PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Accuracy analysis of 3D object shape recovery using depth filtering algorithms

Alexey Ruchay, Konstantin Dorofeev, Anastasia Kober, Vladimir Kolpakov, Vsevolod Kalschikov

Alexey Ruchay, Konstantin Dorofeev, Anastasia Kober, Vladimir Kolpakov, Vsevolod Kalschikov, "Accuracy analysis of 3D object shape recovery using depth filtering algorithms," Proc. SPIE 10752, Applications of Digital Image Processing XLI, 1075221 (17 September 2018); doi: 10.1117/12.2319907

SPIE.

Event: SPIE Optical Engineering + Applications, 2018, San Diego, California, United States

Accuracy analysis of 3D object shape recovery using depth filtering algorithms

Alexey Ruchay[a,b], Konstantin Dorofeev[b], Anastasia Kober[a], Vladimir Kolpakov[a], Vsevolod Kalschikov[b]

[a]Federal Research Centre of Biological Systems and Agro-technologies of the Russian Academy of Sciences, Russian Federation;
[b]Department of Mathematics, Chelyabinsk State University, Russian Federation.

ABSTRACT

In this paper, we estimate the accuracy of 3D object reconstruction using depth filtering and data from a RGB-D sensor. Depth filtering algorithms carry out inpainting and upsampling for defective depth maps from a RGB-D sensor. In order to improve the accuracy of 3D object reconstruction, an efficient and fast method of depth filtering is designed. Various methods of depth filtering are tested and compared with respect to the reconstruction accuracy using real data. The presented results show an improvement in the accuracy of 3D object reconstruction using depth filtering from a RGB-D sensor.

Keywords: RGB-D, depth map, filtering, 3D object shape recovery.

1. INTRODUCTION

The 3D reconstruction of objects is a popular topic, with applications in the field of architecture, games, agriculture, and medicine. The 3D reconstruction has many applications in object retrieval, scene understanding, object recognition, object tracking, navigation, engineering, visualization, human-computer interaction and virtual maintenance.[1–5]

Surface 3D reconstruction of real-world objects by using RGB-D sensors such as the Kinect is one of the most important topics in computer vision.[6–8] These cameras can provide a high-resolution RGB color image with a depth map of the environment.

We are interested in improving the quality of the depth map because the depth map is characterized by piecewise smooth regions bounded by sharp object boundaries. It means that the depth value varies discontinuity, and a small error around object boundary may lead to significant ringing artifacts in rendered views. Also, the depth map provided by a RGB-D camera is often noisy due to imperfections associated with infrared light reflections, and missing pixels without any depth value appear as black holes in depth maps.

The noise and holes can greatly affect the performance of 3D reconstruction, therefore, the noise-reduction and hole-filling enhancement algorithms are intended to serve as a pre-processing step for 3D reconstruction systems using Kinect cameras.[9–12]

In order to fill small holes and to eliminate noise, the median and binomial filters are used.[13–16] Moreover, the use of the color information in the point correspondence process avoids false positives matches and, therefore, leads to a more reliable registration. Note that by adjusting iterative closest point (ICP) algorithm and reconstruction parameters it is possible to improve the registration and appearance of details that are invisible with just one scan due to the sensor limited precision.

Further author information:
A.R.: E-mail: ran@csu.ru
K.D.: E-mail: kostuan1989@mail.ru
A.K.: E-mail: anastejsha11@mail.ru
Vl.K.: E-mail: vkolpakov056@yandex.ru
Vs.K.: E-mail: vkalschikov@gmail.com

Applications of Digital Image Processing XLI, edited by Andrew G. Tescher, Proc. of SPIE Vol. 10752, 1075221 · © 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2319907

Traditional 3D depth denoising methods are focused on fusing multiple consecutive noisy depths to get a higher quality:[17] spatial-temporal denoising approaches;[18,19] a method based on the correlation between aligned color and depth frames provided by such sensors;[20] a deep-learning based approach which makes use of aligned gray images to denoise depth data.[21] However, in many cases, color or gray images aligned with the depth frames may not be available.

