Depth Image Enhancement Using Local Tangent Plane Approximations

Kiyoshi Matsuo[1], Yoshimitsu Aoki[2]
[1]Hokuyo Automatic Co., LTD. [2]Department of Integrated Design Engineering, Keio University.

To enhance both the accuracy and resolution of a noisy low-resolution depth image captured by a consumer RGB-D camera, the pixel-coordinates of the aligned high resolution color image are widely used [1, 3, 4, 5, 6]. However, these global coordinates are not suitable for handling the geometry of the measured shapes. For example, for a depth image, general smoothing/upsampling filters smooth the surfaces to be parallel to the image plane. Because the image plane has no geometric relationship with the measured surfaces, the geometries of the surfaces are not taken into account, and are sometimes corrupted by these filters.

In this paper, we propose a depth image enhancement method using local tangent planes as local coordinates of the measured surfaces. In our method, to reconstruct accurate surfaces, the normal direction components of the measured surfaces are smoothed at each local coordinate using a Gaussian filter. The explicit definition of smoothed point $\tilde{\mathbf{x}}$ of point $\mathbf{x}$ in local coordinates R is

$$\tilde{\mathbf{x}} = \frac{1}{W} \sum_{\mathbf{x}' \in d(R)} w_{\text{Gauss}}(|\mathbf{x}-\mathbf{x}'|_1)(\mathbf{x}',\ \mathbf{n}(R))\mathbf{n}(R) + \mathbf{x}_{\mathcal{T}}, \tag{1}$$

where w_{Gauss} is the Gaussian weight function, $d(R)$ is the shape in region R, $\mathbf{n}(R)$ is the normal vector, $\mathbf{x}_{\mathcal{T}}$ is the tangential component of point $\mathbf{x}$, and W is the total weighted sum. If $\mathbf{t}_1(R)$ and $\mathbf{t}_2(R)$ are the orthonormal tangent vectors of the local tangent plane, then $\mathbf{x}_{\mathcal{T}} = (\mathbf{x},\ \mathbf{t}_1(R))\mathbf{t}_1(R) + (\mathbf{x},\ \mathbf{t}_2(R))\mathbf{t}_2(R)$. We smooth the measured surfaces while preserving the local geometries that are linearly approximated by the local tangent planes. Accurate depth image enhancement is achieved by using the local geometries approximated by the local tangent planes.

Our method is composed of two steps, a step that estimates the local tangent planes of the uncorrupted surfaces from a noisy low-resolution depth image and a step that reconstructs surfaces using the estimated local tangent planes. Because the reconstruction accuracy strongly depends on the approximation accuracy of local tangents, we divide the local tangent estimation process into three additional steps. We first calculate local tangent planes by applying principle component analysis (PCA) to the shapes on each superpixel [2]. Then, we correct the local tangents using neighboring colors and their positional relationships to reduce the influence of outliers caused by measurement noise. For this correction, we introduce two types of PCA, a color heuristic PCA that measures the distribution of 3D points by color similarity and a PCA that is applied to only the centers of the neighboring local tangents. Finally, we connect linearly connectable local tangent planes to detect planar regions. In each planar region, local surface geometries are well approximated by the connected local tangent planes. A conceptual diagram of the calculation of local coordinates is shown in Figure 1. After this local coordinates estimation, in the reconstruction step, we first reconstruct coarse surfaces using segmentation defined by the contact relations between local planar regions, similarly to [5]. We then refine the coarse surfaces by smoothing the normal direction components on each local planar region.

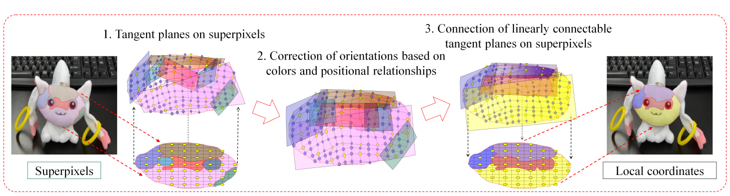

Figure 1: Calculation of local coordinates that approximate uncorrupted surfaces from a noisy low-resolution depth image.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.

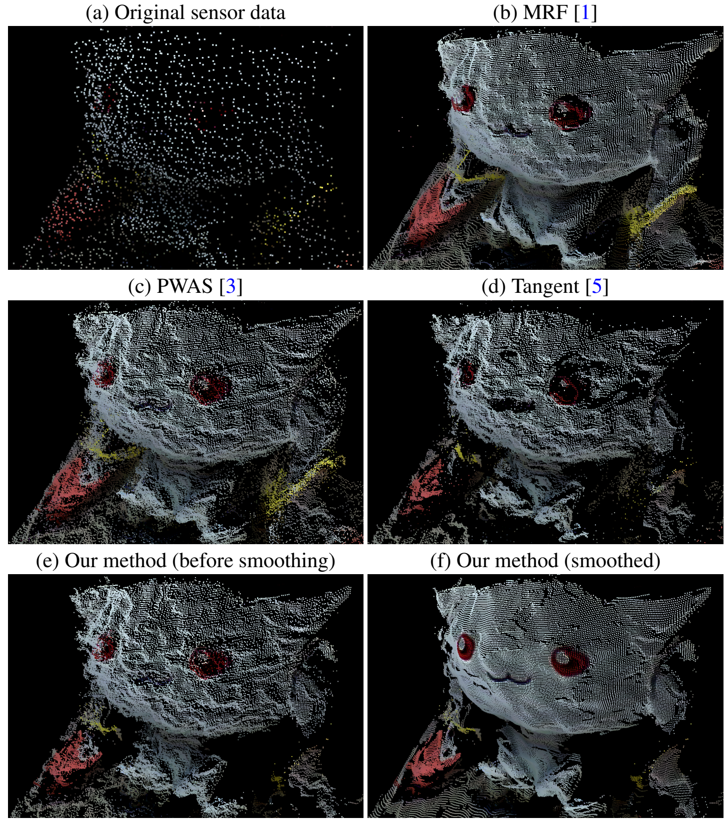

Figure 2: 3D point clouds of a doll head that visualizes the difference between the enhancement results of several methods. The granular noise was reduced by smoothing the normal components on local tangents.

We applied our method to real images captured by a consumer RGB-D camera as well as some images from a well-known dataset. In the paper, both quantitative and qualitative evaluations are described. Figure 2 shows the measured real sensor data and enhanced data as 3D point clouds. Compared with the existing methods that use the global coordinates of the image plane, the interior regions of surfaces were reconstructed more accurately by our method. Our experimental results show the effectiveness of our local tangent-based depth enhancement.

[1] James Diebel and Sebastian Thrun. An application of markov random fields to range sensing. In *Advances in Neural Information Processing Systems 18*, pages 291–298. 2006.

[2] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.

[3] F. Garcia, B. Mirbach, B. Ottersten, F. Grandidier, and A. Cuesta. Pixel weighted average strategy for depth sensor data fusion. In *Proc. ICIP*, pages 2805 –2808, 2010.

[4] Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matt Uyttendaele. Joint bilateral upsampling. *ACM Trans. Graph.*, 26(3), July 2007.

[5] K. Matsuo and Y. Aoki. Depth interpolation via smooth surface segmentation using tangent planes based on the superpixels of a color image. In *Proc. ICCVW*, pages 29–36, 2013.

[6] Jaesik Park, Hyeongwoo Kim, Yu-Wing Tai, M.S. Brown, and Inso Kweon. High quality depth map upsampling for 3d-tof cameras. In *Proc. ICCV*, pages 1623 –1630, 2011.