Edge-aware Normal Estimation by Rotated Bilateral Sampling

Viktor Kovacs, Gabor Tevesz
Budapest University of Technology and Economics
Department of Automation and Applied Informatics
Magyar Tudosok krt. 2. QB207
1117 Budapest, Hungary
{viktor.kovacs},{gabor.tevesz}@aut.bme.hu

ABSTRACT

In this paper we deal with edge preserving surface normal estimation and crease edge detection in discretized range images. Such range images consist of few discrete quantization levels due to the data acquisition method (short base distance stereo), or when the distance variation of the examined surface is low, compared to the disparity quantization levels. We propose a method for normal estimation and crease edge detection using iso-range curves and rotated bilateral filter based sampling. Iso-range curves are used to extract sparse, but reliable range image points. Samples are first selected by a rotated weight matrix and a plane is fitted on such samples. Simple statistics are gathered during the rotation of the weight matrix, in order to find the best fitting plane and extract crease edge measure. Such information may be used for further range image processing: segmentation, mapping, localization, object detection, recognition etc. Results are shown for both synthetic and real range images. It was shown that applying the proposed method resulted in more accurate normal estimations, crease edges were not smoothed and crease edges were successfully detected.

Keywords

range image, normal estimation, edge detection, plane fitting, short base distance

1 INTRODUCTION

The apparatus for acquisition and processing of 3D geometry data became affordable and compact thus appearing in a wide range of applications (mobile robotics, photogrammetry etc.). Several methods and implementations are available for 3D sensing, each balancing with different features. The optimal must be chosen for each application (cost, precision, range etc).

Time of flight (ToF) based methods provide the most accurate results even at long range at high costs. Conventional stereo is widely used to reconstruct 3D geometric data due to the low development price. Stereo triangulation is based on disparity estimation between the two viewpoints. Surface texture, geometry and lighting affects the disparity estimation and thus the reconstructed geometry. Feature points may be used for more accurate matching between viewpoints but it results in a sparse disparity map. In case of homogeneous texture it might be impossible to estimate the geometry. For such reasons structured light based methods are also used: one camera is substituted by a calibrated projector that emits a known or a series of known light patterns thus homogeneous surfaces can be textured in this way. Disadvantages are short range and non-passive operation: it needs to emit enough light that could be detected. In some cases active operation is not admissible.

Stereo camera based methods provide a dense disparity map based on similarity (normalized cross correlation, sum of differences etc). Dense maps may be transformed to a range image by associating a range value based on the disparity measure and known optical properties. Usually these depth maps are stored as range images, where pixel intensity encodes the depth (Z) coordinate value in order to keep beneficial properties of such images: regular sampling, vicinity information, simple surface and triangle mesh generation etc.

In this paper we deal with range images to utilize specific features and errors associated with structured light based image acquisition methods. The research aims to provide a set of methods to handle range images that were acquired using short baseline distance and the disparity map was estimated using traditional (SAD: sum of absolute differences or NCC: normalized cross correlation) methods. Due to the short baseline distance

disparity values are also small and quantized in image space. Such quantization leads to the discretization of range values as well:

$$z_i = Bf\frac{1}{d_i}, \tag{1}$$

where z_i is the calculated depth value from disparity d_i using B baseline distance and f focal length. When the disparity map was quantized (to pixels) the reconstructed depth map would be quantized as well, and would be inversely proportional to the disparity. Range images $Z(u,v),\ 0 \leq u < n, 0 \leq v < m$ suffering from strong quantization consist of few discrete range values:

$$|\{Z(u,v)\}| << nm. \tag{2}$$

Such strong quantization noise cannot be considered random, and has significant effects on algorithms used for low level range image processing. Hough transform or RANSAC based model (ie. plane) fitting may easily find better but improper explanations of surface parts described by discrete range values. Such false surface regions show as quantization levels, planar patches perpendicular to the z axis.

A framework is presented for processing such layered range images, specifically edge-aware normal estimation and edge detection. Our method consists of two major steps: a preprocessing step, where low level features are extracted first and postprocessing, where such features are utilized for further analysis.

The preprocessing step involves layer separation, filtering, skeleton extraction. First the quantized range image is broken into binary images, each image describing a layer. Each binary layer is filtered in order to reduce noise, finally skeletons are extracted using thinning to describe the layers. It is assumed that such skeletons estimate iso-range curves on surfaces where the quantization error is minimal. Due to perspective projection such centerline estimation in image space is biased. However our results show that such error is not significant in practical cases, only at extreme cases where the surface is steep and the quantization step is significantly large.

