Noise aware depth denoising for a time-of-flight camera

Jiyoung Jung, Joon-Young Lee, In So Kweon*

A time-of-flight camera provides depth maps of the scene at video frame rate. However, their depth measurements are severely influenced by random noise and systematic bias. Previous approaches on depth denoising are usually variants of adaptive joint bilateral filtering with the help of a color image of the same scene. In this paper, we access to the raw range measurements of the ToF sensor instead of the transformed depth values, and we acquire range error profile for each pixel along the range measurement by capturing a planar scene at different distances. We correct the range bias using plane fitting and then the remaining noise can be assumed to follow a zero-mean Gaussian distribution with variance according to the pixel location and the range measurement. Since the whole process is done beforehand leaving variance information, any kind of depth denoising algorithm assuming zero-mean Gaussian noise can perform well with our noise estimation.

Keywords: Depth denoising, Time-of-flight camera noise modeling

1. Introduction

Recently, depth sensors have become popular as the sensors are getting cheaper and smaller. They are widely used in computer vision community in the fields of human-computer interaction [1], 3D reconstruction [2], and robot navigation [3]. A 3D time-of-flight (ToF) camera is a type of depth sensors which modulates its illumination LEDs and measures the phase and the amplitude of the returned signal with its CCD/CMOS imaging sensor at each pixel. The new generation Kinect is known to be installed with a ToF camera which has greater accuracy compared with the previous Kinect with structured light [4].

Unfortunately, the depth measurement that a ToF camera provides suffer from severe random noise and systematic bias. Therefore depth denoising is essential to improve the performance of further depth related applications. Many of previous approaches on depth denoising depends on the color image of the same scene assuming an RGB-D input. They are usually variants of adaptive joint bilateral filtering resort to the observation that edges in the color image usually coincide with depth discontinuities [5] [6].

In this paper, we propose a noise aware depth denoising for a time-of-flight camera without using any other sensors. We handle the raw range measurements of the ToF camera instead of the transformed depth values for better accuracy. We acquire range error profile along the range measurement for each pixel by capturing a planar scene (e.g. a wall) at different distances. After range bias correction, the remaining zero-mean noise is assume to be Gaussian having a standard deviation according to the location of the pixel and its range measurement.

We calculate a range standard deviatioin map of the scene and apply adaptive bilateral filter to show the effectiveness of the range error profile of a ToF camera and to present its application on noise aware depth denoising.

* Korea Advanced Institute of Science and Technology (KAIST)
291 Daehak-ro, Yuseong-gu, Daejeon, Korea
{jyjung,jylee}@rcv.kaist.ac.kr, iskweon77@kaist.ac.kr

