Real-Time Range Image Segmentation
Using Adaptive Kernels and Kalman Filtering [†]

F. W. DePiero, EE Dept, CalPoly, San Luis Obispo, CA 93407, fdepiero@calpoly.edu
M. M. Trivedi, ECE Dept, UC-San Diego, La Jolla, CA 92093-0407, trivedi@ece.ucsd.edu

Abstract

Segmentation is a fundamental process affecting the over-all quality and utility of a machine vision system. Range Profile Tracking (RPT) is a systematic approach for stable, accurate and high speed segmentation of range images that is based on Kalman filtering. Tests of RPT have produced stable decompositions of second order surfaces bounded by jump and crease discontinuities, having a volumetric error of a few percent, in under 6 sec. for a wide variety of conditions. Results from over 900 tests on synthetic scenes and 150 real range images are presented.

1. Introduction

Segmentation is a fundamental process affecting the over-all quality and utility of a machine vision system. For range image analysis to be of practical value and utility, segmentation algorithms should provide results which satisfy often conflicting requirements of speed, accuracy and robustness. Segmentation is a pixel-level process for partitioning data into meaningful groups. Measurement noise makes segmentation a challenging problem because it corrupts the uniformity of range data. Outliers, that result from a sensing operation, or that are naturally present at the boundary of two distinct shapes, also make segmentation challenging by corrupting local curvature estimates.

Range image segmentation has been an active area of research for the past 15 years [11, 22]. A number of good survey articles [2, 19] are available. The spectrum of techniques can be described in terms of a footprint, or "kernel," that is used to access range data at the pixel-level [8]. The kernel has an associated shape, placement and processing scheme. A new trend in range image segmentation is the use of adaptive kernels. Adaptive techniques avoid the use of rigid geometries in determining kernel size or shape, for example. These types of techniques provide flexible means to locate kernels and use dynamic procesing mechanisms.

Adaptive kernels have an inherently improved robustness over rigid schemes, via their ability to avoid outliers. Some examples of other adaptive methods include schemes that selectively shape the kernel [5, 16], that prudently position it [3, 18, 24], and that use dynamic processing methods [13]. The technique described herein, **Range Profile Tracking**, (**RPT**) is novel in its use of adaptive mechanisms in all three of the above areas: shape, placement and processing.

2. Range Profile Tracking Approach

In RPT, kernel processing is approached as a real-time process of shape-profile tracking [8]. Tracking is accomplished using a Kalman filter [15]. This type of approach is similar to time-domain predictive tracking, but here, heights of measurements are tracked and the independent variable is spatial, not temporal. This tracking operation involves a single spatial variable and, overall, scene data is tracked in two orthogonal directions resulting in two overlapping sets of segmented vertical profiles. The overlapping sets of profiles are then combined to yield a 3-D segmented scene. Some results of the profile tracking are presented in Figure 2.

Hence the goal of RPT tracking is to segment range data into sets of vertical profiles. Each profile is to contain a series of end-to-end curved segments, known as "strips". The strips designate a portion of a profile having uniform curvature. Using Kalman filtering (KF) provides several benefits. First, the state of the filter provides a description of surface curvature. Secondly, the predictive capability provides a means for segmentation. The height of each subsequent measurement is predicted using the KF state. The state covariance estimate of the filter provides a means to express prediction error in terms of a Mahalanobis distance [12]. When the normalized prediction error exceeds a threshold, the current strip is terminated and a new one begun. This proceedure is depicted in Figure 1. A second criteria is also used to terminate strips. This uses the residual error between the original measurements and the strip curvature model given by the KF state. Another advantage of the KF is its computational efficiency when processing data containing a second order variation in shape. This reduces

Proceedings of ICPR '96

the computational burden that would otherwise be incurred when using least square-based approaches to second order curve fitting [16, 14].

Figure 1. *A Kalman filter is used to accumulate range measurements. The predictive capability of the filter provides a means to determine the boundaries of strips of range data. Strips designate portions of vertical profiles of having consistent curvature.*

In the RPT approach a kernel region is synonmous with a strip. Hence, the size of RPT kernels is adaptive, depending on the extent of consistent curvature. RPT kernels are adaptively located using the discontinuity and consistency (residual error) measures described above. Processing in the RPT kernel is state driven, making it adaptive, too.

Overall, one of the key computational advantages to RPT is the use of strips to form seed regions. This is advantageous because the strips are computed efficiently and because the seed regions tend to be relatively large, compared to the size of the original surface patches.