During postprocessing such skeletons are used in plane, edge and corner detection. Such features can be used in registration problems, mapping, localization, or object detection.

This paper focuses on normal estimation in such sparse range image point sets while keeping both jump and crease edges. Surface normals provide low level features used in subsequent processing steps in range image understanding. The naive approach of local surface estimation by plane fitting on a local neighborhood of pixels lead to significant errors in discretized range images. Based on the layer widths (surface orientation and quantization), the local neighborhood size and the weighting, naive estimation would give significantly different results. Using a small neighborhood for sampling would mostly result in sampling from one layer, thus providing a normal parallel to the depth direction. Near layer edges a perturbation would be observed of the incorrectly estimated surface normal. In order to improve estimation, data uncertainty must be estimated. As quantization is not random, spatial information may be introduced for uncertainty estimation. In the proposed method as a simplification, it is assumed that the centerlines of layers carry reliable depth information, these shall be used for model fitting, other layer pixels are ignored.

In this paper the rotated bilateral sampling method is proposed, by which edge aware fitting of models, in the given example local planar segments for normal estimation are possible. With side information of edges, the estimation process may be sped up.

The paper is organized as follows: in section 2 related work is presented, in section 3 the proposed algorithm is shown. Results of simulation and real images are presented in section 4, finally in section 5 results are discussed and conclusions are drawn.

2 RELATED WORK

Bilateral filtering was introduced in [17]. Such filters combine closeness (spatial) and similarity (value) filtering in one general filter:

$$\mathbf{h}(\mathbf{x}) = k^{-1}(\mathbf{x}) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{f}(\xi)c(\xi,\mathbf{x})s(\mathbf{f}(\xi),\mathbf{f}(\mathbf{x}))d\xi \tag{3}$$

$$k_r(\mathbf{x}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(\xi,\mathbf{x})s(\mathbf{f}(\xi),\mathbf{f}(\mathbf{x}))d\xi \tag{4}$$

where $\mathbf{f}(\mathbf{x})$ denote the input image, $\mathbf{h}(\mathbf{x})$ the filtered output. Functions $c(\xi,\mathbf{x})$ and $s(\mathbf{f}(\xi),\mathbf{f}(\mathbf{x}))$ define the closeness and similarity functions. Weights $k(\mathbf{x})$ are applied in order to preserve units.

On smooth regions where the variance of the values is low, it acts as a standard domain filter. On regions containing a sharp edge, where values differ significantly, values would be taken only from samples similar to the center value. As the kernel is not spatially invariant and is based on the original image contents, FFT and other methods are not applicable to speed up calculation.

Such filters are exceptionally popular for range image processing as object boundaries are not blurred with background information. Typical closeness and similarity filters are Gaussian. In the proposed method the filters are also Gaussian.

A hybrid solution with weighted median filtering is presented in [18] for range image upscaling using high resolution intensity images. Trilateral filters also take the

gradient into consideration, in [14] such filters are used for upscaling range images.

Bilateral filters were also proposed as edge detectors in [8]. A high-pass closeness (domain) kernel is combined with an inverted Gaussian similarity (range) kernel.

Several types of edges can be differentiated in range images. Step or jump edges show the most resemblance to intensity image counterparts, they appear as depth discontinuities at object boundaries. Crease edges show as a significant change in normal direction. Smooth edges are identified by abrupt change of the surface curvature while the normal changes gradually. Most papers do not deal with smooth edges. According to the data acquisition circumstances, false edges may appear around regions of unknown depth values: using stereo triangulation occlusion may happen or distances may be larger than what the rangefinder can handle. Such edges must be omitted or handled accordingly. If layered range images were handled as intensity images an other edge type could be identified. These edges appear between quantization levels but do not represent any type of real edge. Such edges are usually noisy thus cannot be utilized directly, this is a reason skeletons were introduced instead of layer edges.

In [4] the bilateral grid was introduced for edge-aware algorithms. The method involves transforming the image to a higher dimensional grid along the similarity axis, such that pixels representing different patches (separated by an edge) are grouped into different grid cells. This method can also be used to re-express the bilateral filtering problem as a linear filter in a higher dimensional space. Sampling rate over the spatial domain controls the smoothing, sampling rate of the range axis defines the degree of edge preservation. Bilateral filtering can be expressed as 3D convolution between grid cells. The division by the weights are delayed, and data are represented by a variety of homogeneous coordinates. The dual operation of the grid generation is division (by the homogeneous coordinate) and slicing. Our representation is a specialized form of the bilateral grid, where the spatial sampling rates are 1, the range sampling corresponds to the available range layers, thus one range pixel is associated to each grid cell.

