume statistical independence between the two model then the final combined model is simple and is equal to the sum of the covariance matrix estimated by equation (20) and equation (11).

3 Optimal Filtering Based on Maximum Likelihood Estimate

Let Ω be a neighborhood of size $\bar{m} = N \times N$ of a local averaging filter where

$$\mathbf{z} = (x(u_1, v_1), y(u_1, v_1), z(u_1, v_1), \cdots x(u_{\bar{m}}, v_{\bar{m}}), y(u_{\bar{m}}, v_{\bar{m}}), z(u_{\bar{m}}, v_{\bar{m}})) \tag{21}$$

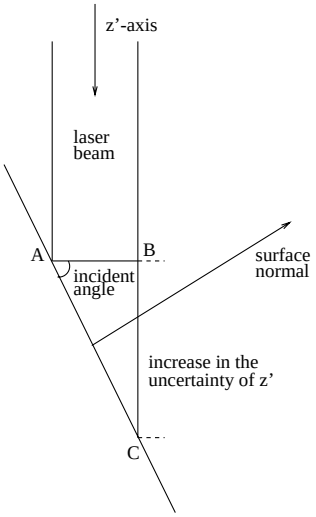

Figure 4: Geometric interpretation of the model proposed.

are the measurements produced by the range sensor in the neighborhood. The problem of filtering in the maximum likelihood framework can be defined as: Given the neighborhood data $\mathbf{Z}$, the maximum likelihood of the center pixel to be $\hat{\mathbf{r}}$ is provided by:

$$p_z(\mathbf{S}|\mathbf{Z}) = \max_{\mathbf{A}} \prod_{u,v\in\Omega} p_{r|A}(\mathbf{r}(u,v)|\mathbf{S}(u,v;\mathbf{A})) \tag{22}$$

where $p_z(\mathbf{S}|\mathbf{Z})$ is the likelihood that the data $\mathbf{Z}$ in the neighborhood including the central pixel can be approximated by a local surface model $\mathbf{S}(u,v;\mathbf{A})$ defined by a set of parameters $\mathbf{A}$. The likelihood is specified by the assumed probability distribution of the fluctuations in the measurements regarding their predicted values in the absence of noise.

The local neighborhood is modelled by the following parametric polynomial:

$$\mathbf{S}(u,v;\mathbf{A}) = \sum_{i=0}^{k}\sum_{j=0}^{l} \mathbf{a}_{ij} B_i(u) B_j(v) \tag{23}$$

where $\mathbf{a}_{ij}$ are the coefficient vectors of each component of $\mathbf{S}(u,v;\mathbf{A})$. The function $B_k(l)$ is the basis function. Let u, v be the parameterization of the surface where $u, v \in U \times U$, and N are assumed to be odd. The set U is defined as:

$$U = \{-(N-1)/2, \ldots, -1, 0, 1, \ldots, (N-1)/2\}. \tag{24}$$

Using a matrix notation the local surface model $\mathbf{S}(u,v;\mathbf{A}) = \mathbf{M}(u,v)\mathbf{A}$ is expressed by the following

equation:

$$\mathbf{S}(u,v;\mathbf{A}) = \begin{bmatrix} \mathbf{H}(u,v) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}(u,v) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}(u,v) \end{bmatrix} \begin{bmatrix} \mathbf{a}_{x|00} \\ \mathbf{a}_{x|01} \\ \vdots \\ \mathbf{a}_{x|kk} \\ \mathbf{a}_{y|00} \\ \mathbf{a}_{y|01} \\ \vdots \\ \mathbf{a}_{y|kk} \\ \mathbf{a}_{z|00} \\ \mathbf{a}_{z|01} \\ \vdots \\ \mathbf{a}_{z|kk} \end{bmatrix} \tag{25}$$

where $\mathbf{A}$ is a columns matrix of size $(k+1)(k+1)3 \times 1$ representing the coefficients of the model. The matrix $\mathbf{M}$ is a $3 \times (k+1)(k+1)$ representing the basis functions of the interpolating polynomial expressed in compact form by
$\mathbf{H}(u,v) = (B_0(u)B_0(v),\ B_0(u)B_1(v),\ \cdots,\ B_k(u)B_k(v))$. The notation $\mathbf{0}$ correspond to a null vector of size $1 \times (k+1)^2$.