Enhancing the quality of the depth map obtained with a single depth frame is an increasingly popular research topic: median filtering based on adaptive weighted Gaussian;[22] wavelet denoising;[23] bilateral filter;[24] total variation regularization;[25] non-Local-Mean(NLM) method.[26]

While the existence of ground-truth data in image denoising literature facilitates quantitative evaluation (e.g., PSNR, MSE, etc) of the final denoised images, it is not the case in depth. We are aware of the literature which pursues a depth data denoising approach similar to approaches from image denoising.

In a common approach of noise reduction in the quality assessment, the authors offered that the raw depth map represents the ground truth, added an artificial noise such as additive or impulse, and then proposed a method to remove the noise.[27] Although this common approach could be used for quantitative comparison, wherein proposed methods reduce only the artificial noise but not the original signal-dependent noise contained in the raw depth. Therefore, we consider denoising depth algorithms for 3D object reconstruction. We propose a signal-dependent denoising approach relies solely on a single depth frame as the input. Contrary to the custom in the literature, we consider the raw depth map as noisy data and we evaluate the performance of the denoising methods based on the enhancement achieved in the accuracy of 3D object reconstruction.

Some denoising methods are not designed to clean the signal-dependent coarse noise contained in the raw depth map. Therefore, our main goal is to evaluate the denoising methods to enhance reconstruction accuracy which depends on the quality of the captured raw depth map. Also, we propose a cascade mechanism to improve 3D object reconstruction which is a consistent application of several methods of noise reduction.

The paper is organized as follows. Section 2 discusses related approaches of denoising depth methods. In Section 3, we describe the proposed cascade algorithm. Computer simulation results are provided in Section 4. Finally, Section 5 summarizes our conclusions.

2. RELATED WORK

A novel and effective divide-and-conquer method for handling disocclusion of the synthesized image is presented,[28] where a binary mask is used to mark the disocclusion region, the depth pixels located in the disocclusion region are modified by a linear interpolation process, and a median filtering is adapted to remove the isolated depth pixels.

A method for depth image enhancement has been developed for reducing noise and filling holes.[29] This method does not utilize any color information to achieve the depth image or map recovery in a computationally efficient manner. This method can be deployed as a preprocessing step in action recognition systems that utilize depth images.

A method for correction of depth maps in order to maintain and improve the quality of synthesized views was proposed.[30] The method uses a three-phase depth map correction, including eliminating anomalies, segmentation, amendment and finally inter-frame and intra-frame filtering.

The paper[25] presents a depth filtering scheme to reduce the distortion by exploiting the temporal information and color information present in a scene. The proposed algorithm increases the point to point correspondence of the depth map with its RGB image by a large extent.

A novel depth map filtering algorithm for high-efficiency 3D video coding was proposed.[31] The depth map filtering consists of a newly designed nonlinear down/upsampling filtering and a depth reconstruction multilateral filtering, which considers a spatial resolution, boundary similarity, and coding artifacts features.

A new 3D collaborative filtering in graph Fourier transform (GFT) domain was proposed.[27] The depth image denoising has the following steps: clustering similar patches and stack them together into 3D groups; constructing a full-connected graph describing the correlations among all the pixels for each group; transforming

each 2D patch to a GFT domain, where the GFT matrix can be computed from the constructed graph; applying a 1D transform across the grouped patches in the same way; formulating an iterative quadratic programming problem by using the smoothness prior in the 3D GFT domain.

The paper[32] uses weighted mode filter and joint bilateral filter for depth denoising where the joint bilateral kernel provides an optimal solution with the help of the joint histogram. Also, an adaptive method to denoise depth data in order to enhance object recognition was proposed[33] based on Differential Histogram of Normal Vectors (DHONV) features along with a linear SVM.

3. PROPOSED ALGORITHM

In this section, we describe the proposed depth denoising algorithm based on cascade mechanism.