RPT has been integrated with a Structured Light (SL) sensor that generates verical profiles of range data. The sensor is known as **PRIME**, the PRofile Imaging ModulE [9]. Because PRIME generates profiles and RPT segments them, the two components are considered to be "well-matched" in terms of their exchanged data. These two components are also well-matched in terms of computational loading. The profile tracking operation of RPT is accomplished on-line, at frame rate, as profiles are generated by PRIME. (Seed growing is accomplished off-line in a post-processing step).

3. RPT Formulation

The Kalman trackers in RPT have been setup with the capability to follow a second order height variation, such as $z = f(x) = ax^2 + bx + c$. To track the second order shape profiles, state vectors for the in the form $\left| \begin{array}{ccc} z & \frac{dz}{dx} & \frac{d^2z}{dx^2} \end{array} \right|^T$

are used with a constant-acceleration type of filter [1]. This formulation is identical to the tracking of a time-domain signal that is modeled with a constant temporal acceleration, but for the exchange of the independent variable.

The KF state describes the curvature of a strip from the perspective of its endpoint. The state includes the current height and slope of the tangent as well as the strip's vertical acceleration (assumed constant). A continuous model of the system with zero noise would have $\frac{d^3z(x)}{dx^3} = 0$. In practice, changes in acceleration are due to process noise $\frac{d^3z(x)}{dx^3} = \tilde{\nu}(x)$ which is assumed to have a zero mean and a Gaussian density [1]. The process noise associated with shape estimation is simply the variation of surfaces beyond second order curvature. The measurement system, expressed in terms of a continuous position variable (x), is given by

$$\frac{da(x)}{dx} = \mathbf{A}a(x) + \left| \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right| \tilde{\nu}(x), \quad z_m = \mathbf{H}a(x) + \left| \begin{array}{c} 0 \\ 0 \\ 1 \end{array} \right| \tilde{w}(x) \quad (1)$$

where

$$\mathbf{A} = \left| \begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right|, \quad \mathbf{H} = \left| \begin{array}{ccc} 1 & 0 & 0 \end{array} \right|. \quad (2)$$

Here, $\mathbf{H}$ is the Kalman filter's measurement matrix and z_m is a height measurement. The measurement noise $\tilde{w}(x)$ is assumed to be Gaussian with zero mean. The state transition matrix associated with a sampling period of δx is

$$\mathbf{\Phi} = e^{\mathbf{A}\delta x} = \left| \begin{array}{ccc} 1 & \delta x & \frac{1}{2}\delta x^2 \\ 0 & 1 & \delta x \\ 0 & 0 & 1 \end{array} \right|. \quad (3)$$

The discrete form of the measurement matrix, $\mathbf{H}$, contains the same numerical values for both the continuous and discrete domains. It is assumed that the measurement uncertainty remains constant, and because measurements are scalar heights, the measurement covariance matrix reduces to a scalar variance σ_m^2. The intensity of the process noise is also assumed to remain constant. This yields a covariance matrix for the process noise of

$$\mathbf{Q} = \left| \begin{array}{ccc} \delta x^5/20 & \delta x^4/8 & \delta x^3/6 \\ \delta x^4/8 & \delta x^3/3 & \delta x^2/2 \\ \delta x^3/6 & \delta x^2/2 & \delta x \end{array} \right| q. \quad (4)$$

where q is the process noise parameter [1].

There are interesting tradeoffs between the measurement uncertainty σ_m^2 and the process noise q [4]. Relatively large q values will encourage RPT to track surfaces closely, but will also permit measurement noise to corrupt strip curvature, tending to shorten strips due to the residual error

limits. Relatively large σ_m discourage the KF from trusting new measurements. This worsens the tracking ability of the strips, but does reject more noise. A basis for selecting σ_m is available using the residual errors associated with sensor calibration [9]. The process noise parameter was determined experimentally. In all tests described herein $q = 100$ and $\sigma_m = 0.010$. Because of the low sensor noise associated with SL and a desire for high accuracy, q was chosen proportionally much higher than σ_m.

The nominal Kalman filter formulation allows for a number of time-varying quantities, such as the measurement matrix, process noise, measurement uncertainty and state transisition matrix. In RPT these quantities are time invariant. Because of this time-invariance and due to certain observability and controlability properties of this formulation [1], a number of the matricies in the Kalman filter equations converge to steady-state values. This is computationally advantageous, as it significantly reduces the necessary computations during RPT tracking.

The nominal formulation of the Kalman filter equations is used to initialize the filter [1]. A zero measurement is applied to the filter on successive iterations until the various KF matrices converge. Convergence is monitored by subtracting subsequent Kalman Gain matrices and applying the infinity norm to the difference. These "stead-state" calculations are given in Eq. 6 through 8.