In order to evaluate the value of the central pixel from the model parameter vector $\mathbf{A}$ a maximum a likelihood estimator is used. In this context, it is assume that $\mathbf{A}$ and $\mathbf{n}(u,v)$ are independent and distributed in a Gaussian fashion at each point so that:

$$\begin{aligned} E(\mathbf{A}) &= \bar{\mathbf{A}} \qquad var(\mathbf{A}) = \mathbf{V}_A \\ E(\mathbf{n}(u,v)) &= \mathbf{0} \quad var(\mathbf{n}(u,v)) = \mathbf{\Sigma}(u,v) \end{aligned} \tag{26}$$

The measurements performed by the range sensor can then be represented by the following equation:

$$\mathbf{r}(u,v) = \mathbf{M}(u,v) \quad \mathbf{A} + \mathbf{n}(u,v) \tag{27}$$

corresponding to a linear combination of the parameter $\mathbf{A}$ and observation noise $\mathbf{n}(u,v)$. All the densities in the foregoing expressions are assumed Gaussian, with

$$E(\mathbf{r}(u,v)) = \mathbf{M}(u,v)\ \bar{\mathbf{A}}\ var(\mathbf{r}(u,v)) = \mathbf{\Sigma}(u,v) \tag{28}$$

The expression of the various density are :

$$\begin{aligned} p_{r|A}(\mathbf{r}|\mathbf{S}) = &\left((2\pi)^{3/2}(det(\hat{w}\mathbf{\Sigma}))^{1/2}\right)^{-1} \\ &\times \exp\left(-\frac{\hat{w}\mathbf{v}^T\mathbf{\Sigma}^{-1}\mathbf{v}}{2}\right) \end{aligned} \tag{29}$$

where $\hat{w} = w(u,v)$ is an intrinsic weighting function which limit the effect of neighbour geometrically distant from the central pixel and $\mathbf{v}(u,v) = (\mathbf{r}(u,v) - \mathbf{M}(u,v)\mathbf{A})$.

The probability distribution of the estimated parameters a are equal to::

$$p_{\mathbf{A}|\mathbf{Z}}(\mathbf{A}|\mathbf{Z}) = \frac{1}{(2\pi)^{N/2} det(\mathbf{\Xi})^{1/2}} \exp(-\frac{1}{2}(\mathbf{A}-\hat{\mathbf{A}})^T \mathbf{\Xi}^{-1}(\mathbf{A}-\hat{\mathbf{A}})) \tag{30}$$

where

$$\begin{aligned} \hat{\mathbf{A}} = &\left(\sum_{u,v\in\Omega} w(u,v)(\mathbf{M}^T(u,v)\ \mathbf{\Sigma}^{-1}(u,v)\ \mathbf{M}(u,v))\right)^{-1} \\ &\left(\sum_{u,v\in\Omega} w(u,v)\mathbf{M}^T(u,v)\mathbf{\Sigma}^{-1}(u,v)\mathbf{r}(u,v)\right) \end{aligned} \tag{31}$$

is the maximum likelihood estimate of $\mathbf{A}$ and

$$\mathbf{\Xi} = (\sum_{u,v\in\Omega} w(u,v)(\mathbf{M}^T(u,v)\ \mathbf{\Sigma}^{-1}(u,v)\ \mathbf{M}(u,v)))^{-1} \tag{32}$$

is maximum likelihood estimate of the covariance matrix on the estimated parameters. The weighting function used in this regression problem was first introduced by Boulanger and Cohen [6], and is defined as:

$$w(u,v) = \frac{\exp(-(d_S^2(\mathbf{p},\mathbf{q})/2\sigma^2)}{\sum_{(u,v)\in\Omega} \exp(-(d_S^2(\mathbf{p},\mathbf{q})/2\sigma^2)}, \tag{33}$$

The parameter σ controls the scope of the function and determine in practice the size of the neighborhood. The present weights have the following attractive properties:

- Larger weights are assigned for the points which are geometrically closer with the center point.
- The weights are sensitive depth discontinuities. Thus, small weights are assigned to points which are in the opposite side of a discontinuity.
- The weights are independent of viewpoint since there are based on an intrinsic surface metric.

