Sparse- and Noisy-to-Dense Depth Map Upsampling Based on Mesh and Colour Consistency

Hanshin Lim
hslim@etri.re.kr
Junseok Lee
leejs@etri.re.kr

VR/AR Research Group
Electronics and Telecommunications
Research Institute
Daejeon, Republic of Korea

Abstract

The paper presents a mesh- and colour-consistency-based depth map upsampling method from sparse depth information with various sampling structures under noisy conditions. In addition, we applied the proposed method to generate spatially consistent depth maps and a dense 3D point cloud from a sparse and noisy initial 3D point cloud. In the proposed method, triangulation is first performed on an image plane, whose sparse depth information is contaminated by noise and have irregular sampling structures. Then, an iterative discontinuity-preserving noise reduction process is enforced in the triangulation. After the noise reduction, a depth assignment method based on colour consistency and triangulation is used to generate a dense depth map. The experiment results show that the proposed method can provide a more accurate depth map than previous sparse-to-dense depth map upsampling methods. Furthermore, the application results verify the applicability and potential of the proposed method to various areas with inherent sparsity and irregularity of the input depth information, such as multi-view stereo.

1 Introduction

In recent years, the production and release of 3D and Virtual Reality (VR) content have dramatically increased, and there has been significant progress in 3D image acquisition, processing, display techniques, and related devices [1, 2, 3]. In 3D research and application areas, one of the most fundamental problems is to obtain a high-quality dense depth map because the quality of the depth information can be crucial in determining the qualities of 3D reconstruction results and synthesised intermediate views, and the 3D pose recognition performance [4].

One approach to obtain a high-quality depth map is to apply an upsampling process to the sparsely sampled depth information, which is obtained using a type of depth camera or computed from the geometric relations of the features or depth cues [5, 6, 7]. In this approach, a key issue is to preserve the depth discontinuities while restoring the planar depth regions because the two regions tend to be clearly separated in most depth maps, and the qualities of the application results such as the intermediate views and 3D reconstruction highly depend on whether the regions are properly addressed [8]. Furthermore, the initial

sparse depth information has unexpectedly irregular sampling structures and various degrees of noise in most applications such as depth map coding, 2D-to-3D video conversion, and multi-view stereo [9], which is an important factor to consider when a sparse-to-dense depth upsampling scheme is developed.

The main goal of the present paper is to reconstruct a high-quality dense depth map from sparse and noisy depth information with irregular sampling structures. To attain the goal, this paper presents a mesh- and colour-consistency-based sparse-to-dense depth map upsampling method, which is applicable to various sampling structures with noise. In addition, to verify the applicability of the proposed method, we apply the proposed method to multi-view stereo, where the initial sparse depth information can be assumed to have a totally irregular sampling structure and noise contamination.

2 Previous Work

2.1 Overview of Sparse-to-Dense Depth Map Upsampling Methods

Two well-known approaches to generate a dense depth map from sparse depth information are the use of colour consistency and Markov Random Field (MRF) model [5, 6, 10, 11, 12, 13, 14, 15]. In colour-consistency-based approaches, depth values are assigned by the guidance of the spatial colour similarity of the corresponding high-resolution colour image. A joint bilateral filter combines the spatial and range kernels to consider the spatial distance and colour consistency between the target and the sparsely sampled pixels [6]. In MRF-based approaches, the design of the smoothness term is a notably crucial factor that determines the algorithm performance in noisy conditions [6].

An important issue in the design of a depth map upsampling method is the consideration of the sampling patterns. From the sparse and irregularly sampled depth information, a dense depth map can be generated using conventional triangular warping [16]. However, the conventional triangular warping has a smoothed and inaccurate representation of the depth values around the object boundaries, which degrades the quality of the application results. To address this problem, a Depth Map Rasterization method performs a simple and efficient depth value assignment using the colour consistency for each triangular patch with depth discontinuities [17]. Although the experiment results show that this method can be used for various sampling structures, the method is not robust to noisy conditions because it does not assume the presence of noise when it estimates and assigns the depth values.