Before accumulating a new measurement z_i^m into the filter, a discontinuity threshold is applied to test the validity of including z_i^m in the current strip. The discontinuity threshold is expressed in terms of the Mahalanobis distance, d_i, that describes the distance of the measurement away from the estimate of the current state [12, 1].

$$d_i = (z_i^m - \mathbf{H}\tilde{\mathbf{a}}_{\mathbf{i}})^T \Lambda^{-1}(z_i^m - \mathbf{H}\tilde{\mathbf{a}}_{\mathbf{i}}) \qquad (5)$$

where Λ is the measurement covariance matrix and $\tilde{\mathbf{a}}_{\mathbf{i}}$ is the estimate of the current state. For RPT, in steady state, this reduces to

$$d_i = (z_i^m - \tilde{z}_i)^2 / \sigma_m^2 \qquad (6)$$

where $\tilde{z}_i$ is the current height estimate, which is available in the vector containing the predicted state. The accumulation of a new measurement into the filter requires only 2 steps in steady state. First, the state must be advanced with

$$\mathbf{a}_{\mathbf{i}} = \tilde{\mathbf{a}}_{\mathbf{i}} + \mathbf{K}_{\mathbf{i}}(z_i^m - \mathbf{H}\tilde{\mathbf{a}}_{\mathbf{i}}) \qquad (7)$$

where, $\mathbf{K}_{\mathbf{i}}$ is the Kalman Gain matrix. Second, the state estimate for the next iteration is found with

$$\tilde{\mathbf{a}}_{\mathbf{i}+1} = \Phi\mathbf{a}_{\mathbf{i}}. \qquad (8)$$

4. Experimental Validation of RPT

Experimental validation of RPT was necessary in order to quantify performance. These metrics include computate

speed, accuracy and stability. Here stability refers to the consistency of results. Stable scene decompositions will remain consistent in size and shape in the presence of small perturbations of the input scene conditions.

A controlled environment was needed that allowed ground truth to be compared with processed results. Synthetic scenes provided this type of condition. In order to achieve statistically valid performance estimates, Monte Carlo analyses [20] were run in large numbers with randomly generated scenes. The tests with real scenes used objects with shapes similar to the synthetic ones.

A 150x150 grid of range data has been used with RPT in testing. Synthetic range data was corrupted with both zero-mean Gaussian and spike noise. Tests ramped the Gaussian noise level from $\sigma_m = 0.010$ to 0.030 inches. Spike noise was added with a probability of 5%. Note that spike noise was practically nonexistent in the real scene data.

Scene conditions for both real and synthetic tests were designed to include the fundamental three types of shape discontinuities: jump, crease and curvature [13]. These have discontinuities in the zeroth, first and second spatial derivatives of height, respectively. Both first and second order surfaces were also used. See Figure 2. The objects generated in each synthetic scene were randomly positioned and oriented, and had random dimensions. Disturbances in position and orientation were also introduced for the objects used in real scenes.

Three measures have been used to quantify RPT performance: (1) percentage of trials with correct number of segmented regions, (2) percent error in volume, and (3) compute speed. The percentage of trials with the correct number of final regions is recorded under "Percent Trials OK", in Table 1. It demonstrates the overall stability of the segmentation. Only on trials with the correct final number of regions were the other performance measures tallied. The tests indicate a volumetric error of less than 10% and processing time under 6 sec for both jump and crease discontinuities. These tests also had a stability measure that scored better than $\approx 90\%$. RPT processing of scenes with curvature discontinuities was somewhat less successful. These types of discontinuities are more difficult for the RPT approach to handle because their boundaries are not revealed as crisply during the initial profile segmentation process.