3.1 Filtering Based on Constant Assumption

One of the most widely used local model for filtering is based on the assumption that the neighborhood can be approximated by a constant value. In this particular case the filter expression is given by:

$$\begin{aligned} \hat{\mathbf{A}} = &\left(\sum_{u,v\in\Omega} w(u,v)\mathbf{\Sigma}^{-1}(u,v)\right)^{-1} \times \\ &\left(\sum_{u,v\in\Omega} w(u,v)\mathbf{\Sigma}^{-1}(u,v)\mathbf{r}(u,v)\right) \end{aligned} \tag{34}$$

which can be generalized for the continuous domain by Equation (1).

4 Intrinsic Surface Distances

To define a distance on a parametric surface $\mathbf{s}(u, v)$, it is convenient to express the first fundamental form of the surface at a point $\mathbf{p} = \mathbf{s}(u_p, v_p)$ in the basis $\{\mathbf{s}_u, \mathbf{s}_v\}$ associated with the parameterization u, v. Let $\mathbf{l}'(t)$ be the tangent vector to a parameterized curve $\mathbf{l}(t) = \mathbf{s}(u(t), v(t)), t \in (-\epsilon, \epsilon)$ embedded in the surface and going through the point $\mathbf{p}$. Then the first fundamental form at $\mathbf{p}$ is given by

$$I_p(\mathbf{l}'(0)) = E(u')^2 + 2Fu'v' + G(v')^2 \qquad (35)$$

where $E =< \mathbf{s}_u, \mathbf{s}_u >_p$, $F =< \mathbf{s}_u, \mathbf{s}_v >_p$, and $G =< \mathbf{s}_v, \mathbf{s}_v >_p$ are the coefficients of the first fundamental form in the basis of the tangent plane at the point $\mathbf{p}$. The operator $< \cdot, \cdot >$ corresponds to the inner product. The variables u' and v' are the derivatives with respect to the parameter t.

By using the first fundamental form, one can treat metric questions on a regular surface without reference to the ambient space. For example, the arc length s_l between two points $\mathbf{p} = \mathbf{l}(t_p)$ and $\mathbf{q} = \mathbf{l}(t_q)$ along a parameterized curve $\mathbf{l}(t)$ is given by

$$s_l(\mathbf{p}, \mathbf{q}) = \int_{t_p}^{t_q} |\mathbf{l}'(t)| dt = \int_{t_p}^{t_q} \sqrt{I_p(\mathbf{l}'(t))} dt \qquad (36)$$

Since only discrete surface measurements are available on discrete parametric coordinates (u_i, v_i), one must have a discrete form of equation (36) to compute the arc length between two points along the curve.

Let us consider a partition of a curve $\mathbf{l}(t)$ defined as $\mathbf{l}(t_i), t_p = t_0 < t_1 < \ldots < t_k < t_{k+1} = t_q$. If the steps of the piecewise approximation are sufficiently small, one can approximate the curve by a linear equation of the form

$$\mathbf{l}(t) = (x(t_i) + a_i(t - t_i), y(t_i) + b_i(t - t_i), z(t_i) + c_i(t - t_i)) \qquad (37)$$

where $a_i = [x(t_{i+1}) - x(t_i)]/(t_{i+1} - t_i)$, b_i, and c_i have similar expressions based on y and z. Then the derivative $\mathbf{l}'(t)$ of $\mathbf{l}$ with respect to t is given by

$$\mathbf{l}'(t) = (a_i, b_i, c_i) \quad \text{for} \quad t_i < t < t_{i+1}$$

Thus a discrete approximation of the arc length between $\mathbf{p}$ and $\mathbf{q}$ is given by

$$s_l(\mathbf{p}, \mathbf{q}) = \sum_{i=0}^{k} \sqrt{a_i^2 + b_i^2 + c_i^2} \qquad (38)$$

This is equivalent to a polygonal approximation of the surface.

Then the surface distance d_S, namely the minimum distance among all trajectories joining the two points, is defined by

$$d_S(\mathbf{p}, \mathbf{q}) = \min_l \, s_l(\mathbf{p}, \mathbf{q}) \qquad (39)$$

Note that the value of d_S will be large if the minimum trajectory goes across a depth discontinuity.