2. Related works

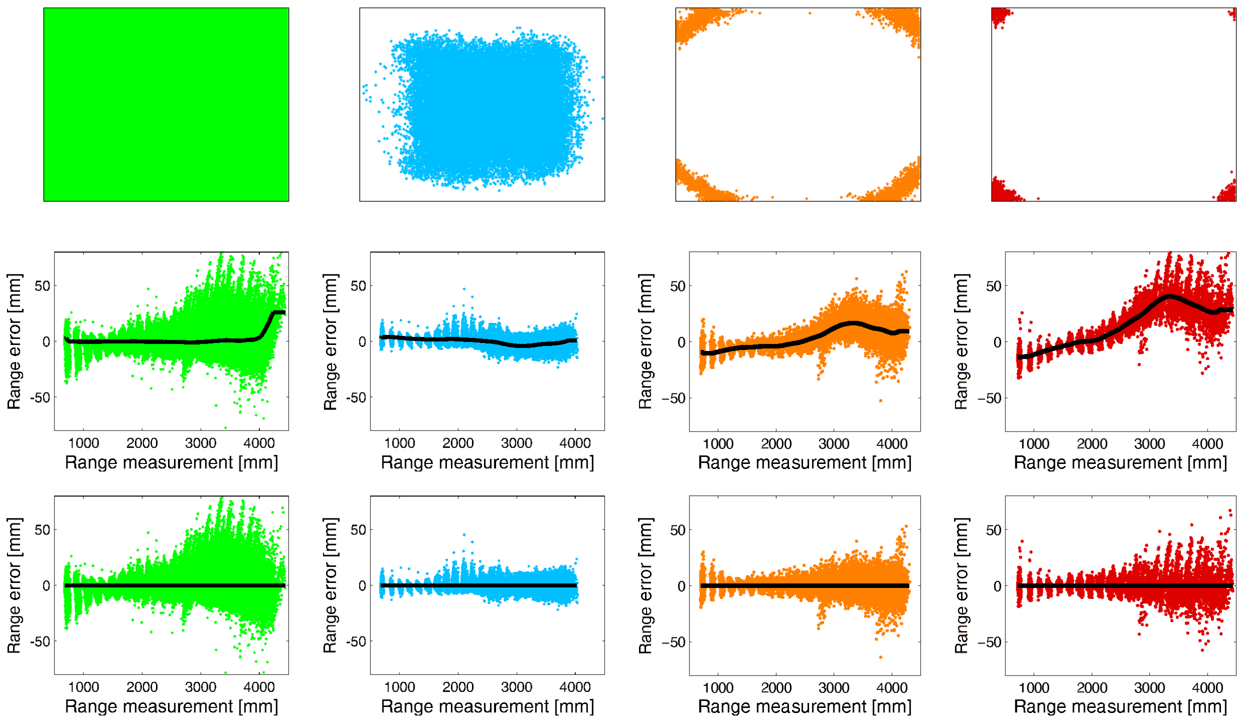
In spite of providing a depth map in video rate, a time-of-flight camera has its limitation on low accuracy and low resolution. There have been several approaches to improve the quality of the depth image of a ToF camera in a modified way of traditional color image enhancement methods.

Schuon *et al.* [7] adopts ideas from traditional color image superresolution to be applied to ToF cameras in order to obtain 3D data of higher X-Y resolution and less noise. Chan *et al.* [5] increase the spatial resolution of the range data using color information of a high resolution video camera. Yeo *et al.* [6] presents an upsampling framework that jointly uses Gaussians of spatial and depth differences of low resolution depth image along with Gaussian of color intensity difference from high resolution 2D color image of the same scene. Frank *et al.* [8] propose an adaptive filter for depth denoising of a ToF camera by adjusting the level of smoothing using the amplitude images as a measure of confidence.

Motivated by Liu *et al.* [9], who estimate an upper bound on the noise level from a single image based on a piecewise smooth image prior model and measured CCD camera responce functions, we have focused on the characteristics of the ToF camera to obtain a hint of its noise behavior from the sensor as well as from the scene.

3. Range error profiles

The range measurements that a time-of-flight sensor provides suffer from random noise and systematic bias. The most crucial error source of a ToF sensor is formed by a systematic wiggling error altering the measured distance by shifting the distance information significantly toward or away from the sensor [10]. The systematic wig-



(a) All pixels (b) Cluster 1 (c) Cluster 2 (d) Cluster 3

Figure 1: (Top) Spatial distribution of the measurements. (Middle row) Range errors and the estimated B-spline functions. (Bottom) Range error after correction. (a) Range bias correction of all the measurement using a single B-spline function. (b-d) Three clusters (Cluster 1,2,3 in Table 1) of the measurements showing similar error profiles and their corrections

gling error arises in the process of distance calculation from the phase difference of the reference signal and the returned signal of a ToF sensor. Due to hardware and cost limitations, the theoretical assumption of a sinusoidal signal shape is generally not suitable in reality. As a result, a range error appears as shown in Fig. 1. Instead of modeling the range bias of all the measurements using a single B-spline function, we focus on the spatial distribution of the range error on the image plane as well.

Range bias correction is important in enhancing the quality of the 3D measurements of a ToF camera because not only it reduces the error in range measurement but also it removes the offset bias from the error to make the error distribution more like a zero-mean gaussian. The bias corrected depth becomes more suitable to depth denoising applications.

To estimate the error of the range measurement, we have captured a wall at different distances, making sure that the wall is captured more than 30 times at each distance. We have averaged those range measurements to obtain a reliable representative frame of the wall at each distance. We model the plane using SVD because a large portion of the 3D measurement after ray correction is already highly accurate. We estimate the range error as the distance between the 3D measurement and the fitted plane along the ray.