We have considered the following depth denoising algorithms which are suitable for only depth data with/without color information:

- Joint Bilateral Filter (JBF):[34] It is a non-linear, edge-preserving, and noise-reducing smoothing filter for depth denoising. It replaces the intensity of each pixel with a weighted average of intensity values from nearby pixels.
- Joint Bilateral Upsampling (JBU):[24] It is a spatial filter (typically a truncated Gaussian) to the low-resolution solution, while a similar range filter is jointly applied on the full resolution image.
- Noise-aware Filter (NF):[35] It is an adaptive multi-lateral upsampling filter that takes into account the inherent noisy nature of real-time depth data.
- Weight Mode Filter (WMF):[36] It is a weighted mode filtering method based on a joint histogram, where the weight based on color similarity between reference and neighboring pixels on the color image is computed and then used for counting each bin on the joint histogram of the depth map, and a final solution is determined by seeking a global mode on the histogram.
- Anisotropic Diffusion (AD):[37] It is a filter based on the heat diffusion framework, where the pixels with known depth values are treated as the heat sources and the depth enhancement is performed via diffusing the depth from these sources to unknown regions. The diffusion conductivity is designed in terms of the guidance color image so that a linear anisotropic diffusion problem is formed. The enhancement is achieved efficiently by solving a sparse linear system.
- Markov Random Field (MRF):[38] It is a filter based on Markov Random Field that integrated high-resolution image data into low-resolution range data, to recover range data at the same resolution as the image data.
- Markov Random Field(Second Order Smoothness) (MRFS):[39] It is a filter based on Markov Random Field that integrated high-resolution image data into low-resolution range data, where the large sparse linear system generated by the second-order MRF upsampling model is solved.
- Markov Random Field(Kernel Data Term) (MRFK):[39] It is a filter based on Markov Random Field that integrated high-resolution image data into low-resolution range data, where the large sparse linear system generated by the MRF upsampling model is solved.
- Markov Random Field(Tensor) (MRFT):[39] It is a filter based on Markov Random Field that integrated high-resolution image data into low-resolution range data, where the large sparse linear system generated by incorporating tensor into the MRF upsampling model is solved.
- Layered Bilateral Filter (LBF):[40] It a new post-processing step to enhance the spatial resolution of range images up to 100x with a registered and potentially high-resolution color image as the reference.
- Kinect depth normalization (KDN):[41] It is a filter for filling the depth images, where this filling process is done by replacing the zero value pixels with the statistical model of the surrounding 25 pixels. Using a statistical mode returns sharper edges than using statistical mean values since mode only inserts the maximum occurring value in the surrounding 25 pixels to the centre pixel.

- Roifill filter (RF):[42] It is a filter of filling in a specified region of interest (ROI) polygon in a grayscale image, which allows erasing elements in an image. The roifill function smoothly interpolates inward from the pixel values on the boundary of the polygon by solving the Laplace's equation.
- Median filter (MF): It is a standard median filter.
- Bilateral Filter (BF): It is a standard Bilateral Filter.
- Okada filter (OF):[43] It is a filter for reducing shot noise in electrophysiological data and photographs, which can be described using a single continuous differentiable equation based on the logistic function and is thus mathematically tractable.

A common algorithm of 3D reconstruction consists of the following steps:

1. Registration a RGB and depth data.
2. Use a depth denoising algorithm: JBF, JBU, NF, WMF, AD, MRF, MRFS, MRFK, MRFT, LBF, KDN, RF, MF, BF, OF.
3. Make a point cloud PC_i and PC_{i-1} using denoising depth data.
4. Detection and matching of keypoints in PC_i and PC_{i-1} with the keypoint detection algorithm SIFT.[44]
5. Remove outliers with correspondence rejectors RANSAC.[44]
6. Count transformation matrix with ICP using the associate 3D points of the inliers.
7. Apply hybrid approach that combines ICP odometry with the RGB-D odometry.
8. Adding results to a general model (dense 3D map).