4.1 Algorithm to Find Minimum Trajectory

To obtain surface distances from the center point to the other points in a moving window, one needs to find minimum trajectories from the center point to all other points in the window. From the definition (39), one can design an efficient algorithm by using the Single-Source Shortest Paths Algorithm for weighted graphs (for example, see [19]).

The vertices of the graph correspond to the points in the window and the edges represent neighboring connections of points. From the equation (38), the arc length of the edge between points (u_{t_i}, v_{t_i}) and $(u_{t_{i+1}}, v_{t_{i+1}})$ is given by

$$\sqrt{a_i^2 + b_i^2 + c_i^2}$$

The algorithm to find minimum distances is given as follows.

Input: $G = (V, E)$ (a weighed graph) and v (the source vertex corresponding to the center point)

Output: for each vertex w, w.SP is the length of the shortest path from v to w and corresponding minimum trajectories w.TR

begin
for all vertices w **do**
    /* w.mark indicate if the vertex distance is determined*/
    w.mark := false;
    w.SP := ∞;
v.SP := 0;
v.TR := 0;
while there exists an unmarked vertex **do**
    let w be an unmarked vertex such that w.SP is minimum;
    w.mark := true;
    for all neighboring edges (w, z) such that z is unmarked **do**
        if w.SP + length$(w, z) < z$.SP **then**
            z.SP := w.SP + length(w, z);
            z.TR := w;
end

In this algorithm, one needs to find the minimum distances among a set of path lengths and to update the path lengths frequently. One can implement this efficiently by using a heap. All unmarked vertices are kept in a heap with their current known shortest path lengths from the center point v as their keys. To find an unmarked vertex w such that the path length w.SP is minimum, one can simply take it from the top of the heap. All the edges connected to the vertex w

can be checked and the path lengths can be updated without difficulty. Since the elements of the heap are the vertices of the graph, the space requirement is only $O(|V|)$, where $|V|$ denotes the number of vertices. The algorithm complexity is $O((|E|+|V|)\log|V|)$, where $|E|$ denotes the number of edges.

5 Experimental Results

5.1 Experimental Verification of the Noise Model

The model in Eq. 19 was tested using the data in [23]. This experiment involved data from a single scan line of a planar surface measured from different viewpoints. The orientation of the sensor and the surface normals were known accurately and furthermore, $n'_{y,i} = 0$, $i = 1, \ldots, N$. Only one scan line was used to avoid errors related to the object movement and the y'-coordinates were considered noise-free. The measuring was repeated several times which provided sample estimates for the covariance matrices. The traces of these matrices have been plotted as a function of the incident angle in Figure 5a.

In this test case, the proposed model yields

$$\begin{aligned}\mathrm{Tr}(\mathrm{Cov}(\mathbf{r}_i)_{\mathrm{new}}) &= \mathrm{Tr}(\mathrm{Cov}(\mathbf{r}'_i)) \\ &+2\alpha\mathrm{Cov}(x'_i, z'_i)\tan(\beta+\delta) \\ &+\alpha^2\mathrm{Var}(x'_i)\tan^2(\beta+\delta)\end{aligned} \tag{40}$$

where δ is an additional shift parameter introduced to obtain a better fit between the observations and the model. Instead of evaluating $\mathrm{Cov}(\mathbf{r}'_i)$ through error propagation from $\Delta p_i, \Delta\theta_i$, each element of the covariance matrix was estimated as a sample mean over observations where the incident angle was less than a threshold. The same sample means were used at all points $\mathbf{r}'_i$, $i = 1, \ldots, N$. A better result might be obtained if the elements were interpolated from curves describing their behavior as a function of the measuring distance.

The axis of the incident angle was divided into T segments S_t of equal length the centers of which were located at $\beta_t = \pi(t-0.5)/(2T)$, $t = 1, \ldots, T$. The model was fitted to the mean traces of covariance matrices in each segment. More precisely, the parameters α and δ were estimated minimizing the merit function.

$$f(\alpha,\delta) = \sum_{t=1}^{T}[\mathrm{Tr}(\mathrm{Cov}(\mathbf{r})_{\mathrm{new}})(\beta_t,\alpha,\delta) \tag{41}$$

$$- \sum_{\beta_i \in S_t} \mathrm{Tr}(\mathrm{Cov}(\mathbf{r}_i))/N_t]^2, \tag{42}$$

where N_t is the number of points in S_t, $t = 1, \ldots, T$. The nonlinear least squares problem in Eq. 42 was solved using the Newton-Gauss method. The result of fitting is illustrated in Figure 5b.