2.2 Applications of Sparse-to-Dense Depth Map Upsampling Methods

The quality of the reconstructed 3D models or synthesised intermediate-view images is strongly affected by the quality of the reconstructed dense depth maps. Therefore, as briefly mentioned in the Introduction, upsampling to high-quality dense depth maps from the sparse depth information has been a key technology in many application areas such as multi-view stereo [18, 19, 20] and depth map coding [8, 21, 22]. In multi-view stereo, the sparse depth information of each image plane is obtained from a sparse 3D point cloud and the corresponding camera parameters. A dense depth map for each image plane is generated from the sparse depth information using various depth estimation schemes such as feature matching and depth propagation [19, 20, 23]. In depth map coding, the original depth maps can be encoded after the downsampling process to reduce the bitrate. Then, the decoded depth in-

Figure 1: Overall procedure of the proposed mesh- and colour-consistency-based sparse-to-dense depth map upsampling method.

formation is scaled to have the same resolution as the original one by enforcing upsampling processes, for which various sparse-to-dense depth map upsampling techniques have been applied [24, 25]. The reconstructed dense depth maps can be used to convert single-view videos into stereoscopic or multi-view videos for a 3D display [24, 25, 26]. In addition, in the area of object recognition and tracking, sparse-to-dense depth map upsampling techniques have been applied to obtain a high-quality dense depth map from low-resolution and noisy depth information acquired from depth sensors [4, 27].

3 Proposed Method

The basic concept of the proposed method is to combine a discontinuity-preserving noise reduction process suitable for a mesh structure with the Depth Map Rasterization method [12]. In the Depth Map Rasterization, the edges in a triangulation are classified into continuous and discontinuous edges, and based on the information, Colour Representative Blocks are selected for each node to compare the colour consistency between the node and a pixel. A depth value for each pixel is estimated by planar extrapolation from a node with the best colour similarity. As mentioned, although the Depth Map Rasterization method shows promising results for various sampling structures, it is vulnerable to noisy depth inputs since the classification of continuous and discontinuous edges based on plane fitting from the neighbour nodes rapidly worsens in noisy conditions, thereby giving rise to erroneous plane parameters and depth value assignment. To resolve this problem, a discontinuity-preserving noise reduction process was designed to reduce noise in the sparse depth information and provide a more reliable classification result to the Depth Map Rasterization process. Concretely, the proposed method initially estimates the continuous edges among adjacent edges for each node. Then, the refined depth value is computed by weighted averaging the depth values of the selected nodes based on the Joint Bilateral Upsampling. The process is iteratively performed until a given criterion is satisfied. After the noise reduction process, we can more reliably estimate whether there are depth discontinuities along each edge. Then, we generate a noise-reduced and dense depth map while preserving the depth discontinuities using the Depth Map Rasterization. In addition, we modify the similarity measure in the Depth Map Rasterization to improve the stability of the depth assignment. The main advantages of the proposed method can be summarised as follows:

1. Because the proposed method is performed on a mesh structure constructed from noisy and sparse depth information, it can be applied to various areas where a triangulation process is possible.
2. Even for regular or near-regular sampling structures, the proposed method can reconstruct higher quality dense depth maps compared to previous methods generally used for regular sampling structures

Figure 2: An example of the noise reduction and depth assignment processes. (a) In Step 2, The maximum distance $dist_{MAX}(\mathbf{a}_1) = dist(\mathbf{a}_1, \mathbf{a}_2)$ and $G_1 = \{\mathbf{a}_1, \mathbf{a}_5, \mathbf{a}_6, \mathbf{a}_7, \mathbf{a}_8, \mathbf{a}_9\}$ for node $\mathbf{a}_1$. The refined depth value d'_i is computed by weighted averaging of the depth values of the nodes in G_1 . (b) In Step 4, the similarities between a pixel $\mathbf{p} = (x_p, y_p)$ and the nodes $\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_9$ are respectively computed, where the Colour Representative Blocks (RBs) are used for colour similarity calculation between a pixel and each of the nodes. If the similarity value with $\mathbf{a}_1$ is larger than those with other nodes, the depth value of $\mathbf{p}$ is estimated by extrapolation using the planar depth parameters of $\mathbf{a}_1$. More details of Depth Map Rasterization can be shown in [12].