5 Concluding Remarks

The objective of this research was to develop a stable, accurate and high speed approach to range image segmentation. Available techniques demonstrate very good shape characterization abilities, but tend to be somewhat slow relative to RPT, typically requiring at least ≈ 1 minute of processing. RPT is an order of magnitude faster than these techniques. RPT has demonstrated a volumetric accuracy
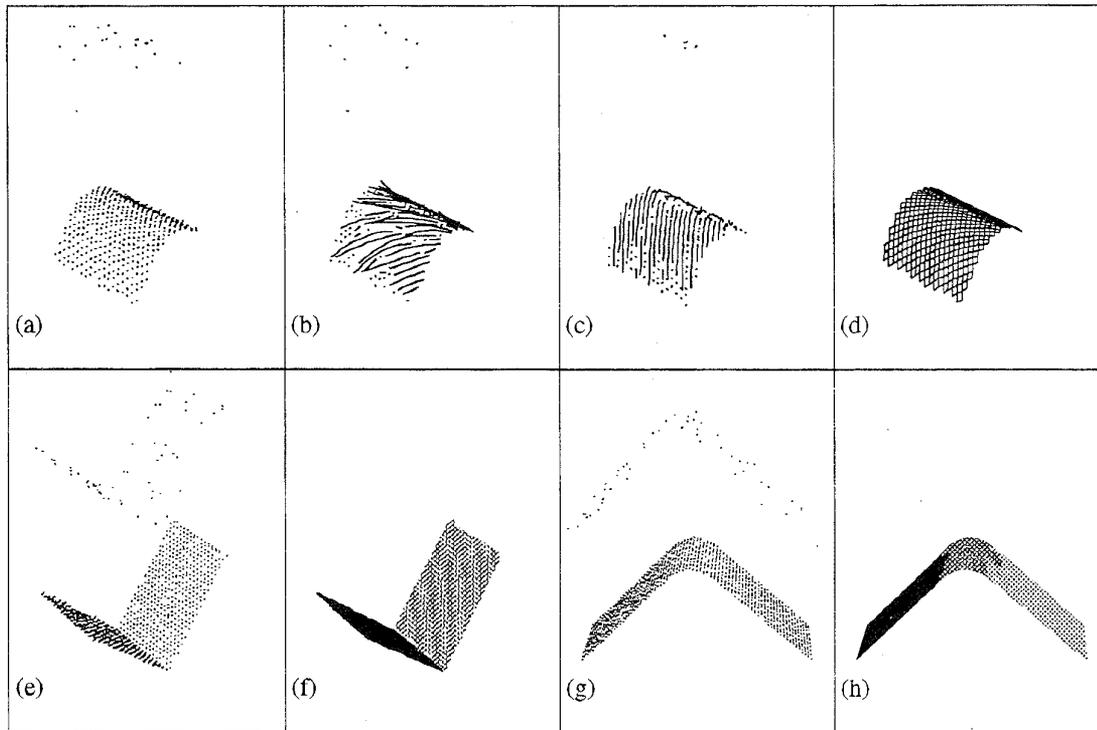
Figure 2. *Synthetic scenes used in RPT testing. Measurements of a second order surface bounded by jump discontinuities appears in (a). Orthogonal sets of strips appear in (b) and (c), and the final segmented surface in (d). The measurements and surfaces of two planar regions separated by a crease discontinuity are shown in (e) and (f). Measurements and surfaces for tests of curvature discontinuities appear in (g) and (h). Note that measurements and strips along the "floor" of the synthetic scenes were not plotted.*

Table 1. *Results from Monte Carlo analyses of synthetic scenes.*

Test Type	Std. Dev. of Gaussian Noise (inches)	Trials OK / Total (% / Num)	Volumetric Error (%)	Duration On-Line (Sec)	Duration Off-Line (Sec)
Jump	0.010	90 / 100	0.3	0.012	3.2
	0.020	89 / 100	0.3	0.012	3.4
	0.030	95 / 100	0.3	0.012	3.2
Crease	0.010	99 / 100	2.4	0.012	5.9
	0.020	97 / 100	3.7	0.013	5.9
	0.030	89 / 100	2.4	0.012	5.9
Curv.	0.010	98 / 100	10.1	0.012	6.5
	0.020	93 / 100	14.2	0.012	6.7
	0.030	90 / 100	32.1	0.012	6.7

Table 2. *Results from RPT processing of real scenes.*

Test Type	Trials OK / Total (% / Num)	Volumetric Error (%)	Duration On-Line (Sec)	Duration Off-Line (Sec)
Jump	94 / 50	1.8	0.012	3.9
Crease	86 / 50	7.2	0.012	5.9
Curv.	76 / 50	54.0	0.013	8.2

of a few percent and a stability measure above 90% under a wide variety of conditions and in many hundereds of tests.

The novelty of the RPT approach is an adaptive kernel that is based on Kalman filtering. This allows the RPT kernel to become arbitrarily large - provided consistency of the underlying surface curvature estimate is maintained. The large kernel size helps to improve curvature estimation. The use of Kalman filtering provides a computationally efficient means of modeling second order curvature.

The ability to sense objects as they move continuously past a structured light sensor is an important capability for real applications as it eliminates the need for any stop and go motion to support sensing. This objective supports a wealth of application areas in industry [23, 7]. Future efforts in the area of geographically-distributed advanced manufacturing are of interest. Because of the speed of RPT processing, applications of RPT in Active Vision systems [17, 6, 21] are also of interest. Future extensions to PRIME and RPT include integration with a novel graph-matching technique for object recognition [10].

