Reconstruction of high-resolution Depth Map using Sparse Linear Model

Hanqi Fan, Dexing Kong, Jinhong Li

College of Computer Science, North China University of Technology, Beijing, China
E-mail: hanchy@gmail.com

Keywords: Depth Image; Sparse Representation

Abstract. In this paper, we propose a method that constructs a high-resolution depth image with high quality from a low-resolution depth image that is noisy and contains holes. We believe that the high-resolution depth map is generated by sparse linear combination of atoms from an over-complete dictionary, and the low-resolution depth map are the samples from the high-resolution depth map. Under Bayesian framework, we find the optimal sparse coefficient vector that represents the high-resolution map best. Comprehensive quantitative comparisons show that our method outperforms existing approaches when applied on Middlebury dataset, and qualitative comparison on real scenes indicates that our algorithm performs best.

Introduction

Capturing accurate depth information of a scene is very important in computer vision and computer graphics, and it has a lot of applications, such as 3D model reconstruction, Human–Computer Interaction, 3D television, etc. There are mainly three categories of methods to obtain depth information. Laser scanner obtain depth information highly accurately, but the processing of scanning takes too much time and the scanner is too expensive. Stereo matching method obtains depth information with high level of noise and demands the targets covered by textures. As technology advances in recent years, depth cameras such as ToF (Time of Flight) and Kinect have appeared which are cheap and widely used in household applications. They capture depth map of a scene very quickly with accuracy between the two methods above. However, the resolutions of the depth maps are very low. The resolution of a ToF camera is usually 320×240, and Kinect provides 640×480 resolution depth maps, which is far lower than color cameras. Besides, the depth information captured by these devices is strongly noisy and often contains holes.

In order to generate a high-resolution depth map from a low-resolution depth map, a color camera is needed usually. Once a color camera and a depth camera are calibrated, we can project the depth information onto the color image. Since the resolution of the depth map is much lower than the color image, some part of pixels of color image has no depth information, and sometimes some connected regions have no depth information (Fig 1. b). In order to fill the missing depth value, most existing algorithms assume that pixels with similar colors have similar depth values, and otherwise, they have very different depth value. Besides, these algorithms cannot process depth map containing holes very well. These inherent flaws make these algorithms very susceptible to the texture of color image and hole-filling results are practically bad (Fig. 3).

We observe that not all of the color pixels have depth value, and we interpret this phenomenon as that observed depth information are samples from high-resolution depth map and these depth values are assigned to the corresponding color pixels. Since the observation is a non-ideal random process, the depth information of the color image is highly corrupted by noise and contains holes. Inspired by multi-morphology signal analysis [1, 2] and compressed sensing theory [3, 4, 5], we use a set of sparse linear combinations of over-complete bases [6] to express high-resolution depth map, and the reconstruction problem is reduced to solving the high-dimensional sparse signals, that is the coefficient vector of the linear combination. When pursuing this sparse coefficient vector, we only use the observed depth values as constraints. Because no information from color image is used, textures have no impact on the depth image reconstruction. The quantitative experiments on Middlebury dataset show that our algorithm outperforms typical algorithms. In the experiments of

reconstructing depth maps of real scenes, our proposed algorithm generates the most accurate depth maps when viewed as point clouds.

Related Work

The existing methods can be divided into two categories, which are filter-based and optimization-based.

Filter-based algorithms are the most common, such as JBU (Joint Bilateral Upsampling) [7], NAFDU (Noise-Aware Filter for Depth Upsampling) [8] and JABDU (Joint-Adaptive Bilateral Depth map Upsampling) [9]. JBU is very popular because of its simplicity. It can even be done in real-time [10]. However, it is so susceptible to scene textures that texture copying often occurs. NAFDU takes a further consideration of noises in depth map to compute hybrid weights. It can reduce the impact of color textures and improve the reconstruction quality. JABDU can only improve edges of depth maps that are generated by bilinear interpolation. Filter-based algorithms are simple and with low complexity. They can generate relatively smooth depth maps. However, they are more likely to blur boundaries of objects in depth maps.

(a) (b) (c) (d)

Figure 1. Illustration of high-resolution depth map reconstruction: (a) the low-resolution depth map projected to the color image (the red points are the low-resolution depth map); (b)the zooming view of green box(a); (c) reconstructed depth map; (d) point cloud model of the depth map.The low-resolution depth map contains a lot of holes and is noisy, and our method fills the holes when reconstructing the high-resolution depth map.