We propose the following depth denoising algorithm based on cascade mechanism:

1. Registration a RGB and depth data.
2. Use a depth denoising algorithm: JBF, JBU, NF, WMF, AD, MRF, MRFS, MRFK, MRFT, LBF, KDN, RF, MF, BF, OF.
3. Use other depth denoising algorithm: JBF, JBU, NF, WMF, AD, MRF, MRFS, MRFK, MRFT, LBF, KDN, RF, MF, BF, OF.
4. Make a point cloud PC_i and PC_{i-1} using denoising depth data.
5. Detection and matching of keypoints in PC_i and PC_{i-1} with the keypoint detection algorithm SIFT.
6. Remove outliers with correspondence rejectors RANSAC.
7. Count transformation matrix with ICP using the associate 3D points of the inliers.

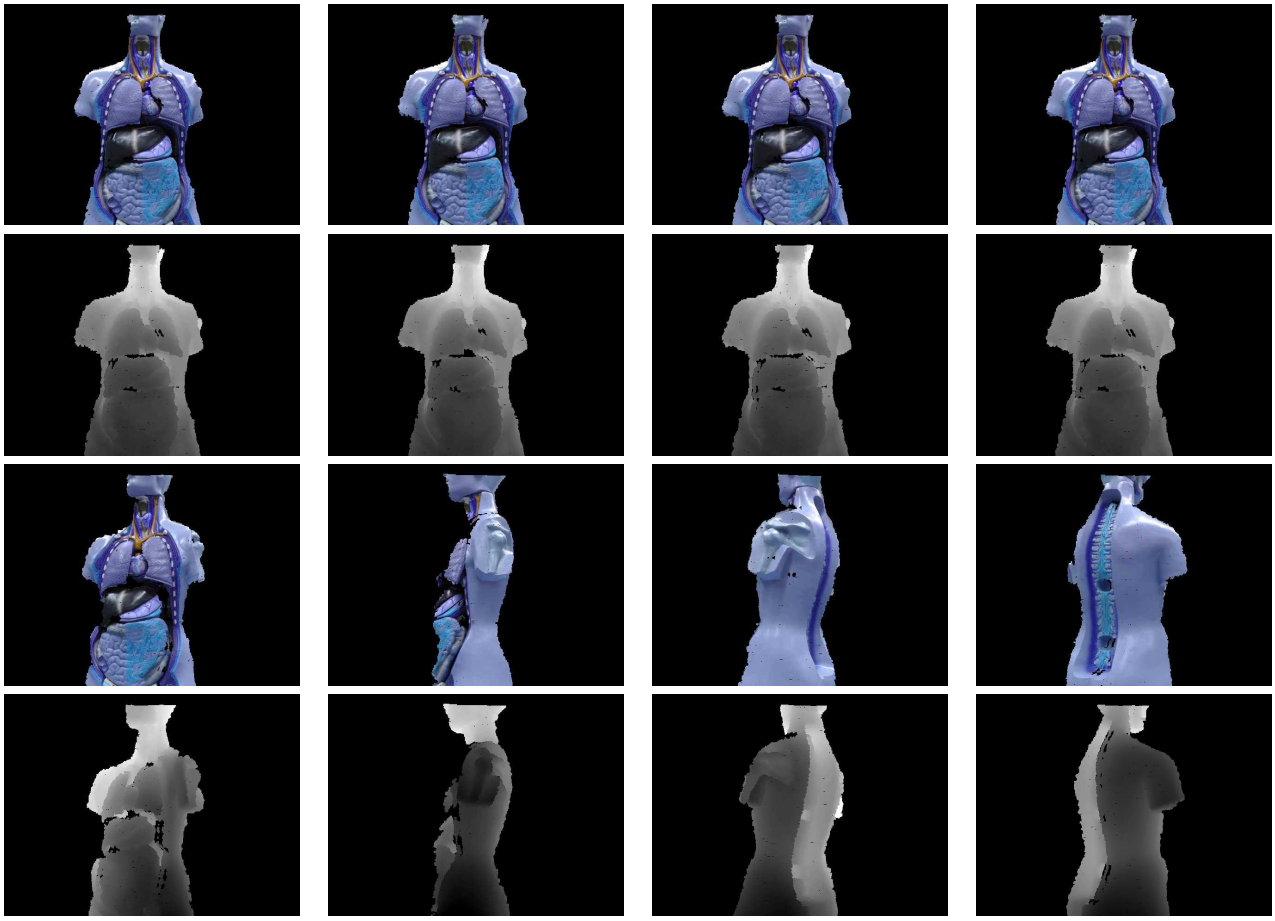

Figure 1. The RGB images and depth maps of an anatomy with 1, 2, 4, 10, 18, 28, 42, 60 step are taken by a Kinect sensor.

4. COMPUTER SIMULATION

In this section, computer simulation results of the accuracy of 3D object reconstruction based on the proposed depth denoising algorithm with a cascade mechanism using real data are presented and discussed.

Raw 3D captured depth data are noisy and the digital ground-truth data are not available. Therefore, we evaluate the performance of our proposed denoising filter against other state-of-the-art filters based on the enhancement of reconstruction accuracy achieved by each filter. We have experimental results for evaluation of the performance of the ICP algorithm with a keypoint detection algorithm for object 3D reconstruction. The metric of evaluation is the root mean square error (RMSE) of measurements.

In our experiments, we select 15 different depth denoising algorithms which are widely cited and used in comparison. The experiments are carried out on a PC with Intel(R) Core(TM) i7-4790CPU @ 3.60 GHz and 16 GB memory.