3.1 Overall Procedure of the Proposed Method

The overall procedure of the proposed upsampling method is shown in Fig.1. Each step of the proposed method is described as follows:

Step 1: 2D Triangulation is performed on a sparse depth map in an image plane, where the initial depth value of node $\mathbf{a}_i$ is d_i , and the edge between nodes $\mathbf{a}_i$ and $\mathbf{a}_j$ is $\langle \mathbf{a}_i, \mathbf{a}_j \rangle$.

Step 2: Let an $n \times n$ block centred at node $\mathbf{a}_i$ be BL_i in the triangulation in the image plane. For each node $\mathbf{a}_i$, $dist_{MAX}(\mathbf{a}_i)$ is determined as follows:

$$dist_{MAX}(\mathbf{a}_i) = \max(dist(\mathbf{a}_i, \mathbf{a}_k), thres_{MAX}) \text{ for all } \mathbf{a}_k, \exists \langle \mathbf{a}_i, \mathbf{a}_k \rangle \quad (1)$$

where $dist(\mathbf{a}_i, \mathbf{a}_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (d_i - d_j)^2}$ and $thres_{MAX}$ is the maximum distance and was set as 300.

A set G_i is then determined as follows:

$$G_i = \{\mathbf{a}_i\} \cup \{\mathbf{a}_j \mid |d_i - d_j| \leq \text{Mean}(|d_i - d_k|) + \sigma(|d_i - d_k|), \text{dist}(\mathbf{a}_i, \mathbf{a}_j) \leq \text{dist}_{\text{MAX}}(\mathbf{a}_i)\}, \quad (2)$$

where $\text{Mean}(|d_i - d_k|)$ and $\sigma(|d_i - d_k|)$ are respectively the mean and standard deviation of the absolute differences between the depth values d_i and d_k of nodes $\mathbf{a}_i$ and $\mathbf{a}_k$ satisfying $\text{dist}(\mathbf{a}_i, \mathbf{a}_k) \leq \text{dist}_{\text{MAX}}(\mathbf{a}_i)$.

A colour consistency between the nodes $\mathbf{a}_i$ and $\mathbf{a}_j \in G_i$ is computed as follows:

$$\text{Colour Consistency } CC = \sum_{(x,y) \in \text{BL}_j} |C(x_i, y_i) - C(x, y)|, \quad (3)$$

where $C(x, y)$ is a bilateral filtered colour value at (x, y) .

The refined depth value d'_i is then computed as follows:

$$d'_i = \frac{\sum_{\mathbf{a}_j \in G_i} w_j d_j}{\sum_{\mathbf{a}_j \in G_i} w_j}, \quad (4)$$

where

$$w_j = \exp\left(-\frac{\text{dist}(\mathbf{a}_i, \mathbf{a}_j)^2}{2\sigma_s^2}\right) \exp\left(-\frac{CC(\mathbf{a}_i, \mathbf{a}_j)^2}{2\sigma_c^2}\right). \quad (5)$$

We refine the depth values in descending order of $|G_i|$ at each iteration, i.e., the nodes with more continuous nodes are refined before the other nodes are processed. Step 2 is iteratively performed until the following criterion is satisfied:

$$\frac{|\text{Mean}_{\text{current}}|\langle \mathbf{a}_i, \mathbf{a}_j \rangle| - \text{Mean}_{\text{prev}}|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|}{\text{Mean}_{\text{prev}}|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|} \leq \text{thres}, \quad (6)$$

where $\text{Mean}_{\text{prev}}|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|$ and $\text{Mean}_{\text{current}}|\langle \mathbf{a}_i, \mathbf{a}_j \rangle|$ are the previous and current means of the absolute depth differences for all edges in the triangulation, respectively, and $\text{thres} = 0.1$ in this paper. Fig. 2 (a) shows an example of the noise reduction process for node $\mathbf{a}_1$ when $G_1 = \{\mathbf{a}_1, \mathbf{a}_5, \mathbf{a}_6, \mathbf{a}_7, \mathbf{a}_8, \mathbf{a}_9\}$ in a triangulation.