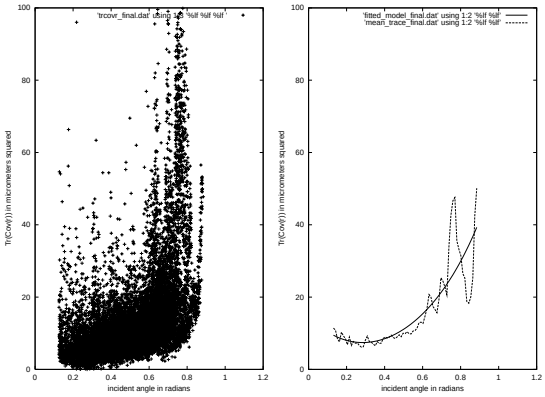

Figure 5: a) $\mathrm{Tr}(\mathrm{Cov}(\mathbf{r}_i))$, $i = 1, \ldots, N$. b) Mean traces of the sample covariance matrices in each segment and the model fitted.

5.2 Range Image Filtering Results

In order to illustrate the property of this new filter we did a series of experiments on a complex scene composed of quadric surface of various shapes and scales. One can see in Figure 6 the evolution of the surface reconstruction as a function of σ for an image corrupted by a noise of 5 db. On can see in Figure 7 (a),(c), and (d) the scale-space evolution of a standard Gaussian kernel and in Figure 7 (b),(d), and (e) the same evolution with the intrinsic filter. From observation it is obvious that contrary to the Gaussian kernel, the new filter preserves the position of the major discontinuities and smooth-out the uniform regions. In order to illustrate this unique property further, on can see in Figure 8 the evolution of depth discontinuity for a Gaussian kernel as compared to the new one. From observation, one can notice that the corner edge were significantly displaced with the Gaussian kernel and that no such phenomena appended with the new operator.

Finally, one can see in Figure 9 an illustration of the evolution of the filter for a range image of a face scanned with a range sensor rotating around the person. One can see that some of the surface detail disappear as a function of scale without the problem of blurring the large depth discontinuities. This property allows us to reduce significantly the noise to compute surface curvature using a method described in [7].

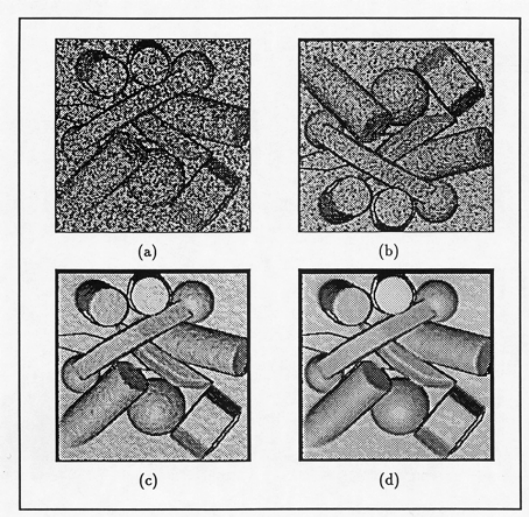

Figure 6: Evolution as a function of σ of the surface reconstruction of a range image composed of quadrics corrupted by a noise of 5db: (a) $\sigma = 0.0$, (b) $\sigma = 1$, (c) $\sigma = 2.0$, (d) $\sigma = 3.0$.