RGB-D datasets for different applications including object reconstruction and 3D-simultaneous localization and mapping were proposed.[45] We choose the special RGB-D datasets.[46]

To evaluate the performance of 3D object reconstruction based on the proposed depth denoising algorithm with cascade mechanism in our experiments, we carried out the point cloud fusion and 3D reconstruction of an anatomy from dataset.[46] Fig. 1 shows RGB images and depth maps of an anatomy taken with a step of 1, 2, 4, 10, 18, 28, 42, 60.

Table 1. Results of measurements using a common ICP algorithm with JBF, JBU, NF, WMF, AD, MRF, MRFS, MRFK, MRFT, LBF, RF, KDN, MF, BF, OF depth denoising algorithms (DDA) for each pair of 1,2,4,10,18,28,42 and 60 step. This table presents RMSE, an average time of processing in sec. (Time).

DDA	1–2	2–4	4–10	10–18	18–28	28–42	42–60	Time
Without	4.68e-04	5.87e-04	5.34e-04	1.05e-03	-	-	2.41e-02	0
JBF	5.58e-04	6.85e-04	6.40e-04	1.04e-03	1.94e-02	1.59e-02	2.18e-03	8.53
JBU	5.71e-04	6.25e-04	6.21e-04	1.04e-03	2.10e-02	8.64e-03	2.23e-03	5.18
NF	5.86e-04	6.03e-04	6.03e-04	1.03e-03	2.09e-02	8.47e-03	2.23e-03	7.18
WMF	-	-	-	-	-	-	-	12.8
AD	-	-	-	-	-	-	-	8.96
MRF	-	-	-	-	-	-	-	3.67
MRFS	4.65e-04	4.41e-04	4.73e-04	9.70e-04	1.91e-02	7.17e-03	2.15e-03	4.14
MRFK	1.87e-03	2.79e-03	8.41e-03	9.49e-04	1.91e-02	1.23e-02	4.21e-03	28.9
MRFT	9.56e-04	1.36e-02	-	-	-	-	-	152.1
LBF	1.44e-03	1.76e-03	2.42e-03	1.64e-04	5.40e-03	1.65e-02	5.59e-03	98.5
KDN	1.82e-03	7.92e-04	1.37e-03	1.05e-03	2.28e-02	7.84e-03	3.03e-03	1.59
RF	5.12e-04	6.89e-04	6.34e-04	1.05e-03	1.76e-02	6.66e-03	2.99e-03	0.78
MF	4.54e-04	5.47e-04	4.70e-04	1.03e-03	-	-	2.41e-02	0.10
BF	3.85e-04	4.11e-04	4.00e-04	9.96e-04	-	-	2.41e-02	2.83
OF	1.49e-03	1.69e-03	1.27e-03	1.04e-03	2.20e-02	1.32e-02	2.21e-02	0.21