Step 3: Although the proposed method reduces noise in the sparse depth values while preserving the depth discontinuities in Step 2, the planar depth regions can remain uneven because the weights w_j are different for each node. To improve the flatness in these regions, the proposed method enforces a simple planar fitting after the iterative noise reduction process in Step 2. In the same manner with Step2, G_i is recomputed for each node. The planar parameters of α , β , and γ of $d''_i = \alpha x_i + \beta y_i + \gamma$ are then computed from the nodes in G_i , where (x_i, y_i) is the position of node $\mathbf{a}_i$ and d''_i is the final depth estimate of $\mathbf{a}_i$.

Step 4: In the previous steps, the sparse depth values have been refined, and the edges have been separated into continuous and discontinuous edges for each node. Given the processed depth and mesh information, depth values are assigned in the image plane using the Depth Map Rasterization method [14] in Step 4. Fig. 2 (b) shows an example of the depth assignment process for a pixel $\mathbf{p} = (x_p, y_p)$. In the previous approach, the spatial distance between each node and a pixel was not considered in the similarity measure [14]. However, if the distribution of sample points is highly uneven, a similarity measure with a spatial distance

	Art		Books		Dolls		Moebius		Reindeer		Bowling2	
	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)
Bilinear	38.3	60.0	27.0	53.2	30.1	54.5	30.1	54.6	28.8	54.7	25.0	52.6
JBF	25.3	40.1	18.3	38.9	20.5	31.0	23.6	31.8	19.8	39.4	15.5	26.2
IBF	17.1	27.0	12.8	30.6	22.2	27.7	12.1	23.1	16.5	29.6	11.6	22.0
WMF	15.8	28.7	12.2	33.1	19.9	32.7	17.1	27.0	24.9	33.2	11.8	21.0
DMR	15.3	28.1	11.7	32.4	13.0	35.5	15.4	28.6	9.66	34.1	9.79	29.4
Proposed	10.2	19.8	7.69	17.1	9.19	21.3	10.1	18.0	8.56	17.3	7.97	14.9

(a)

	Art		Books		Dolls		Moebius		Reindeer		Bowling2	
	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)	(0, 1)	(0, 4)
JBF	38.4	51.8	20.5	29.9	23.2	42.7	24.3	42.0	21.9	40.8	17.4	36.6
IBF	14.4	31.9	21.5	27.5	25.2	42.0	14.5	27.5	18.3	34.3	15.1	28.9
WMF	18.4	40.0	24.1	33.1	23.1	43.8	18.4	27.7	17.7	45.9	13.7	33.5
DMR	17.8	38.7	13.3	33.1	16.6	46.2	16.7	27.8	12.8	43.9	11.4	29.0
Proposed	13.2	22.2	10.8	19.0	13.9	26.3	12.2	19.0	10.2	21.9	7.22	14.3

(b)

Table 1: Quantitative evaluation for the upsampled disparity maps. We measured percentage of bad pixels for Middlebury test set with (a) 8x8 regular and (b) irregular sampling structures, where Gaussian noise of (0, 1) and (0, 4) are added to the downsampled ground truth disparity values, respectively.

term tends to show more stable results. To improve the stability, we modify the similarity measure between the Colour Representative Blocks (RBs) of node $\mathbf{a}_h$ and pixel $\mathbf{p}$ by adopting the spatial distance term as follows:

$$Sim(\mathbf{p}, \mathbf{RB}_h^{l_h}) = \exp\left(\frac{-CS(\mathbf{p}, \mathbf{RB}_h^{l_h})^2}{2\sigma_s^2}\right) \exp\left(\frac{-d(\mathbf{p}, \mathbf{RB}_h^{l_h})^2}{2\sigma_c^2}\right) \quad (7)$$

for $h = i, j$, and k , and $1 \leq l_h \leq$ the number of RBs of $\mathbf{a}_h$,

where $CS(\mathbf{p}, \mathbf{RB}_h^{l_h}) = \sum_{(x,y) \in \mathbf{BL}_p} |C(x,y) - \mathbf{MRB}_h^{l_h}|$, $\mathbf{MRB}_h^{l_h}$ is the mean of the colour values of $\mathbf{RB}_h^{l_h}$, $\mathbf{BL}_p$ is an $n \times n$ block centred at $\mathbf{p}$, and $d(A, B)$ is the Euclidean distance between the centroids of A and B. In the Depth Map Rasterization method, after the similarity calculation for each pixel, a depth estimate is computed from the depth information of a neighbour node with the maximum similarity.