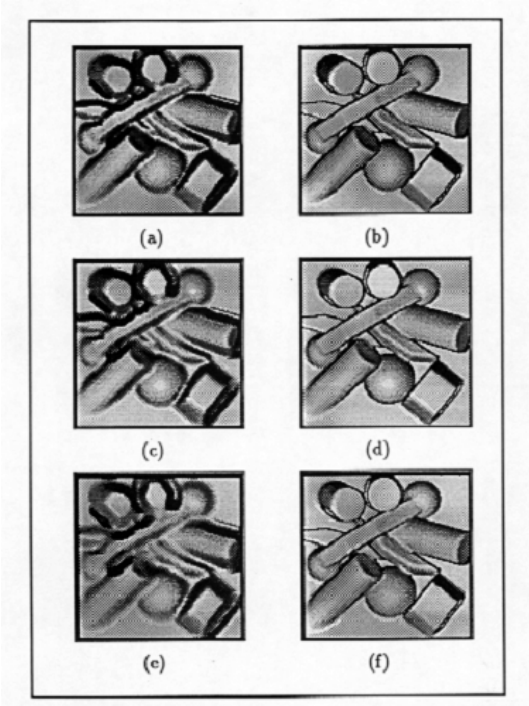

Figure 7: Evolution of the surface reconstruction of a range image composed of quadrics for a Gaussian kernel (right) and an intrinsic kernel (left): (a)-(b) $\sigma = 2.5$, (c)-(d)$\sigma = 4.5$, (e)-(f)$\sigma = 6.0$.

6 Conclusion

One of the goal of this paper is to present a coherent framework for the filtering of range images based on fundamental principles. As discussed in the introduction range image is a fundamentally different signal from normal intensity images produced by CCD cameras. We have demonstrated experimentally that the noise of these signal is indeed very different and much more complex. The simple noise model presented here is a relatively good approximation of the real one but there are many more factors contributing to the precision of these range sensors that is not modelled here. For instance, the contribution of laser intensity fluctuations and the variation of the surface albedo does have an influence on the precision of the laser scanner and should be modelled.

This paper is a new and improved version of a publication that was made over ten years ago [6]. In the previous paper, we introduced the concept of intrinsic surface distance and some of its very interesting properties. However, at the time we did not had any knowledge of the real stochastic properties of this signals and its consequence to the process of filtering. In this paper, we demonstrate that one can create a scale-space based on intrinsic properties that also take into account the very different nature of these signals. We also experimentally demonstrate in this paper that the new scale space has some of the key properties necessary for a good scale-space i.e., Causality, Localization, and maximum principle. Over the years we have applied this filter to various applications and demonstrated its usefulness for curvature segmentation [7], edge detection [8], and many others.

7 Acknowledgments

The model for pointing error was developed when O. Jokinen participated in the visiting researcher program of National Research Council Canada.

References

[1] L. Alvarez, F. Guichard, P.L. Lions,and J.M. Morel, *Axiomes et equations fondamentales du traitement d'image*, C.R. Acad. Sci. Paros, Vol. 315, pp. 135–138, 1992.

[2] H. Asada and M. Brady, *The curvature primal sketch*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8, pp. 2–14, 1986.

[3] J.-A. Beraldin, M. Rioux, F. Blais, G. Godin, and R. Baribeau, *Calibration of an Auto-synchronized Range Camera with Oblique Planes and Collinearity Equation Fitting*, NRC/ERB-1041, National

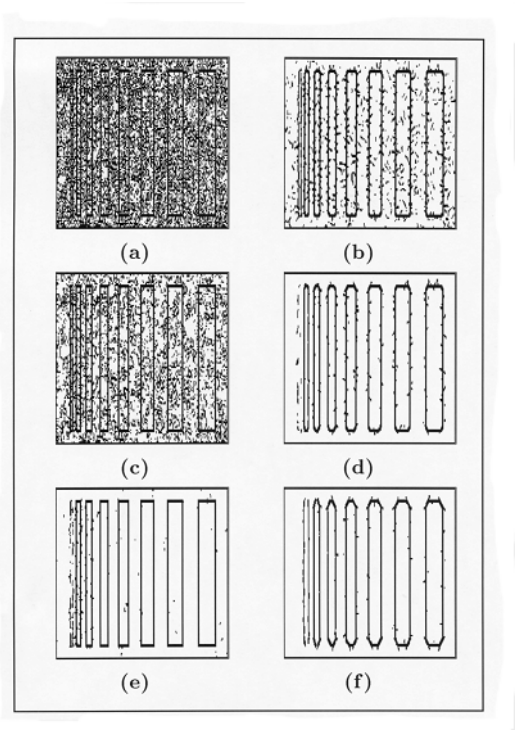

Figure 8: Evolution of the location of depth discontinuity for a Gaussian kernel (right) and an intrinsic kernel (left) for a signal corrupted by a noise of 6.5db: (a)-(b) $\sigma = 2.5$, (c)-(d)$\sigma = 4.5$, (e)-(f)$\sigma = 6.0$.