Surface normals provide basic features for higher level range image understanding such as segmentation, mapping, navigation, object recognition or detection. Authors of [9] compare several methods (different variations of singular value decomposition and principal component analysis, triangle based averaging) for surface normal estimation evaluating the tradeoff between precision and speed. Joint surface and surface normal reconstruction is shown in [19] using statistical methods for improved robustness. In [1] normal estimation is optimized for reduced computational demand transforming range images to spherical coordinate system, giving spherical range images. Normals may be directly extracted from such representation. In [13] directional joint bilateral filters are introduced to take edge direction into account during filtering. State of the art methods [6] involve integral images for surface normal estimation in point clouds. Integral images simplify summing over a rectangular region as only the values at the corners of the rectangle is needed.

The sampling matrix we proposed shows some resemblance to the one used in the Kuwahara filter and it's modifications [12]. The original filter uses four square regions around the sampled point where mean of the subregion is applied to the center pixel where the standard deviation is the lowest. Generalizations involve rotated circular and elliptical filter kernels.

Several methods are used for fitting planar surfaces to sample data. Hough-trasformation may be extended to 3D for plane detection [7, 2]. As the accumulator space has a higher dimensionality accumulator space design must be made carefully [3]. Model fitting in noisy data and high number of outliers are usually done by using a variation of the RANSAC algorithm [5]. Hybrid methods were also developed to fuse advantages of the methods. In [16] both Hough and RANSAC based methods are used at multiple resolutions.

3 PROPOSED METHOD

In this section we present a new method for local surface normal estimation in discretized range images. Estimated normals can be used for further analysis of the images, such as segmentation [11], edge detection [10], smoothing etc. while it also provides basic information for higher level semantic analysis such as object or landmark recognition.

3.1 Skeletonization

First the layered range image is broken into binary images representing each quantization value. Next these binary images are filtered in order to reduce noise that might be present near the edges of layers. Small detached patches are removed from the layers using the connected components algorithm. Morphological operations are applied in the resulting binary image: dilation and erosion in order to smooth the transitions between layers.

Next a thinning algorithm is applied for skeleton extractions. Skeletons are one pixel width lines that represent the centerline of binary images. There are numerous thinning algorithms, [15] was implemented, that produces few side branches. The resulting skeletons are broken into skeleton segments: such segments consist of skeleton pixels between junctions or endpoints. Skeletons are pruned by removing remaining unwanted short segments.

In our framework we utilize these iso-range skeletons for further processing. It is assumed that these iso-range lines describe surface crossections that have minimal quantization noise on linear surfaces. Centerline estimation error due to projection is minimal in practical cases. Figure 1 shows the iso-range skeletons of a sample simulated scene.

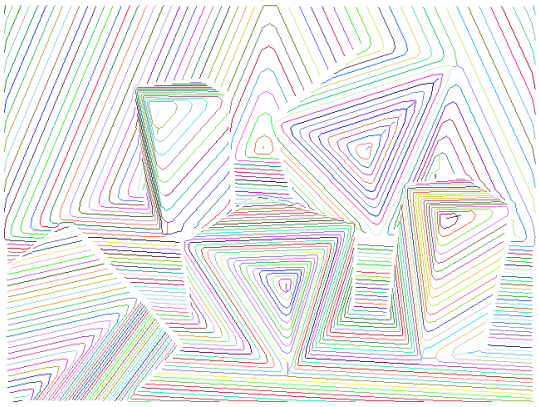

Figure 1: *Iso-range skeletons extracted from each layer of the discretized range image.*

3.2 Surface normal estimation

Normal estimation is one of the most fundamental steps in range image processing. Normals can be used for plane segmentation, edge detection etc.

Previously we proposed a method [10] for surface normal estimation by a variation of the forward difference method but adapted for 3D iso-range curves. The method can also be extended to utilize multiple layer information to estimate the gradient. Figure 2 illustrates the process.