Fig. 1 shows the range error along the measurement of the different pixels on the image. The figure in the middle of the column (a) shows the range error profiles of all the pixels. The error is wiggling along the range measurement, showing a wide range of variance. Lindner *et al.* [10] model this error using a single B-spline function. Instead, we have applied k-means clustering to the error data to classify them according to shape of fluctuation. As a result, the wiggling errors are successfully divided into a number of clusters, each of which having a similar shape to be modeled by a single B-spline function. The measurements in each cluster are radially distributed on the image plane, as shown in the top row of Fig. 1. The columns (b-d) show the three out of five ($k = 5$) clusters of pixels' spatial distribution of the measurements and their corresponding range error profiles with the estimated B-spline functions.

Table 1: RMS Error After Range Bias Correction [mm]

Before range bias correction	6.18
All pixels corrected together with a single B-spline function [10]	6.12
Each cluster corrected separately (proposed)	3.80
Cluster 1 (Fig. 1(b))	2.92
Cluster 2 (Fig. 1(c))	5.61
Cluster 3 (Fig. 1(d))	8.95
Cluster 4	3.72
Cluster 5	4.14

The RMS errors before and after bias correction are shown in Table 1. It is shown to be much more effective when range error is corrected each cluster independently than when corrected as a whole. The average RMS error of spatially varying range bias correction is 3.80mm, which is much less than that of a single function correction. Note that the small number of pixels on the image boundary (Cluster 3) show large error compensation due to range bias correction. The range bias clearly has a spatial distribution and it is more effective when corrected by using different B-spline functions.

4. Depth denoising

Bilateral filtering [11] is a simple and popular algorithm for edge-preserving and noise reducing smoothing filter for an intensity or a color image. Generally, when a bilater filter is applied to a grayscaled intensity image, the intensity value at each pixel in an image is replaced by a weighted average of intensity values from nearby pixels. The weight depends both on Euclidean distance of pixels and on the intensity(range) difference. Therefore, the bilateral filter has two parameters, the spatial and range standard deviations, σ_s and σ_r.

In our case, we have applied adaptive bilater filtering to a depth image obtained by a time-of-flight camera. We have applied a fixed value for the spatial standard deviation $\sigma_s = 3$, whereas different values for the range standard deviation. σ_r is determined adaptively for each pixel according to the pixel location in the image and its range measurement.

Given the range error profiles shown in Fig. 1, all the pixels are classified into k clusters purely by its location in the image as shown in Fig. 1(top). Given a frame of range measurements Fig. 2(b,i), σ_r is determined for each range measurement by the remaining variance of the range error profiles after range bias correction shown in Fig. 1(bottom). The resulting σ_r map for the scene is shown in Fig. 2(c,j).

The results of adaptive bilateral filtering using σ_r maps of the scenes are shown in Fig. 2(g,n). Fig. 2(d,e,k,l) are the results of the same algorithm using a fixed value for σ_r for comparison. Since σ_r tend to be large for large range measurements and the pixels on the image boundaries, a small fixed σ_r leaves noise on the image boundaries and where the scene has a large depth and a large fixed σ_r tend to oversmooth the scene in the center. Frank *et al.* [8] tend to preserve details but leaves some noise where the input depth is large. It shows advantage in specular region rejection because it uses the amplitude image of the ToF camera for confidence measure. The proposed method shows effective noise reduction while preserving depth discontinuities without blurring edges. A similar level of edge-preserving smoothing is present on the boundary of the image as well as around its center.

5. Discussion and conclusion

We present noise aware depth denoising application on a depth image of a time-of-flight camera using adaptive bilateral filtering. The standard deviation of the range measurement is adaptively determined according to the scene and the sensor's range error profiles. Given the range error profiles and the range measurement of the scene, the range standard deviation map is calculated without the help of any other sensors. Since the range standard deviation for a ToF camera is tend to be large on the image boundaries and for the large range measurements, the adaptive bilateral filtering is shown to be effective in edge-preserving noise reduction for the entire parts of the image.