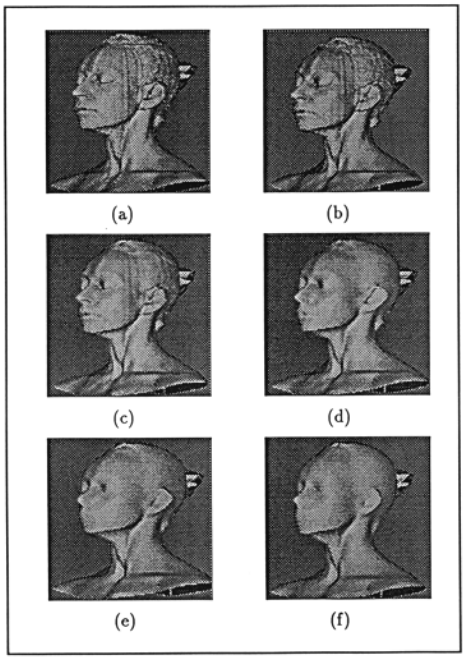

Figure 9: Evolution of surface reconstruction of a human face as a function of σ: (a) $\sigma = 0.0$, (b) $\sigma = 1.0$, (c) $\sigma = 2.5$, (d) $\sigma = 5.0$, (e) $\sigma = 7.5$, and (f) $\sigma = 10$.

Research Council of Canada, Ottawa, Ontario, November 1994.

[4] P.J. Besl, *Surfaces in Early Range Image Understanding*, PhD thesis,University of Michigan. 1986.

[5] P.J. Besl and R.C. Jain, *Invariant surface characteristics for three-dimensional object recognition in range images*, Comput. Vision, Graphics Image Process, 33, 1, pp. 33–80, 1986.

[6] P. Boulanger and P. Cohen, *Adaptive smoothing of range images based on intrinsic surface properties*, Proc. of the SPIE Technical Symposium on Optical Engineering and Photonics in Aerospace Sensing, Orlando, FL, April 16–20, pp. 254–263, 1990.

[7] P. Boulanger and P. Cohen, *Viewpoint invariant computation of surface curvatures in range images*, Proceedings of Vision Interface '94, Banff, Alta., pp. 145–154, May 16-20, 1994.

[8] Boulanger, P. *Multiscale edge detection based on a new geometrically intrinsic filter*, SPIE Proceedings, Videometrics III, International Symposium on Photonic and Sensors and Controls for Commercial Applications, Boston, MA, Vol. 2350. pp. 264-278, October 31 - November 4, 1994.

[9] J. Brady, J. Ponce, A. Yuille, and H. Asada, *Describing Surfaces* Journal of Computer Vision, Graphics, and Image Processing,32, pp. 1–28, 1985.

[10] R.T. Chin and C.L. Yeh, *Quantitative evaluation of some edge preserving noise smoothing techniques* Journal of Computer Vision, Graphics, and Image Processing, 23, pp. 67–91, 1983.

[11] M. Desbrun, M. Meyer, P. Schrder, A. Barr, *Implicit Fairing of Irregular Meshes using Diffusion and Curvature Flow*, in Siggraph '99, August, 1999.

[12] T.-J. Fan, G. Medioni, and R. Nevatia, *Segmented Description of 3-D Surfaces*, IEEE J. of Robotics and Automation, RA-3, 6, pp. 527-538, 1987.

[13] R.M. Haralick, L.T. Watson, and T.J. Laffey, *The topographic primal sketch*, Int. J. Robotics Res., 2, 1, pp. 50–72, 1983.

[14] R.A. Hummel and B.C. Gidas, *Zero crossing and heat equation*, Tech. Rep. 111, New York Univ.

Courant Intitute of Math. Sciences, Computer Sciences Division, 1984.

[15] J.J. Koenderink, *The stucture of images*, Biol. Cybern., Vol. 50, pp. 363–370, 1984.

[16] J.-G. Leu, I.K. Sethi and G. Yu,*Range Image Segmentation from Equidistant Contours*, Proc. IEEE Comput. Soc. Workshop on Comput. Vision, Miami Beach, Florida, Nov. 30- Dec. 2, pp. 339-341, 1987.