First the skeleton tangential orientation v is estimated by sampling $n_{FitLine}$ number of skeleton points in image space. The total least squares method is applied to find the best fitting direction: the eigenvector related to the larger eigenvalue of the corvariance matrix is evaluated. This tangential direction is also the tangential of the skeleton in 3D, and it specifies a plane P. In order to estimate the binormal $\mathbf{p}^*$ must be identified on an adjacent layer (l_{i+1}), where the skeleton of the adjacent layer crosses P plane. By knowing the tangential and the binormal, the surface normal $\mathbf{n} = \mathbf{b} \times \mathbf{v}$ can be calculated. By identifying several binormals, not only on one adjacent layer, but on several, the normal estimation may be improved for planar regions, but near edges normals may be smoothed. By default this method identified one binormal on the following layer thus being noise sensitive. Estimation failed or was ignored near crease edges, due to the significant change in layer skeleton orientation.

In this paper we present a method that overcomes such problem, providing edge preserving surface normal es-

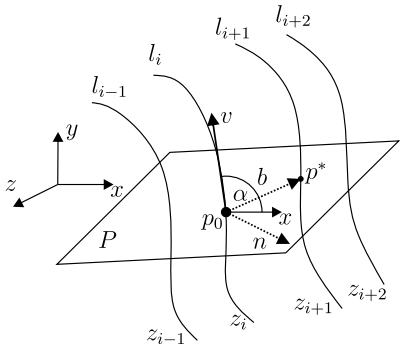

Figure 2: *l_i: layer skeletons, v tangent, b binormal, n normal direction*

timation, using a rotated bilateral filter kernel for sampling and principal component analysis (PCA). PCA is an orthogonal transformation, which maximizes the variance of the samples along the principal directions. The eigenvectors and eigenvalues of the covariance matrix define the principal directions and components.

First a set of sampling matrices are constructed as a function of θ and N.

$$\mathbf{S}_N(\theta)_{i,j} = \begin{cases} e^{-(a^2+b^2)} & \text{if } 0 \le \phi \le \pi \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$
$$i, j \in 1..(2N+1)$$

where

$$a = \frac{\sin(\phi)R}{Ns_2}, \quad b = \frac{\cos(\phi)R}{Ns_1}, \qquad (6)$$

$$R = \sqrt{(i-N-1)^2 + (j-N-1)^2}, \qquad (7)$$

$$\phi = \text{atan2}(j-N-1, i-N-1) + \pi + \theta. \qquad (8)$$

s_1 and s_2 modify the shape and weight falloff of the sampled pixels. In our implementation we set $\theta = i/18\pi$, where $i = 0..35$. The constructed sampling matrices are illustrated in Figure 3. $\mathbf{S}_N(\theta)$ matrices are calculated only once and stored in a look-up table. To modify the behavior around edges or corners the opening angle may be changed by modifying the π constant in (eq. 5).

Such matrices are used for sampling skeleton points in image space for plane fitting. For each skeleton point a sampling matrix is selected by taking skeleton distance into account. Skeleton distance $d_s(\mathbf{p})$ is given during skeletonization by the number of steps after the centerline point is reached:

$$N(\mathbf{p}) = \max(N_{min}, \min(N_{max}, [s_N d_s(\mathbf{p})])) \qquad (9)$$

where N_{min} and N_{max} denote the minimal and maximal size parameter of the sampling matrices, s_N denote the scale multiplier for selecting the size. Such selection of the sampling size enables better adaptation to skeleton (surface information) density.

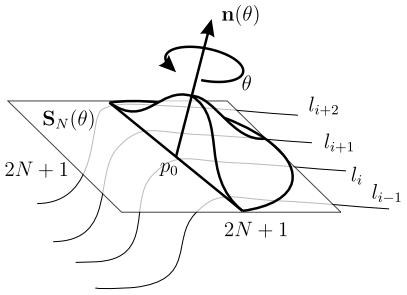

Figure 3: *The constructed sampling matrices are similar to folded bivariate Gaussian distribution pdfs, but rotated by θ around the center point. It serves as defining weights for sampled skeleton pixels but the weights are also modified by a bilateral filter. l_i denote layer skeletons.*

In order to estimate the local surface normal we first estimate the covariance matrix of the sampled skeleton points.

$$\Sigma(\theta) = \frac{0.5\sum_{k=1}^{N} w_k}{\left(\sum_{k=1}^{N} w_k\right)^2 - \sum_{k=1}^{N} w_k^2} \sum_{k=1}^{N} w_k \left(\mathbf{p}_k - \mathbf{p}_0\right)^T \left(\mathbf{p}_k - \mathbf{p}_0\right) \tag{10}$$

where $\mathbf{p}_0$ is the selected center point, $\mathbf{p}_k$ are the sampled skeleton points around the center in N radius. The estimation of the cross-variances relative to $\mathbf{p}_0$ means that the algorithm assumes that there are points mirrored to $\mathbf{p}_0$ on the zero side of $\mathbf{S}(\theta)$. To accept the samples N_s number of points must be sampled from at least N_l number of layers.