Table 2. Results of measurements using the ICP algorithm with the proposed depth denoising algorithm based on cascade RF, MF, BF algorithms (DDA) for each pair of 1,2,4,10,18,28,42 and 60 step. This table presents RMSE, an average time of processing in sec. (Time).

DDA	1–2	2–4	4–10	10–18	18–28	28–42	42–60	Time
Without	4.68e-04	5.87e-04	5.34e-04	1.05e-03	-	-	2.41e-02	0.00
BF, RF	1.89e-03	1.95e-03	2.05e-03	3.42e-03	1.04e-02	1.20e-02	1.38e-02	4.10
RF, BF	4.42e-04	5.71e-04	5.30e-04	9.90e-04	1.75e-02	6.26e-03	2.98e-03	3.76
BF, MF	1.84e-03	2.19e-03	2.11e-03	3.40e-03	-	1.54e-02	2.09e-02	2.87
MF, BF	1.85e-03	2.03e-03	1.96e-03	3.81e-03	-	1.53e-02	2.08e-02	2.82
RF, MF	5.11e-04	6.58e-04	5.85e-04	1.03e-03	1.76e-02	6.66e-03	3.05e-03	0.80
MF, RF	5.07e-04	6.46e-04	5.71e-04	1.03e-03	1.76e-02	6.71e-03	3.10e-03	0.76

Corresponding RMSE values calculated for each pair of 1, 2, 4, 10, 18, 28, 42, 60 step in the ICP algorithm with JBF, JBU, NF, WMF, AD, MRF, MRFS, MRFK, MRFT, LBF, KDN, RF, MF, BF, OF depth denoising algorithms are shown in Table 1.

After denoising WMF, AD, MRF filters the ICP algorithm cannot normally work because of too smoothed depths. The quality of depth denoising we can also evaluate visually looking at the restored point cloud. Fig. 2 shows the 3D point clouds of an anatomy after denoising JBF, JBU, NF, MRFS, MRFK, RF, KDN, MF, BF filters. RF, MF, and BF yield the best result in terms of RMSE and visual evaluation among all depth denoising algorithms.

Based on the results of previous experiments, we chose the three best RF, MF and BF depth denoising algorithms for our proposed algorithm based on cascade mechanism, which first uses one of the RF, MF, BF algorithms and then another algorithm. Corresponding RMSE values calculated for each pair of 1,2,4,10,18,28,42 and 60 step in the ICP algorithm with the proposed algorithm are shown in Table 2. The cascade with MF and RF depth denoising algorithm yields the best result in term of RMSE and speed. Fig. 3 shows the 3D point clouds of an anatomy after the proposed depth denoising algorithm based on cascade MF and RF filters.