References

[1] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, New York, 1989.

[2] P. J. Besl and R. C. Jain. Three-dimensional object recognition. *ACM Computing Surveys*, 17(1):75–145, 1985.

[3] R. C. Bolles and M. A. Fischler. A ransac-based approach to model fitting and its application to finding cylinders in range data. In *Proceedings of the 7th International Joint Conf. on Artificial Intelligence*, pages 637–643, Vancouver, B.C., Canada, Aug 24-28 1981. IJCAI.

[4] J. A. Borrie. *Stochastic Systems for Engineers: Modeling, Estimation and Control*. Prentice Hall, New York, 1992.

[5] K. L. Boyer, M. J. Mirza, and G. Ganguly. The robust sequential estimator: A general approach and its application to surface organization in range data. *IEEE Trans. on Pattern Anal. Machine Intell.*, 16(10):987–1001, 1994.

[6] A. K. Dalmia and M. M. Trivedi. Integrating depth from motion and binocular stereo approaches. *Machine Vision and Applications*, To appear 1996.

[7] F. W. DePiero and R. L. Kress. Design and in situ calibration of a structured light sensor. In *Proceedings of Intl. Conf. on Intelligent Robotics and Systems*, pages 513–518, Pittsburgh, PA, August 5-9 1995. IEEE/RSJ.

[8] F. W. DePiero and M. M. Trivedi. Range profile tracking (RPT): A profile-based approach for robust and real-time range image segmentation. *Under Revision for IEEE Pattern Analysis and Machine Intelligence*, Resubmitting in 1996.

[9] F. W. DePiero and M. M. Trivedi. 3-d computer vision using structured light: Design, calibration and implementation issues. *Advances in Computers*, 43, To Appear in 1996.

[10] F. W. DePiero, M. M. Trivedi, and S. Serbin. Graph matching using a direct classification of node attendance. *Pattern Recognition*, 29(6), 1996.

[11] R. O. Duda, D. Nitzan, and P. Barnett. Use of range and reflectance data to find planar surface regions. *IEEE Trans. on Pattern Anal. Machine Intell.*, 1(3):254–271, 1979.

[12] O. D. Faugeras. *Three-Dimensional Computer Vision, A Geometric Viewpoint*. MIT Press, London, 1993.

[13] A. K. Jain and S. G. Nadabar. *Markov Random Fields Theory and Application*, chapter Range Image Segmentation Using MRF Models, pages 543–572. Academic Press, Boston, 1993.

[14] X. Jiang and H. Bunke. Fast segmentation of range images into planar regions by scan line grouping. *Machine Vision and Applications*, 7(2):115–122, 1994.

[15] R. E. Kalman. A new approach to linear filtering and prediction problems. *Trans ASME J. Basic Eng.*, 83:35–45, 1960.

[16] Y. Mainguy, J. B. Birch, and L. T. Watson. A robust variable order facet model for image data. *Machine Vision and Applications*, 8:141–162, 1995.

[17] S. B. Marapane and M. M. Trivedi. Experiments in active vision with real and virtual robot heads. *International Journal of Applied Intelligence, Special issue on Sensor Fusion*, 5(3):237–250, July 1995.

[18] G. Roth and M. D. Levine. A genetic algorithm for primative extraction. In *Proc. 4th Int'l Conf. on Genetic Algorithms*, pages 487–494, 1991.

[19] S. S. Sinha and R. Jain. *Handbook of Pattern Recognition and Image Processing*. Wiley, New York, 1994.

[20] I. M. Sobol. *A Primer for the Monte Carlo Method*. CRC Press, Boca Raton, 1994.

[21] M. M. Trivedi and C. Chen. *Advances in Computers*, volume 32, chapter Sensor-Driven Intelligent Robotics, pages 105–148. Academic Press, Boston, 1988.

[22] E. Trucco and R. B. Fisher. Experiments in curvature-based segmentation of range data. *IEEE Trans. on Pattern Anal. Machine Intell.*, 17(2):177–182, 1995.

[23] M. R. Ward, D. P. Rheaume, S. W. Holland, and J. H. Dunseth. Production plant consight installations. Technical Report GMR-4156, General Motors Research Laboratories, Warren, Michigan, August, 1982.

[24] X. Yu, T. D. Bui, and A. Krzyzak. Robust estimation for range image segmentation and reconstruction. *IEEE Trans. on Pattern Anal. Machine Intell.*, 16(5):530–538, 1994.