Weights w_k are given by the appropriate element of the sampling matrix and a scaled difference in depth (Z):

$$w_k = \mathbf{S}(\theta)_{i,j} \cdot \exp\left(-\frac{(Z(p_k) - Z(p_0))^2}{s_3^2}\right) \tag{11}$$

where $Z(\mathbf{p})$ denotes the range component (z) of a pixel, s_3 defines the weight scale in depth coordinates and the kth sampled pixel corresponds to the i,jth component of $\mathbf{S}(\theta)$.

Next eigenvalue-eigenvector decomposition is applied for $\Sigma(\theta)$:

$$\Sigma(\theta) = \mathbf{V}(\theta)\Lambda(\theta)\mathbf{V}^T(\theta) \tag{12}$$

The eigenvector of the smallest eigenvalue is selected as the normal $\mathbf{n}(\theta)$. The smallest eigenvalue corresponds to the least significant direction, which is the normal of the best fitting plane at θ direction:

$$i^*(\theta) = \arg\min_{i=1..3}\left(|\Lambda_{i,i}(\theta)|\right) \tag{13}$$

$$\lambda^*(\theta) = |\Lambda_{i^*(\theta),i^*(\theta)}| \tag{14}$$

$$\mathbf{n}(\theta) = \mathbf{V}(\theta)^{i^*(\theta)} \tag{15}$$

At a given $\mathbf{p}_0$ point the normal $\mathbf{n}^*$ is selected which was estimated with the best fit:

$$\mathbf{n}^* = \mathbf{n}(\arg\min_{\theta}(\lambda^*(\theta)) \tag{16}$$

By using a-priori information about edges (position and orientation), the rotation process may be ignored and the appropriate $\mathbf{S}(\theta_{edge})$ sampling matrix may be used. As mentioned in Section 2 other methods exist for normal estimation, which are based on different error functions. Such methods can be easily integrated with the sampling technique given in this paper. PCA was selected to provide a baseline algorithm.

3.3 Edge detection

Edge detection in range images differs from how edges appear in intensity images. Edge types have been summarized in Section 2.

In our previous research [10] we have shown a method for detection and classification of edges in such discretized range images. The method was also based on skeleton extraction. Jump edges were detected by evaluating pixel local neighborhood for significant depth changes and taking quantization levels into account.

In case the range map was acquired using stereo disparity map, the quantization function is usually not linear, but may be already known or it can be identified from the image itself.

For jump edge detection $Z_{th}(z) = \sigma dZ(z)$ is selected where $dZ(z)$ is the quantization step at given z depth, $\sigma > 1$. Equation

$$\max_{||p-p_0||<r}(Z(p) - Z(p_0)) > Z_{th}(Z(p_0)) \tag{17}$$

is satisfied near jump edges on the foreground surface. To deal with false edges we not only look the adjacent pixels, but in case of missing data the adjacent side of the unknown region.

Crease edge detection was based on abrupt changes of skeleton orientation. Such orientation and distances between skeletons encode the surface normal as the orientation is a projection of the normal to the *xy* plane, distances carry information of the z component. Changes in the orientation means changes in the normal but not vice versa. In order to detect all create edges normal reconstruction is needed.

We also propose a new method for crease edge detection. For surface normal estimation a rotated bilateral sampling was used, resulting in a normal function of rotation $\mathbf{n}(\theta)$. We assume that along planar surfaces the variance of such function is low, but on crease edges the variance increases. An edge measure e is introduced as the mean square error of normals at different rotations:

$$e = \frac{1}{N} \sum_{\theta \in \Theta} \left(\cos^{-1}(\mathbf{n}(\theta) \cdot \overline{\mathbf{n}})\right)^2. \qquad (18)$$

4 RESULTS

In this section results of the proposed algorithm is presented. Both synthetic and real captured data are evaluated [1].

4.1 Synthetic data

Sample scenes were constructed and rendered with ground truth data: surface normals were available directly from the modeling software. Values are usually estimated for skeleton pixels only but all pixels are filled in the image according to the closest skeleton pixel (in image space). This leads to incorrect visualization in many cases near endpoints of skeletons.