As optimization-based algorithms, Diebel and Thrun [11] constructed a double Markov random Field. They used the color similarities of pixels to weight nodes. Based on that, Park et al. [12] proposed HQDMU (High Quality Depth Map Upsampling for 3D-tof cameras) algorithm. HQDMU takes multiple factors to weight nodes on depth edges and takes similarities of textures in local regions to weight nodes on other parts of depth maps. In order to generate better depth edges, Ferstl et al. [13] used a second-order TGV(Total Generative Variation) of color textures to segment the color image into slices. Then they reconstructed depth maps by solving a convex optimization problem. Optimization-based algorithms can bring out better results than filter-based algorithms. However, color textures still affect the depth map reconstruction and texture copying still occurs. Besides, all of these algorithms assume that the low-resolution depth map contains no holes and is projected to the color image evenly.

I. Our Algorithm

Before we apply our algorithm, the low-resolution depth camera and the high-resolution camera should be calibrated together. We achieve this task by adopting the algorithm proposed by Raposo et al. [14] who improved the accuracy of the algorithm proposed by Herrera et al. [15].

In order to reduce the time-space complexity of the proposed algorithm, we reconstruct the sparse high-resolution depth map block by block, and the reconstructed blocks compose to the whole new high-resolution depth map.

A. Sparse Linear Representation of the Depth Map

In this paper, the high-resolution depth map is denoted as $\mathbf{x} \in \mathbb{R}^n$, where n is the number of pixels. $\mathbf{x}$ can be a linear combination of l n-dimensional atoms:

$$\mathbf{x} = \mathbf{\Phi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \in N(\mathbf{0}, \sigma^2\mathbf{I}), \qquad (1)$$

where $\mathbf{\Phi} \in \mathbb{R}^{n \times l}$ is usually refereed as dictionary, each column vector of which is refereed as atom. $\boldsymbol{\alpha} \in \mathbb{R}^{l}$ is the corresponding coefficient vector and $\boldsymbol{\varepsilon}$ is the noise.

The expression capability of $\mathbf{\Phi}$ is critical because it directly affects the quality of reconstructed depth maps. Starck et al. [16] discussed the capabilities of different dictionaries to express different image shape forms. In order to obtain a stronger expression, different kinds of dictionaries can be merged into $\mathbf{\Phi}$. The dictionary to express high-resolution depth maps in this paper contains four different types of transformation: Daubechies wavelet to express small local features suitably, LDCT (Local Discrete Cosine Transfrom) to express repetitive texture, Grouplets [17, 18] to capture the directional structural features; Curvelet [19, 20] to express depth map with smooth piecewise. We develop the high-resolution depth map into a 1-dimensional vector $\mathbf{x} \in \mathbb{R}^{n}$. The four types of transformation matrixes above are generated according to the length of $\mathbf{x}$. Then they are merged into $\mathbf{\Phi}$. Obviously, the number of columns of $\mathbf{\Phi}$ is much greater than the number of rows ($l \gg n$), which makes this system to be a high degree of redundancy linear system. As indicated by Starck et al [16], it is more suitable to use a sparse linear combination of atoms to express high-resolution depth maps in this redundant system. The term of sparseness means that there are only small part of elements in $\boldsymbol{\alpha}$ is not zero when the dictionary $\mathbf{\Phi}$ can express the depth map $\mathbf{x}$ perfectly.

The high-resolution depth map $\mathbf{x}$ consists of two parts: the observed depth values $\mathbf{x}_{obs} \in \mathbb{R}^{m} = \{\mathbf{x}_i\}$, $i \in I_{obs}$ and the missing depth values $\mathbf{x}_{miss} = \{\mathbf{x}_i\}$, $i \in I_{miss}$. I_{obs} and I_{miss} are indexes of the observed depth values and the missing depth values in $\mathbf{x}$. $\mathbf{x}_{obs}$ is available and $\mathbf{x}_{miss}$ is to be solved. The relationship between $\mathbf{x}_{obs}$ and $\mathbf{x}$ is

$$\mathbf{x}_{obs} = \mathbf{S}\mathbf{x}, \qquad (2)$$

where $\mathbf{S} \in \mathbb{R}^{m \times n}$ is the sample matrix. Thus, $\mathbf{x}_{obs}$ can be expressed as

$$\mathbf{x}_{obs} = \mathbf{S}\mathbf{x} = \mathbf{S}\mathbf{\Phi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}_{obs} = \mathbf{A}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}_{obs}. \qquad (3)$$