We believe that this framework can be an advantageous in further ToF camera related applications since the actual field of view of a ToF camera is even smaller with the pixels on the image corners providing much less accurate range measurements. To enhance the quality of its measurements, the range error profiles can be obtained for each sensors beforehand and any depth denoising algorithms can be used to take advantage of the error profile information

Acknowledgments

This research was supported by the MOTIE (The Ministry of Trade, Industry and Energy), Korea, under the Human Resources Development Program for Convergence Robot Specialists support program supervised by the NIPA(National IT Industry Promotion Agency) (H1502-13-1001)

References

(1) Jamie Shotton, Andrew Fitzgibbon, Mar Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, Andrew Blake:"Real-Time Human Pose Recognition in Parts from a Single Depth Image", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2011)

(2) Yan Cui, Sebastian Schuon, Derek Chan, Sebastian Thrun, Christian Theobalt:"3D Shape Scanning with a Time-of-Flight Camera", Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (2010)

(3) A.Prusak, O. Melnychuk, H. Roth, I. Schiller, R. Koch:"Pose estimation and map building with a Time-Of-Flight-camera for robot navigation", International Journal of Intelligent Systems Technologies and Applications, Vol.5, No.3, pp.355-364 (2008)

(4) Rob Knies:"Collaboration, expertise produce enhanced sensing in Xbox One", The Official Microsoft Blog (2013)

(5) Derek Chan, Hylke Buisman, Christian Theobalt, and Sebastian Thrun:"A Noise-Aware Filter for Real-Time Depth Upsamping", ECCV Workshop on Multicamera and Multimodal Sensor Fusion Algorithms and Applications (2008)

(6) Donghoon Yeo, Ehsan ul haq, Jongdae Kim, Mirza Waqar Baig, and Hyunchul Shin:"Adaptive Bilateral Filtering for Noise Removal in Depth Upsamping", International SoC Design Conference (2010)

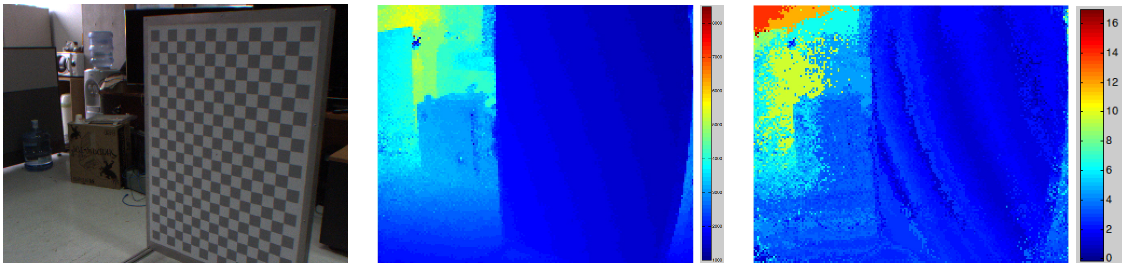
(7) Sebastian Schuon, Christian Theobalt, James Davis, and Sebastian Thrun:"High-Quality Scanning Using Time-Of-Flight Depth Superresolution", IEEE Computer Society Conference on CVPR Workshops (2008)

(8) Mario Frank, Matthias Plaue, Fred A. Hamprecht:"Denoising of continuous-wave time-of-flight depth images using confidence measures", Optical Engineering, Vol.48, No.7 (2009)

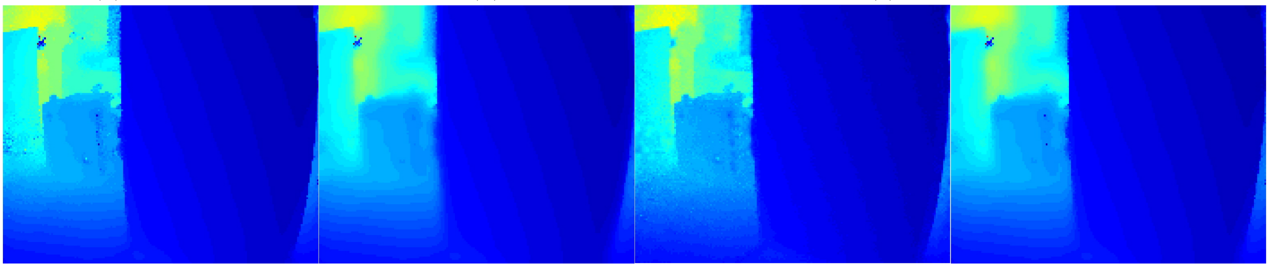
(9) Ce Liu, Willian T. Freeman, Richard Szeliski, Sing Bing Kang:"Noise Estimation from a Single Image", IEEE Conference on Computer Vision and Pattern Recognition (2006)