A simple bilateral-type of sampling is used in Figure 4. It can be seen that jump edges are correctly kept, but crease edges are smoothed. This is normal as the depth function part of the bilateral weight function still produces high weights because of relatively small depth differences on both sides of the edge. The sampling matrix size was adapting to the local density of the skeletons.

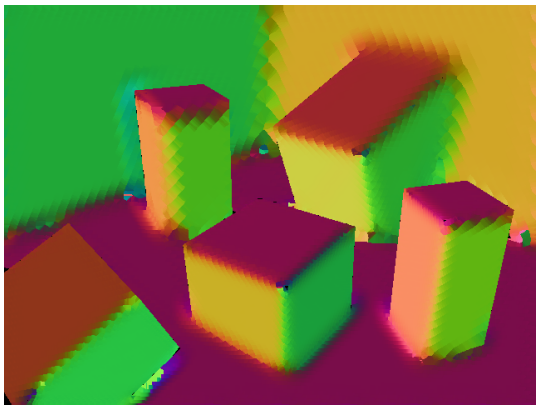

Figure 4: *Normal estimation using PCA and weighting based on distance in images space and depth. Boundaries are mostly kept intact, crease edges are blurred.*

Results using the proposed method is shown in Figure 5. Neither boundary nor crease edges were smoothed. The number of missing normals is very low. Some corners were smoothed.

Normal estimation error distribution is shown in Figure 6. Evaluating only skeleton pixels shows faster error fall-off as edges are not smoothed. The forward difference method shows slower fall-off, while the difference of the error distribution for all pixels or skeleton pixels do not show significant difference. For the proposed method the distribution for all pixels is similar to the compared method but still shows less uncertainty.

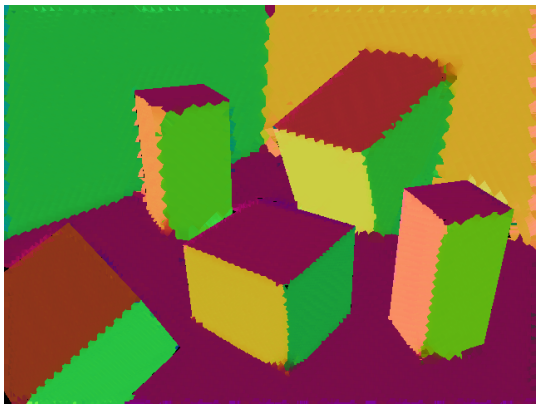

Figure 5: *Normal estimation using the proposed method. Neither boundaries nor crease edges are smoothed. The number of unestimated normals is minimal.*

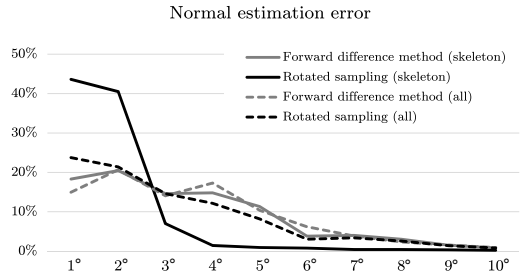

Figure 6: *Normal estimation error distribution. Forward difference method is colored gray, proposed sampling method is black. Errors for only skeletons pixels are shown as solid, for all pixels as dashed lines.*

Identified crease edges are shown in Figure 7. Crease edges are highlighted very well while object boundaries are not highlighted due to the bilateral behavior of the sampling matrix.

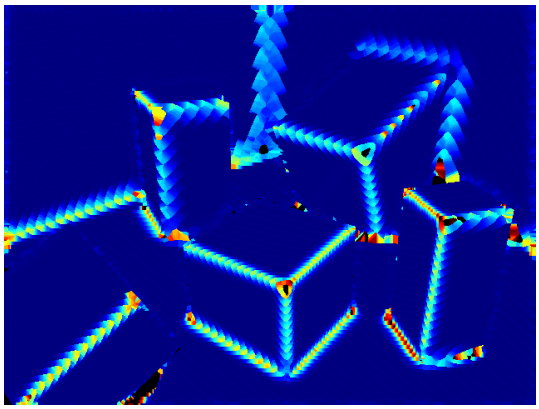

Figure 7: *Crease edge detection using edge measure given in equation (18)*

[1] http://dx.doi.org/10.6084/m9.figshare.1409432

4.2 Captured data

Data were acquired using the Microsoft Kinect sensor. In order to simulate even lower number of range layers the depth resolution was reduced manually. When observing a surface where the variance in depth direction is small, the resulting scene may also contain a low number of layers.