The sparseness of $\boldsymbol{\alpha}$ is very important. If $\mathbf{x}$ can be expressed very well when some atoms in $\mathbf{\Phi}$ are not needed, then the corresponding elements in $\boldsymbol{\alpha}$ should be set zero. Otherwise, the corresponding elements in $\boldsymbol{\alpha}$ can have large values. In order to minimize the coherence between $\mathbf{S}$ and $\mathbf{\Phi}$, we amend the dictionary which is to express $\mathbf{x}$ to this:

$$\mathbf{x} = \mathbf{G}\mathbf{\Phi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon} = \mathbf{\Psi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \qquad (4)$$

where $\mathbf{G}$ is a Gaussian matrix suggested by yang et al.[21]. In (3), $\boldsymbol{\alpha}$ is required to be sparse. Therefore, it can be understood from (4) that the high-resolution depth map $\mathbf{x}$ can be sparsely and linearly expressed by atoms in the dictionary $\mathbf{\Psi}$.

The sparseness of $\boldsymbol{\alpha}$ discussed above can be considered as the priori knowledge about $\boldsymbol{\alpha}$, and we use the Laplace distribution:

$$Pr(\boldsymbol{\alpha}) = \frac{\lambda}{2} e^{-\lambda \sum_i |\alpha_i|}. \qquad (5)$$

B. The Algorithm to Solve Sparse Vector

Since $\mathbf{x} = \mathbf{\Psi\alpha} + \mathbf{\varepsilon}$, it can be considered that $\mathbf{x}$ obeys a normal distribution like this:

$$Pr(\mathbf{x} \mid \sigma, \mathbf{\Psi}, \mathbf{\alpha}) = \frac{1}{Z} \exp(-\frac{\left\|\mathbf{x} - \mathbf{\Psi\alpha}\right\|_2^2}{2\sigma^2}). \qquad (6)$$

The joint probability distribution of $\mathbf{x}$ and $\mathbf{\alpha}$ is

$$Pr(\mathbf{x}, \mathbf{\alpha} \mid \sigma, \mathbf{\Psi}) = Pr(\mathbf{x} \mid \sigma, \mathbf{\Psi}, \mathbf{\alpha}) Pr(\mathbf{\alpha}). \qquad (7)$$

We obtain the missing depth values $\mathbf{x}_{miss}$ and coefficient vector $\mathbf{\alpha}$ by solving this optimization problem:

$$\begin{aligned} \{\mathbf{x}, \mathbf{\alpha}\}_{opt} &= \arg\max_{\mathbf{x},\mathbf{\alpha}} Pr(\mathbf{x}, \mathbf{\alpha} \mid \sigma, \mathbf{\Psi}) \\ &= \arg\max_{\mathbf{x},\mathbf{\alpha}} \log Pr(\mathbf{x}, \mathbf{\alpha} \mid \sigma, \mathbf{\Psi}) \qquad (8) \\ &= \arg\max_{\mathbf{x},\mathbf{\alpha}} \left\|\mathbf{x} - \mathbf{\Psi\alpha}\right\|_2^2 + \lambda \left\|\mathbf{\alpha}\right\|_1 . \end{aligned}$$

We use the Expectation Maximization (EM) algorithm to solve (8). EM algorithm solves $\mathbf{x}_{miss}$ and $\mathbf{\alpha}$ iteratively. In detail, the algorithm consists of two main steps.

1) The E Step

From the probability distribution of $\mathbf{x}$ (6), the probability density distribution of $\mathbf{x}_{miss}$ can be obtained:

$$\begin{aligned} f(\mathbf{x}_{miss}) &= Pr(\mathbf{x}_{miss} \mid \mathbf{x}_{obs}, \mathbf{\Psi}, \mathbf{\alpha}^{old}, \sigma) \\ &= Pr(\mathbf{x}_{miss} \mid \mathbf{\Psi}, \mathbf{\alpha}^{old}, \sigma) \qquad (9) \\ &= N(\mathbf{x}_{miss} \mid (\mathbf{\Psi\alpha}^{old})_{miss}, \sigma). \end{aligned}$$

Under this probability density distribution:

$$E_{f(\mathbf{x}_{miss})}(\ln Pr(\mathbf{\alpha})) = \lambda \left\|\mathbf{\alpha}\right\|_1 , \qquad (10)$$

we define probability distribution of $\mathbf{x}_{miss}$ (9) as this equivalent form [22]:

$$\mathbf{x}_{miss} = (\mathbf{\Psi\alpha}^{old})_{miss} + \mathbf{\varepsilon}_{miss}. \qquad (11)$$

Equation (11) indicates that $\mathbf{x}_{miss}$ can be directly replaced with $(\mathbf{\Psi\alpha}^{old})_{miss} = \{(\mathbf{\Psi\alpha}^{old})_i\}$, $i \in I_{miss}$ in each E step of the EM algorithm iterations.