[17] A. Lev, S.W. Zucker, and A. Rosenfeld, *Iterative enhancement of noisy images*,IEEE Transactions on Systems, Man, and Cybernetics, 7, pp. 435–441, 1977.

[18] T. Lindeberg, *Scale-sapce for discrete signals*,IEEE Tans. Pattern Analysis and Machine Intelligence, vol 12, no.3, pp. 234–245, 1990.

[19] U. Manber, *Introduction to Algorithms, A Creative Approach*, Addison-Wesley, 1989.

[20] G.A. Mastin,*Adaptive filter for digital image noise smoothing: An evaluation*, Journal of Computer Vision, Graphic, and Image Processing, 31, pp. 103–121, 1985.

[21] P. Olver, G. Sapiro, A. Tannenbaum, *Differential invariant signatures and flow in computer vision: A symmetry group approach*, In Geometry-Driven Diffusion in Computer Vision (B.M. ter Haar Romery, ed.),Computational Imaging and Vision, pp.255–306, Dordrecht: Kluwer Academic Publishers, 1994.

[22] P. Perona and J. Malik,*Scale space and edge detection using anisotropic diffusion*, Proceedings of IEEE Workshop on Computer Vision, Miami, pp. 16–22, 1987.

[23] F. Prieto, T. Redarce, P. Boulanger, and R. Lepage, *Tolerance Control with high quality 3D data*, Third Conference on 3D Digital Imaging and Modeling, Qubec, Canada, May 28th-June 1st, 2001.

[24] J. Ponce and M. Brady, *Toward a surface primal sketch*, Proceeding of IEEE International Conference on Robotics and Automation, pp. 420–425, 1985.

[25] T.H. Romery, L.MJ. Florack, J.J. Koenderink, and M.A. Viergever, *Scale-Space Theory in Computer Vision*. Lecture Notes in Computer Science 1252, Sprigner Verlag, Berlin, 1997.

[26] M. Rioux, F. Blais, J.A. Beraldin and P. Boulanger, *Range image sensors development at NRC laboratories*, Proc. of the Workshop on Interpretation of 3D Scenes, Austin, TX, November 27–29, 1989.

[27] P. Saint-Marc and M. Richetin, *Structural filtering from curvature information*, Proceedings of IEEE Computer Vision and Pattern Recognition, Miami Fl., pp. 338–343,1986.

[28] P. Saint-Marc, J.S. Chen, and G. Medioni, *Adaptive smoothing a general tool for early vision*, Proceedings of IEEE Computer Vision and Pattern Recognition, pp. 618–624, 1989.

[29] G. Taubin,*Estimation of Planar Curves, Surfaces, and Nonplanar Space Curves Defined by Implicit Equations with Applications to Edge and Range Image Segmentation*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 13, No. 11, November 1991.

[30] D. Terzopoulos, *The role of constraints and discontinuities in visible surface reconstruction*, Proceedings of International Joint Conference on Artificial Intelligence, pp. 1019–1022, 1983.

[31] A.P. Witkin, *Scale-space filtering*, Proceedings of International Joint Conference on Artificial Intelligence, pp. 1019–1022, 1983.

[32] D.C.C. Wang, A.H. Vagnucci, and C.C. Li, *Gradiant inverse weighted smoothing scheme and the evaluation of its performance*, Journal of Computer Graphics, and Image Processing, 15, pp. 167–181, 1981.

[33] J. Weickert, *A review of non-linear diffusion filtering*, In Scale-Space Theory in Computer Vision, Vol. 1252 of Lecture Notes in Computer Science, pp.. 3–28, Sprigner, Berlin, 1997.

[34] R. Witaker and G. Gerig, *Vector-value diffusion* In Geometry-Driven Diffusion in Computer Vision, Computational Imaging and Vision, pp. 93–134, Kluwer Academic Publishing B.V., 1994.

[35] B.H. Zhuang, J.H. Zhang, C. Jiang, Z. Li, and W. Zhang, *Precision laser triangulation range sensor with double detectors for measurement on CMMs*, Industrial Optical Sensors for Metrology and Inspection, Proc. SPIE 2349, Boston, 1995, pp. 44–52.