Figure 8 shows the reconstructed normal map of a captured scene using the forward differences method. The algorithm could not estimate normals around object boundaries due to limited skeleton information near borders and abrupt changes in skeleton directions cause incorrect estimates which are rejected. The range image contained only 28 depth layers. Due to noise normal estimation shows significant variance along planar surfaces.

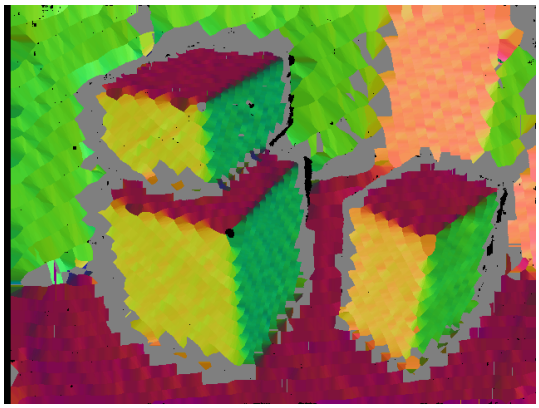

Figure 8: *Normal estimation based on forward difference method. Normal estimation failed around object borders. Crease edges appear smoothed.*

Results of the rotated bilateral sampling algorithm is shown in Figure 9. Edges are preserved and normals show less variance on planar surfaces. Such smoothing is due to the higher number of samples used in fitting compared to the one (or few) skeleton point on adjacent layer(s).

Estimated crease edges are shown in Figure 10. Due to noise and the very low number of range layers some false positive regions appear. Again this evaluation was run only on skeleton pixels but for visualization purposes data were interpolated using the nearest neighbor method. Hence the large curved positive regions.

5 CONCLUSIONS

In this paper we have shown a method for surface normal estimation and crease edge detection. The method is based on a set of sampling matrices that contain weights, and are constructed in advance as a function of size and orientation. Plane fitting is evaluated using adaptive size for the sampling matrix and also incorporating distance weights, similar to bilateral filtering.

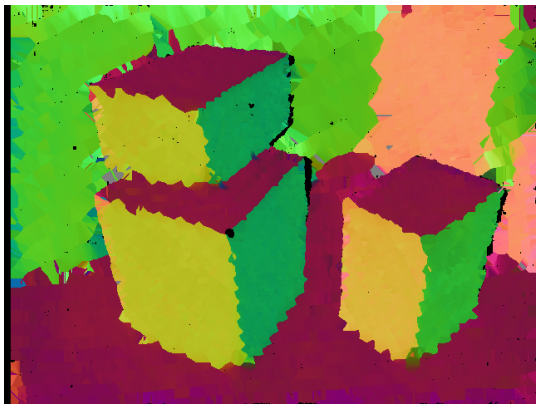

Figure 9: *Normal estimation based on the proposed rotated sampling and fitting. More edge points are preserved.*

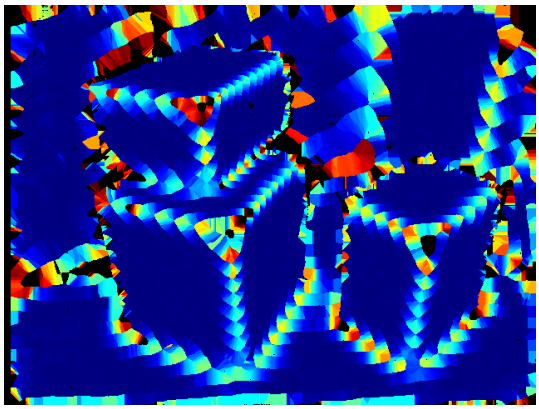

Figure 10: *Crease edge map estimated using the edge measure. Due to noise and the little number of layers false positive regions appear.*

This step produces a surface normal and a fitting error measure for each orientation. Simple statistics are used to select the best fitting plane and to identify a crease edge score. The method can be easily extended to incorporate other fitting methods: such as simple least squares, or apply two-step methods such as RANSAC.

We have presented examples of the output of the algorithm for both simulation and real data. Results show significant improvement compared to bilateral filtering as crease edges are less prone to blurring. We also compared a previously implemented method (forward differences between layers) that was used for normal estimation in heavily quantized range images. Although the proposed algorithm runs slower than the forward difference method, the results are more accurate: crease edges are less blurred and normal variation is lower on planar surfaces and data are estimated for more pixels. A method for crease edge detection was also presented based on the output of the normal estimation algorithm. Future research involves smoothing heavily quantized range images. Such results can successfully be used for further range image processing: segmentation, mapping, localization, object detection, recognition etc.