2) The M Step

After the replacement of the missing values in the E step, all elements of $\mathbf{x}$ is available. Then, a new sparse coefficient vector $\mathbf{\alpha}^{new}$ can be obtained by solving this optimization problem:

$$\begin{aligned} \mathbf{\alpha}^{new} &= \arg\min_{\mathbf{\alpha}} - E_{f(\mathbf{x}_{miss})}(\ln Pr(\mathbf{x}, \mathbf{\alpha} \mid \mathbf{\Psi}, \sigma)) \\ &= \arg\min_{\mathbf{\alpha}} (\left\|\mathbf{x} - \mathbf{\Psi\alpha}\right\|_2^2 + \lambda \left\|\mathbf{\alpha}\right\|_1). \end{aligned} \qquad (12)$$

An iterative greedy algorithm proposed by Li and Osher [23] can very quickly solve (12).

After the convergence of the EM algorithm, the optimal coefficient vector $\boldsymbol{\alpha}^{opt}$ can be obtained. Thereby the high-resolution depth map is $\boldsymbol{\Psi}\boldsymbol{\alpha}^{opt}$.

In initial stages of the EM algorithm, the existence of $\mathbf{x}_{miss}$ makes $\boldsymbol{\alpha}$ very unstable and we use relatively larger λ, which sparsifies $\boldsymbol{\alpha}$ quickly. As the iteration goes on, we gradually decrease λ, so the optimization results can be closer to the observed values.

Experiment and Analysis

We implement the proposed algorithm, and apply it on Middlebury dataset. Because the ordinary high-resolution depth map (the ground truth) exists, we conduct a quantitative research on our results, and then compare them with results generated by some other conventional algorithms. We also apply the proposed algorithm on real scenes and have it compared with some typical algorithms in visual effects.

C. Middlebury Dataset

We randomly select five scenes from Middlebury 2006 [24] dataset to test, which are Tsukuba, Art, Books, Laundry and Bowling1. In order to test our algorithm, we add noises and holes to the ordinary depth map and down-sample it to 1/4 its ordinary size. Then we reconstruct it using the proposed algorithm and other conventional algorithms. Finally, we compare the reconstructed depth maps with the ground truth. As it can be seen in Fig. 2, our algorithm outperforms other algorithms.

In order to evaluate each algorithm accurately, we use three different methods to measure the quality of reconstructed depth maps: 1) PSNR (Peak Signal-to-Noise Ratio) to measure the overall quality; 2) DISC to measure the quality of discontinuous regions; 3) SRMS to measure the quality of smooth regions.

For a single-channel image I of m × n size and a reconstructed image K to be measured, MSE (Mean Squared Error) is

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(I_{ij} - K_{ij})^2. \qquad (13)$$

Based on MSE, PSNR is defined as

$$PSNR = 10(2\log_{10} I_{max} - \log_{10} MSE), \qquad (14)$$

where I_{max} is the maximum value in I. Obviously, the reconstructed image K is closer to the ordinary image I when PSNR is larger.

DISC typically is used to measure the quality of discontinuous regions. In these regions, a pixel is considered to be a wrong point if the corresponding deviation between K and I is larger than a certain threshold. The ratio between the number of wrong points and the number of total pixels in discontinuous regions is defined as DISC. It means that there are more wrong points when the DISC is larger.

After the discontinuous regions are removed, the MSE on the remaining regions (referred as smooth regions usually) is defined as SRMS:

$$SRMS = \sqrt{MSE}, (i, j) \in SmoothRegion. \qquad (15)$$

Obviously, the SRMS is smaller when the two smooth regions of K and I are similar.

We implement JABDU [9] and use the source codes provided by the authors of HQDMU [12] and ATGV [13]. Then we compare the results generated by these algorithms with ours. Table I shows the comparison of PSNR, which indicates that our results have the best quality overall. In

aspects of DISC and SRNS, our algorithm is better than other algorithms in most cases, which can be seen in Table II and Table III. In conclusion, our algorithm can generate more accurate results than other algorithms.