4 Results

For the quantitative evaluation, the error rates of the upsampled disparity maps were measured for the Middlebury test set (Art, Books, Dolls, Moebius, Reindeer, and Bowling2) [28]. Two different types of sampling structures were set for the performance evaluation. The ground truth disparity maps were down-sampled at a ratio of 8 for each direction for regular 8x8 sampling. On the other hand, for the irregular sampled disparity maps, a pixel position with maximum edge value within an 8x8 block was selected as a sample position.

Figure 3: Disparity map upsampling from 8x8 regularly sampled disparity values with Gaussian noise of (0,4) for Dolls: (a) Colour image, (b) Ground truth, (c) Bilinear, (d) JBF, (e) IBF, (f) WMF, (g) DMR, and (h) the proposed method.

Figure 4: Disparity map upsampling from 8x8 irregularly sampled disparity values with Gaussian noise of (0,4) for Moebius: (a) Colour image, (b) Ground truth, (c) triangulation on the sampling structure, (d) JBF, (e) IBF, (f) WMF, (g) DMR, and (h) the proposed method.

For each sampling structure, Gaussian noise of (0,1) and (0,4) were added to the sparse disparity values, respectively. The percentage of bad pixels (error tolerance $\delta_d = 1.0$) of the generated disparity maps was evaluated through a comparison with ground truth disparity maps. The parameters used in the experiment were assigned as follows: the block size of BL_i was 3x3, and σ_s and σ_c were 5 and 10, respectively.

4.1 Quantitative Evaluation for Regular Sampling Structures

In the quantitative evaluation for regular sampling structures under noisy conditions, the disparity maps were upsampled using a Bilinear Filter, Joint Bilateral Filter (JBF) [14], Iterative Bilateral Filter (IBF) [15], Weighted Mode Filter (WMF) [16], Depth Map Rasterization [17], and the proposed method from the 8x8 regularly down-sampled ground truth disparity maps with Gaussian noise of (0,1) and (0,4), respectively. Fig. 3 shows the disparity

maps upsampled using the comparison and the proposed methods for 8x8 regular sampling structures with Gaussian noise of (0, 4) for Dolls.

4.2 Quantitative Evaluation for Irregular Sampling Structures

In the quantitative evaluation for irregular sampling structures under noisy conditions, the disparity maps were upsampled using JBF [5], IBF [10], WMF [10], DMR [10], and the proposed method from the irregularly down-sampled ground truth disparity maps with Gaussian noise. Table 1 (b) shows the percentage of bad pixels of the reconstructed disparity maps of the test set for the irregular sampling structures with Gaussian noise of (0, 1) and (0, 4), respectively. Fig. 4 shows the sampling structure (Fig. 4 (c)) and the disparity maps upsampled using the comparison and the proposed methods for irregular sampling structures with Gaussian noise of (0, 4) for Moebius. The experimental results show that the proposed method outperforms the compared methods for both sampling structures with moderate (0, 1) and severe (0, 4) noise conditions.

5 Application to Multi-View Stereo

We applied the proposed method to reconstruct spatially consistent and dense depth maps from a sparse and noisy point cloud, and generated a dense 3D point cloud. The resolutions of the test sequences were 1920x1080, which were taken by a mobile phone. In the application, σ_s and σ_c were 20 and 30, respectively, for the test sequences.