(10) Marvin Lindner, Ingo Schiller, Andreas Kolb, Reinhard Koch:"Time-of-Flight sensor calibration for accurate range sensing", Computer Vision and Image Understanding, Vol.114, No.12, pp.1318-1328 (2010)

(11) C. Tomasi and R. Manduchi:"Bilateral filtering for gay and color images", In Proc. IEEE International Conference on Computer Vision (1998)

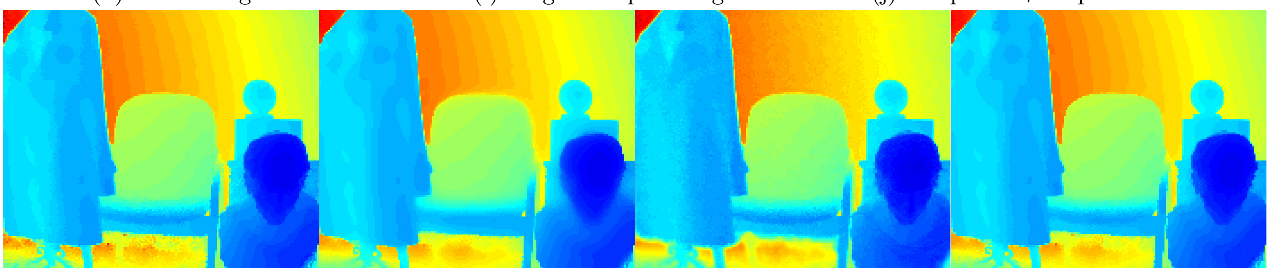


(a) Color image of the scene (b) Original depth image (c) Adaptive σ_r map



(d) σ_r=1 (e) σ_r=5.5 (f) Frank *et al.* [8] (g) Adaptive σ_r

(h) Color image of the scene (i) Original depth image (j) Adaptive σ_r map



(k) σ_r=1 (l) σ_r=5.5 (m) Frank *et al.* [8] (n) Adaptive σ_r

Figure 2: Depth denoising results. (d,e,k,l) General bilater filtering results using the specified range standard deviations. (g,n) The proposed method of adaptive bilteral filtering for a ToF depth image.

Jiyoung Jung received the B.S and the M.S degree in Electrical Engineering from KAIST, Korea in 2008 and 2010 respectively. She is currently a Ph.D. candidate in Robotics and Computer Vision Lab in the Department of Electrical Engineering in KAIST. Her research interests include photometric stereo, multiple sensor fusion, and 3D modeling of indoor and outdoor scenes. She is a recipient of Presidential Science Scholarship by Korea Student Aid Foundation. She is a student member of the IEEE.

Joon-Young Lee received the B.S degree in Electronic Engineering from Yonsei University in 2008, and the M.S degree in Electrical Engineering from KAIST in 2009. He is currently working toward the Ph.D. degree in Electrical Engineering at KAIST. His research interests include photometric methods in computer vision and 3D modeling. He is a student member of the IEEE.

In So Kweon received the B.S. and M.S. degrees in mechanical design and production engineering from Seoul National University, Korea, in 1981 and 1983, respectively, and the PhD degree in robotics from the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1990. He worked for Toshiba R&D Center, Japan, and joined the Department of Automation and Design Engineering, KAIST, Korea in 1992, where he is now a Professor with the Department of Electrical Engineering. He is a recipient of the best student paper runner-up award at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 09). His research interests are in camera and 3-D sensor fusion, color modeling and analysis, visual tracking, and visual SLAM. He was the program co-chair for the Asian Conference on Computer Vision (ACCV 07) and is the general chair for the ACCV 12. He is also on the editorial board of the International Journal of Computer Vision. He is a member of the IEEE and the KROS.