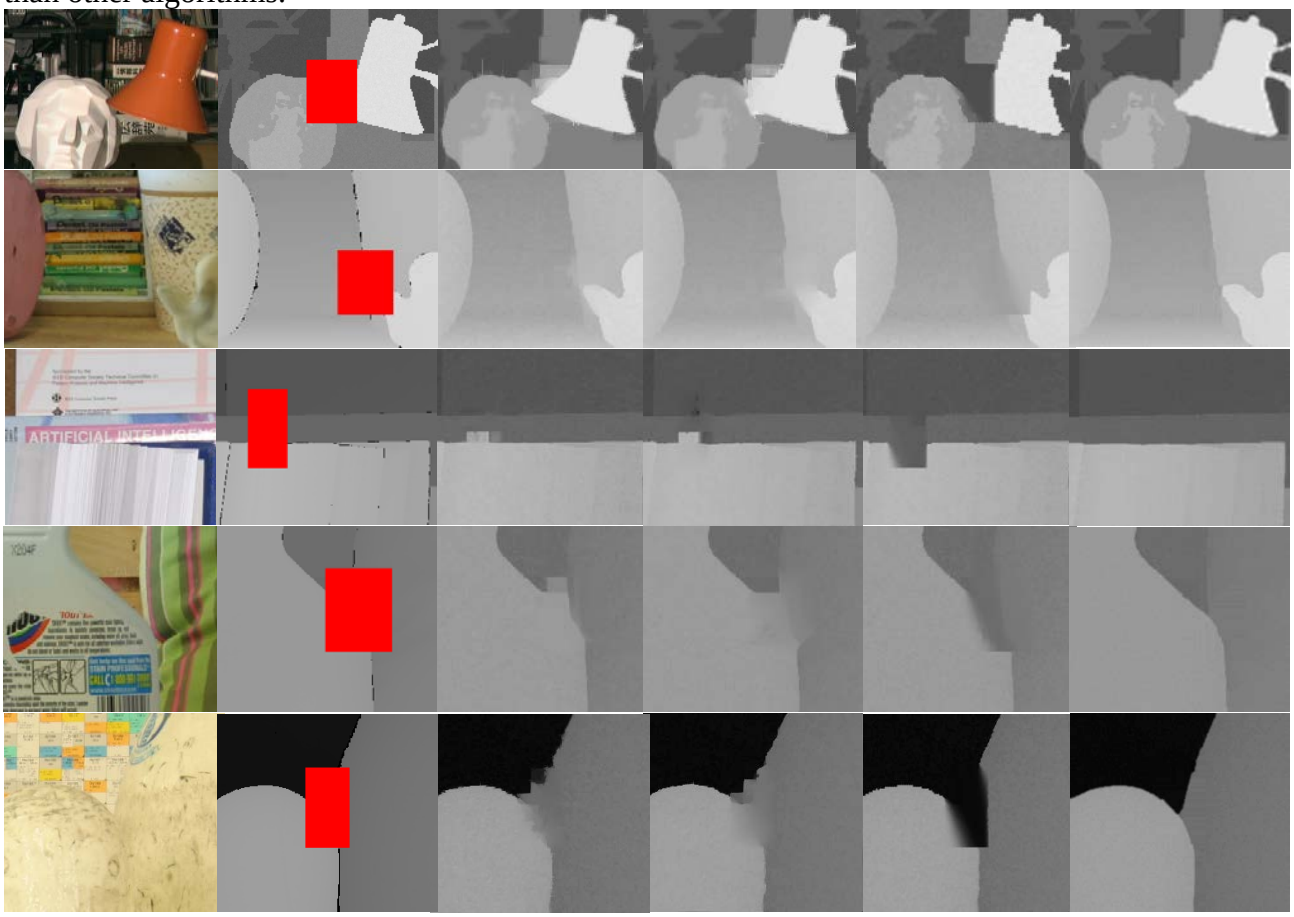


Figure 2. The comparison of the visual effects of upsampling results using typical algorithms applied on Middlebury dataset. From left to right, each are the color image, the damaged(the red region) ground truth and upsampling results using HQDMU[12], ATGV[13], JABDU[9] and the proposed method.

TABLE I. THE PSNR (PEAK SIGNAL TO NOISE RATIO, IN DB) OF THE UPSAMPLING RESULTS OF TYPICAL ALGORITHMS APPLIED ON MIDDLEBURY DATASET(*THE RED NUMBER IS THE BEST VALUE*).