In the multi-view stereo process, sparse 3D points and camera parameters were first computed using the Structure from Motion method [24, 30], where we applied CeresSolver [31] for a bundle adjustment process. The sparse 3D points were projected onto each image plane to acquire a sparse depth map for each image plane. To reconstruct dense depth maps from the obtained sparse depth maps with the corresponding colour images, the proposed mesh- and colour-consistency-based depth map upsampling method was used for each image plane. In the depth map upsampling, the depth value Z is assigned as $d = 1/Z$ for each node. Fig. 5 shows the overall procedure of the multi-view stereo, where spatially consistent depth maps are reconstructed by the proposed method. A dense 3D point cloud is then obtained from the dense depth maps by a depth map merging process. To generate a 3D model from the 3D point cloud, we can apply various surface reconstruction and texture mapping methods such as the Poisson surface reconstruction [62].

Fig. 6 shows the selected colour images of a test sequence (kitchen), initial and refined sparse 3D point clouds, generated depth maps, and the dense 3D point cloud. The results verify the applicability and potential of the proposed method to reconstruct dense depth maps from various sampling structures with noise conditions, such as multi-view stereo of indoor scenes with partially homogeneous regions. The supplementary video of the results is available at www.youtube.com/channel/UCy2jD71OxUyPBtqWATdzW7g/.

6 Conclusion

In this paper, we proposed a mesh- and colour-consistency-based sparse-to-dense depth map upsampling method for use under noisy conditions and with various sampling structures. Furthermore, we applied the proposed method to multi-view stereo to verify its applicability.

Figure 5: Application of the proposed method to multi-view stereo: (a) selected frames from a test sequence, (b) an initial sparse 3D point cloud with noise, which was obtained by using a SfM method (c) triangulation on the sampling structure for each image plane, (d) a refined 3D point cloud and (e) reconstructed dense depth maps by the proposed method, and (f) a generated dense 3D point cloud.

As the results show, the proposed method successfully generates dense depth maps from sparse depth information with various degrees of noise, and can be applied to reconstructing spatially consistent depth maps and 3D point clouds.

Our future work is to apply the proposed method to various areas such as depth map coding and 3D modelling of indoor scenes.

ACKNOWLEDGEMENT This work was supported by Institute for Information and communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) and Korea Creative Content Agency(KOCCA) grant funded by the Korea government(MCST) (R0132-15-1006, Developing the technology of open composable content editors for realistic media).

(a)

(b)

(c)

(d)

(e)

(f)

Figure 6: Dense 3D point cloud generation from kitchen sequence: (a) selected frames from the sequence, (b) an initial sparse 3D point cloud with noise, (c) triangulation on the sampling structure for each image plane, (d) a refined 3D point cloud and (e) reconstructed dense depth maps by the proposed method, and (f) a generated dense 3D point cloud.