6 ACKNOWLEDGMENTS

This work was partially supported by the European Union and the European Social Fund through project FuturICT.hu (grant no.: TAMOP-4.2.2.C-11/1/KONV-2012-0013) organized by VIKING Zrt. Balatonfured. This work was partially supported by the Hungarian Government, managed by the National Development Agency, and financed by the Research and Technology Innovation Fund (grant no.: KMR 12-1-2012-0441).

7 REFERENCES

[1] H. Badino, D. Huber, Y. Park, and T. Kanade. Fast and accurate computation of surface normals from range images. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 3084–3091, May 2011.

[2] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nüchter. A data structure for the 3d hough transform for plane detection. In *Proceedings of the 7th IFAC symposium on intelligent autonomous vehicles (IAV 2010), Lecce, Italy*, 2010.

[3] D. Borrmann, J. Elseberg, K. Lingemann, and A. Nuchter. The 3d hough transform for plane detection in point clouds: A review and a new accumulator design. *3D Research*, 2(2), 2011.

[4] Jiawen Chen, Sylvain Paris, and Frédo Durand. Real-time edge-aware image processing with the bilateral grid. *ACM Trans. Graph.*, 26(3), July 2007.

[5] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, June 1981.

[6] S. Holzer, R.B. Rusu, M. Dixon, S. Gedikli, and N. Navab. Adaptive neighborhood selection for real-time surface normal estimation from organized point cloud data using integral images. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pages 2684–2689, Oct 2012.

[7] L. Iocchi, K. Konolige, and M. Bajracharya. Visually realistic mapping of a planar environment with stereo. In *Experimental Robotics VII*, ISER '00, pages 521–532, London, UK, UK, 2001. Springer-Verlag.

[8] A. Jose and C.S. Seelamantula. Bilateral edge detectors. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 1449–1453, May 2013.

[9] K. Klasing, D. Althoff, D. Wollherr, and M. Buss. Comparison of surface normal estimation methods for range sensing applications. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3206–3211, May 2009.

[10] V. Kovacs and G. Tevesz. Edge detection in discretized range images. In *Computational Intelligence and Informatics (CINTI), 2014 IEEE 15th International Symposium on*, pages 203–208, Nov 2014.

[11] V. Kovacs and G. Tevesz. Plane segmentation in discretized range images. In *Workshops on Electrical and Computer Engineering Subfields, Proceedings of*, pages 184–189, Aug 2014.

[12] J. E. Kyprianidis and H. Kang. Image and video abstraction by coherence-enhancing filtering. *Computer Graphics Forum*, 30(2):593–602, 2011.

[13] Anh Vu Le, Seung-Won Jung, and Chee Sun Won. Directional joint bilateral filter for depth images. *Sensors*, 14(7):11362–11378, 2014.

[14] Kai-Han Lo, Y.-C.F. Wang, and Kai-Lung Hua. Joint trilateral filtering for depth map super-resolution. In *Visual Communications and Image Processing (VCIP), 2013*, pages 1–6, Nov 2013.

[15] G. Németh and K. Palágyi. 2d parallel thinning algorithms based on isthmus-preservation. In *ISPA 2011: 7th International Symposium on Image and Signal Processing and Analysis: Dubrovnik, Croatia, 4 - 6 September 2011. IEEE*, pages 585–590, 2011.

[16] B. Oehler, J. Stueckler, J. Welle, D. Schulz, and S. Behnke. Efficient multi-resolution plane segmentation of 3d point clouds. In Sabina Jeschke, Honghai Liu, and Daniel Schilberg, editors, *Intelligent Robotics and Applications*, volume 7102 of *Lecture Notes in Computer Science*, pages 145–156. Springer Berlin Heidelberg, 2011.

[17] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on*, pages 839–846, Jan 1998.

[18] Qingxiong Yang, N. Ahuja, Ruigang Yang, Kar-Han Tan, J. Davis, B. Culbertson, J. Apostolopoulos, and Gang Wang. Fusion of median and bilateral filtering for range image upsampling. *Image Processing, IEEE Transactions on*, 22(12):4841–4852, Dec 2013.

[19] Mincheol Yoon, Yunjin Lee, Seungyong Lee, Ioannis Ivrissimtzis, and Hans-Peter Seidel. Surface and normal ensembles for surface reconstruction. *Computer-Aided Design*, 39(5):408 – 420, 2007.