Scene	HQDMU [12]	ATGV[13]	JABDU [9]	Our Algorithm
Tsukuba	27.1566	25.7923	19.9521	30.5765
Art	39.5957	38.3369	33.4176	42.0352
Books	33.4420	32.1950	25.0264	34.7531
Laundry	34.7523	36.0453	31.6916	41.9068
Bowling1	27.5619	27.0297	21.9769	27.8943

TABLE II. THE DISC (THRESHOLD: 3) OF THE UPSAMPLING RESULTS OF TYPICAL ALGORITHMS APPLIED ON MIDDLEBURY DATASET(THE DATA IN RED IS THE BEST VALUE).

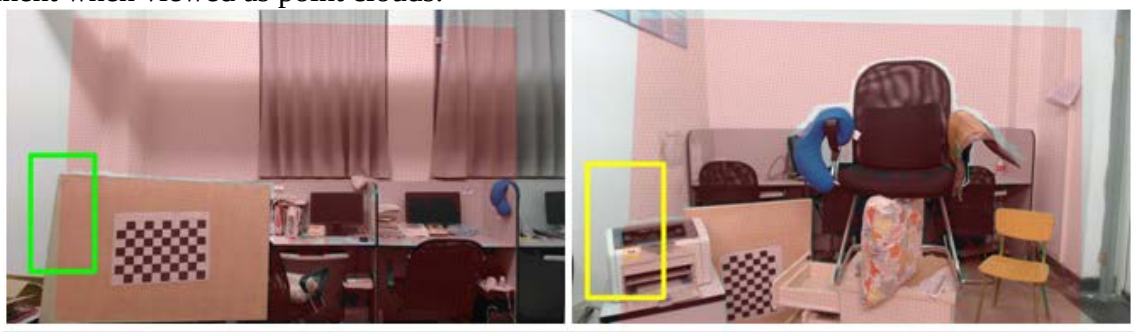
Scene	HQDMU[12]	ATGV[13]	JABDU[9]	Our Algorithm
Tsukuba	0.545709	0.524584	0.413348	0.589235
Art	0.415194	0.393110	0.470848	0.218322
Books	0.251009	0.319633	0.354128	0.128977
Laundry	0.303883	0.444660	0.466990	0.155926
Bowling1	0.584756	0.502439	0.400610	0.219847

TABLE III. THE SRMS OF THE UPSAMPLING RESULTS OF TYPICAL ALGORITHMS APPLIED ON MIDDLEBURY DATASET(THE DATA IN RED IS THE BEST VALUE).

Scene	HQDMU[12]	ATGV[13]	JABDU[9]	Our Algorithm
Tsukuba	5.38272	7.24882	21.8286	2.98375
Art	2.18021	2.59023	4.79732	1.72906
Books	4.62262	5.22905	13.8256	4.35283
Laundry	4.35383	3.74181	6.37146	1.83971
Bowling1	8.65488	9.36677	18.4914	9.54324

D. Real Scenes

We also apply HQDMU [12], ATGV [13] and our algorithm on real scenes. Because the low-resolution depth map contains a lot of noise and holes and is not projected onto the color image evenly (Fig. 3), the reconstruction is very changeling. Fig. 3 shows the comparison of the results. It can be seen that our algorithm can keep depth edges while suppressing noise and filling holes. HQDMU and ATGV act particularly bad on depth edges. It is mainly because that there are a lot of holes around depth edges, and these two algorithms fill holes using color textures. Thus, wrong results are generated when the textures of the holes corresponds are similar to the surrounding textures (Fig. 3). In fact, the interaction of HQDMU can only improve the visual effect of depth maps, but cannot fundamentally remove the impacts from color textures. In scene 1 of Fig. 3, the rectangular plate and the wall have similar textures (the green box), so results of HQDMU and ATGV are far worse than ours. In scene 2 of Fig. 3, although the printer and the wall share the same color (the yellow box), our results are still very accurate. The quality of our results is particularly prominent when viewed as point clouds.