References

- [1] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-Time Dense Surface Mapping and Tracking," ISMAR, 2011.
- [2] P. Tanskanen, K. Kolev, L. Meier, F. Camposeco, O. Saurer, and M. Pollefeys, "Live Metric 3D Reconstruction on Mobile Phones," IEEE International Conference on Computer Vision, 2013.
- [3] H. Urey, K. V. Chellappan, E. Erden, and P. Surman, "State of the Art in Stereoscopic and Autostereoscopic Displays," Proceedings of the IEEE, Vol. 99, No. 4, pp. 540-555, 2011.
- [4] J.Y. Jang and S.W. Nam, "Fast Random-Forest-Based Human Pose Estimation Using a Multi-scale and Cascade Approach," ETRI Journal, Vol. 35, No. 6, pp. 949-959, 2013.
- [5] J. Kopf, M. Cohen, D. Lischinski, and M. Uyttendaele, "Joint Bilateral Upsampling," SIGGRAPH, 2007.
- [6] J. Diebel and S. Thrun, "An Application of Markov Random Fields to Range Sensing," In NIPS, 2005.
- [7] X. Cao, Z. Li, and Q. Dai, "Semi-Automatic 2D-to-3D Conversion Using Disparity Propagation," IEEE Trans. on Broadcasting, Vol. 57, No. 2, pp. 491-499, 2011.
- [8] K. Müller, Ph. Merkle, and T. Wiegand, "3-D Video Representation Using Depth Maps," Proceedings of the IEEE, Vol. 99, No. 4, pp. 643-656, 2011.
- [9] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multiview Stereopsis," IEEE Trans. on PAMI, Vol. 32, No. 8, pp. 1362-1376, 2010.
- [10] Q. Yang, R. Yang, J. Davis, and D. Nistér, "Spatial-Depth Super Resolution for Range images," IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- [11] D. Min, J. Lu, and M. N. Do, "Depth Video Enhancement Based on Weighted Mode Filtering," IEEE Trans. on IP, Vol. 21, No. 3, pp.1176-1190, 2012.
- [12] H.S. Lim and H.W. Park, "Depth Map Rasterization Based on Triangulation and Color Consistency," IEEE Trans. on CSVT, Vol. 25, No. 7, pp. 1081-1098, 2015.
- [13] J. Dolson, J. Baek, C. Plagemann, and S. Thrun, "Upsampling range data in dynamic environments," IEEE Conference on Computer Vision and Pattern Recognition, 2010.
- [14] K. He, J. Sun, and X. Tang, "Guidance image filtering," European Conference on Computer Vision, 2010.
- [15] J. Yang, X. Ye, K. Li, C. Hou, and Y. Yang, "Color-guided depth recovery from RGBD data using an adaptive autoregressive model," IEEE Trans. on Image Processing, Vol. 23, No. 8, pp. 3443-3458, 2014.
- [16] G. Wolberg, Digital Image Warping, IEEE Computer Society Press, 1990.

- [17] D. Gallup, J. M. Frahm, P. Mordohai, Q. Yang, and M. Pollefeys, "Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions," IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- [18] J. Wei, B. Resch, and H. Lensch, "Multi-view depth map estimation with cross-view consistency," BMVC 2014.
- [19] S. Galliani, K. Lasinger, and K. Schindler, "Massively Parallel Multiview Stereopsis by Surface Normal Diffusion," IEEE International Conference on Computer Vision, 2015.
- [20] J. L. Schönberger, E. Zhang, J. M. Frahm, and M. Pollefeys, "Pixelwise View Selection for Unstructured Multi-View Stereo," European Conference on Computer Vision, 2016.
- [21] V. A. Nguyen, D. Min, and M. N. Do, "Efficient Techniques for Depth Video Compression Using Weighted Mode Filtering," IEEE Trans. on CSVT, Vol. 23, No. 2222, pp. 189-202, 2013.
- [22] S. Schwarz, R. Olsson, M. Sjöström, and S. Tourancheau, "Adaptive depth filtering for HEVC 3D video coding," Picture Coding Symposium 2012.
- [23] C. Zhang, Z. Li, Y. Cheng, R. Cai, H. Chao, and Y. Ruk, "MeshStereo: A Global Stereo Model with Mesh Alignment Regularization for View Interpolation," IEEE International Conference on Computer Vision, 2015.
- [24] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," Proc. SPIE, 5291, 93-104 (2004).
- [25] N. Plath, S. Knorr, L. Goldmann, and T. Sikora, "Adaptive Image Warping for Hole Prevention in 3D View Synthesis," IEEE Trans. on Image Processing, Vol. 22, No. 9, pp. 3420-3432, 2013.
- [26] D. Doyen, S. Thiébaud, "Towards A Free Viewpoint and 3D Intensity Adjustment on Multi-View Display," ICIP 2014.
- [27] Y. Park, V. Lepetit, W. Woo, "Texture-Less Object Tracking with Online Training using An RGB-D Camera," ISMAR 2011.
- [28] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereo matching," IEEE Conference on Computer Vision and Pattern Recognition, 2007.
- [29] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge, 2003.
- [30] M. Pollefeys, L.V. Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, "Visual modeling with a hand-held camera," International Journal of Computer Vision, Vol. 59, No. 3, pp. 207-232, 2004.
- [31] S. Agarwal, Keir Mierle, and Others, Ceres Solver, <http://ceres-solver.org>.
- [32] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson Surface Reconstruction," Eurographics Symposium on Geometry Processing, 2006.