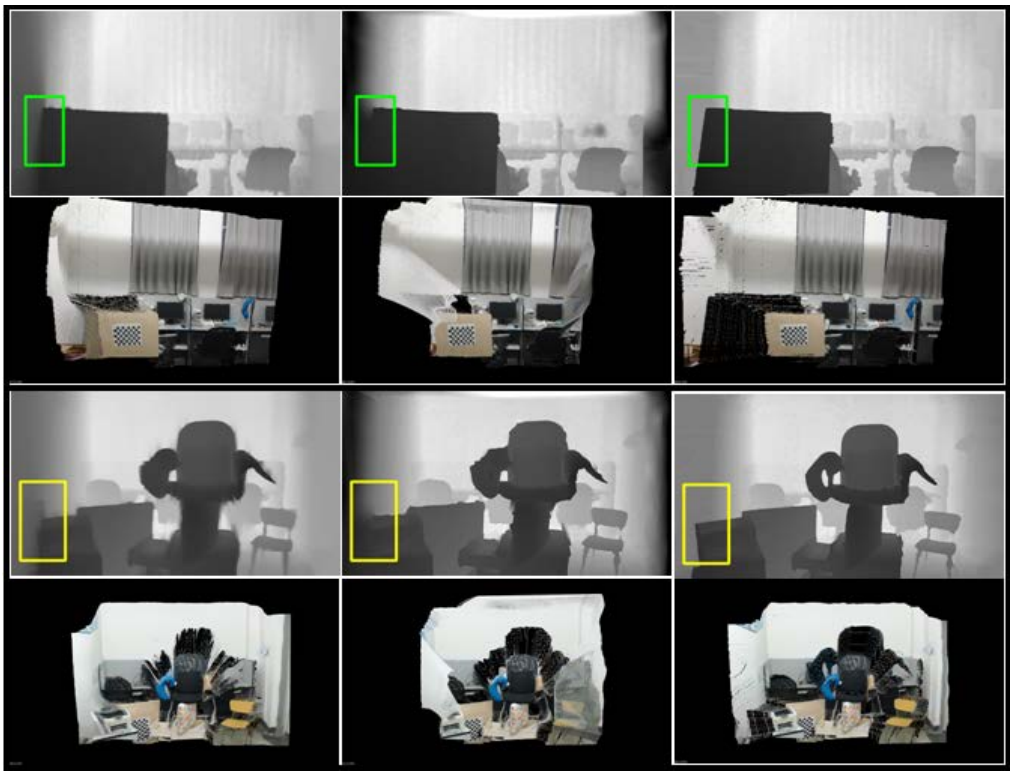


Figure3.The comparisons of the results produced by three different algorithms. Top row shows two different scenes with color image and low-resolution depth maps. The reconstructed results are shown in black box. From left to right column, each are reconstructed depth maps and point clouds generated by HQDMU[12], ATGV[13] and our method. In the first scene (top left), the low-resolution depth map does not cover whole rectangular plane and the depth between the plane and the wall differs too much as shown in green box. In the second scene (top right), the texture of the printer is almost the same as textur of the wall, while the depths are different as shown in yellow box. These experimental results indicate that our algorithm perfroms much better than other algorithms.

Conclusion

In this paper, we construct a dictionary that can express isotropic structure, repeating structure, and directional structure and piecewise smooth structure. The sparse linear combination of atoms in the dictionary can express various forms of depth maps. After a low-resolution of depth map that contains strong noises and lots of holes is captured, we use the EM algorithm to obtain a sparse coefficient vector, and then reconstruct the high-resolution depth map. We totally get rid of scene textures during the reconstruction. Thus, our algorithm can fill holes very well and reconstruct depth edges very accurately. The point clouds of real scenes experiment show that our algorithm is much better than other algorithms when viewed as. The quantitative experiments on Middlebury dataset demonstrate that our algorithm is superior to other conventional algorithms.

Acknowledgment

We express our sincere thanks to teachers and colleagues who give support and advice on this paper.

References

[1] J. Starck, M. Elad and D. Donoho, “Redundant multiscale transforms and their application for morphological component separation,” Advances in Imaging and Electron Physics, vol. 132, Jan. 2004, pp. 287-348, doi: 10.1016/S1076-5670(04)32006-9.

[2] M. Elad, J. Starck, P. Querre and D. L. Donoho, “Simultaneous cartoon and texture image inpainting using morphological component analysis(mca),” Applied and Computational Harmonic Analysis, vol. 19, Nov. 2005, pp. 340-358, doi:10.1016/j.acha.2005.03.005.

[3] D. L. Donoho, “Compressed sensing,” IEEE Transactions on Information Theory, vol. 52, Apr. 2006, pp. 1289-1306, doi:10.1109/TIT.2006.871582.

[4] A. M. Bruckstein, D. L. Donoho and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” SIAM review, vol. 51, Feb. 2009, pp. 34-81, doi:10.1137/060657704.

[5] M. F. Duarte, Y. C. Eldar. “Structured compressed sensing: From theory to applications,” Signal Processing, IEEE Transactions on, vol. 59, Jul. 2011, pp. 4053-4085, doi:10.1109/TSP.2011.2161982.

[6] M. W. Seeger, “Bayesian inference and optimal design for the sparse linear model,” The Journal of Machine Learning Research, vol. 9, Apr. 2008, pp. 759-813.

[7] J. Kopf, M. F. Cohen, D. Lischinski and M. Uyttendaele. “Joint bilateral upsampling,” ACM Transactions on Graphics (TOG). vol. 26, Aug. 2007, pp. 96, doi: 10.1145/1276377.1276497.

[8] D. Chan, H. Buisman, C. Theobalt and S. Thrun, “A noise-aware filter for real-time depth upsampling,” Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications-M2SFA2 2008.

[9] J. Kim, G. Jeon and J. Jeong, “Joint-adaptive bilateral depth map upsampling,” Signal Processing: Image Communication, vol. 29, Apr. 2014, pp. 506-513, doi:10.1016/j.image.2014.01.011.

[10] A. Adams, J. Baek and M. A. Davis, “Fast high-dimensional filtering using the permutohedral lattice”, Computer Graphics Forum, vol. 29. Jun. 2010, pp. 753-762, doi: 10.1111/j.1467-8659.2009.01645.x.

[11] J. Diebel and S. Thrun, “An application of markov random fields to range sensing,” Advances in neural information processing systems, 2005, pp. 291-298.

[12] J. Park, H. Kim, Y Tai, M. S. Brown and I. Kweon, “High quality depth map upsampling for 3d-tof cameras,” Proc. Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE Press, Nov. 2011, pp. 1623-1630, doi: 10.1109/ICCV.2011.6126423.

[13] D. Ferstl, C. Reinbacher, R. Ranftl, M. Rüther and H. Bischof, “Image guided depth upsampling using anisotropic total generalized variation,” Proc. Computer Vision (ICCV), 2013 IEEE International Conference on, IEEE Press, Dec. 2013, pp. 993-1000, doi:10.1109/ICCV.2013.127.

[14] C. Raposo, J. P. Barreto and U. Nunes, “Fast and accurate calibration of a kinect sensor,” Proc. 2013 International Conference on 3DTV-Conference, IEEE Press, Jun. 2013, pp. 342-349, doi: 10.1109/3DV.2013.52.

[15] C. Herrera, J. Kannala and J. Heikkilä, “Joint depth and color camera calibration with distortion correction,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 34, May. 2012, pp. 2058-2064, doi:10.1109/TPAMI.2012.125.

[16] J-L Starck, M. Elad and D. L. Donoho, “Image decomposition via the combination of sparse representations and a variational approach”, Image Processing, IEEE Transactions on, vol. 14, Oct. 2005, pp. 1570-1582, doi: 10.1109/TIP.2005.852206.

[17] S. Mallat, “Geometrical grouplets”, Applied and Computational Harmonic Analysis, vol. 26, Mar. 2009, pp. 161-180, doi: doi:10.1016/j.acha.2008.03.004.

[18] A. Maalouf, P. Carré, B. Augereau and C. Fernandez-Maloigne, “Inpainting using geometrical grouplets,” Proc. EUSIPCO2008, Aug. 2008, pp. 15-20.

[19] J-L Starck, E. J. Candès and D. L. Donoho, “The curvelet transform for image denoising,” Image Processing, IEEE Transactions on, vol. 11, Jun. 2002, pp. 670-684, doi: 10.1109/TIP.2002.1014998.
[20] E. Candes, L. Demanet, D. Donoho and L. Ying, “Fast discrete curvelet transforms,” Multiscale Modeling & Simulation, vol. 5, Sep. 2006, pp. 861-899, doi: 10.1137/05064182X.
[21] J. Yang, J. Wright, T. S. Huang and Y. Ma, “Image super-resolution via sparse representation,” Image Processing, IEEE Transactions on, vol. 19, May. 2010, pp. 2861-2873, doi: 10.1109/TIP.2010.2050625.
[22] Y. C. Eldar and G. Kutyniok, Compressed sensing: theory and applications. Cambridge University Press, 2012.
[23] Y. Li and S. Osher, “Coordinate descent optimization for l1 minimization with application to compressed sensing; a greedy algorithm,” Inverse Probl. Imaging, vol. 3, Aug. 2009, pp. 487-503, doi:10.3934/ipi.2009.3.487.
[24] H. Hirschmuller and D. Scharstein, “Evaluation of cost functions for stereo matching,” Proc. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on, IEEE Press, Jun. 2007, pp. 1-8, doi: 10.1109/CVPR.2007.383248.