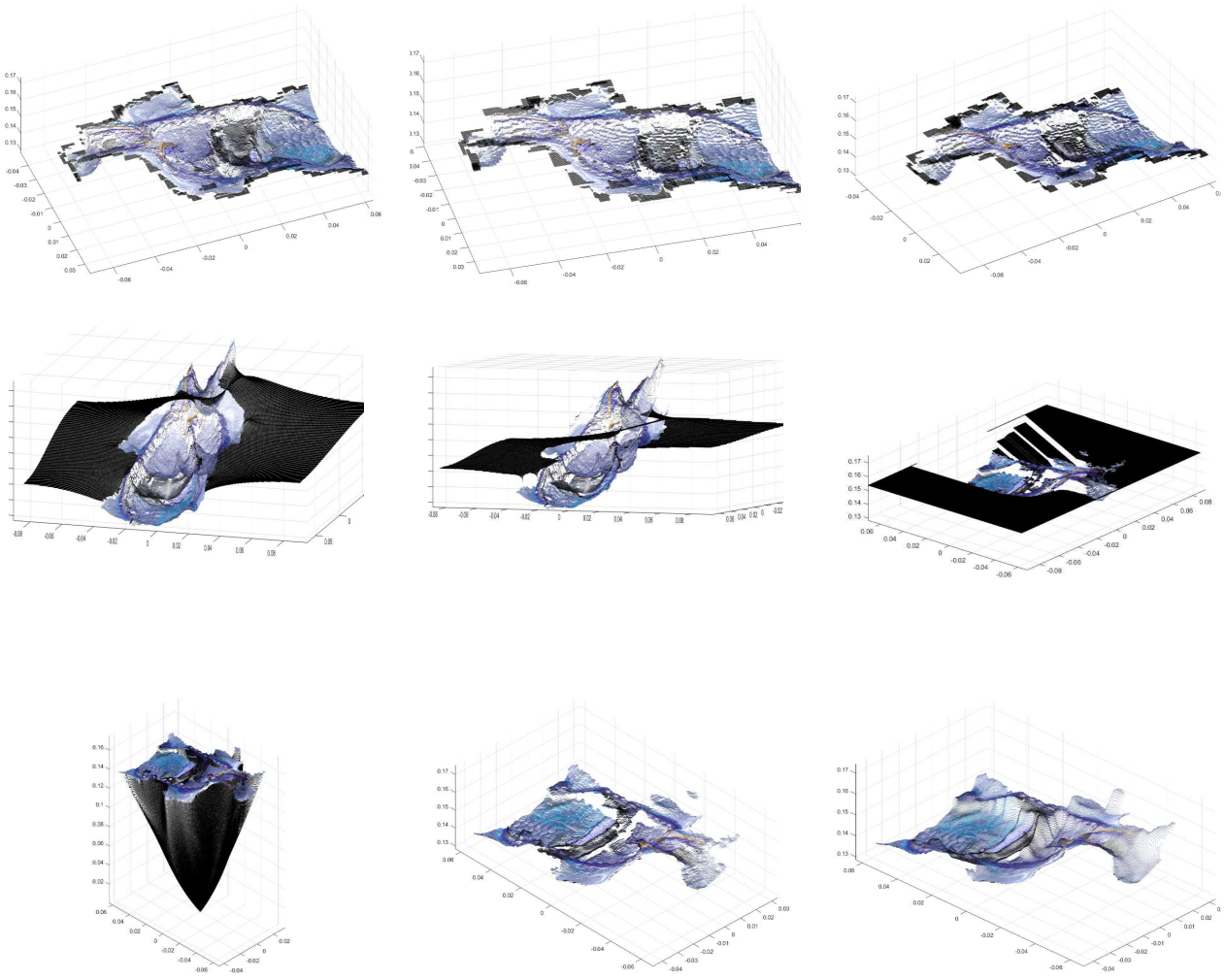

Figure 2. The restored point clouds of an anatomy after denoising JBF, JBU, NF, MRFS, MRFK, RF, KDN, MF, BF filters (from left to right from top to bottom).

5. CONCLUSION

In this paper, we presented the depth denoising algorithm based on a cascade mechanism for object 3D reconstruction. The proposed depth denoising algorithm uses RF and MF algorithms. We evaluated the performance of the ICP algorithm with the proposed depth denoising algorithm for object 3D reconstruction using real data. The experiment has shown that the ICP algorithm with the proposed denoising algorithm based on cascade mechanism for object 3D reconstruction increases the accuracy in terms of RMSE.

ACKNOWLEDGMENTS

This work was supported by the Russian Science Foundation, grant no. 17-76-20045.

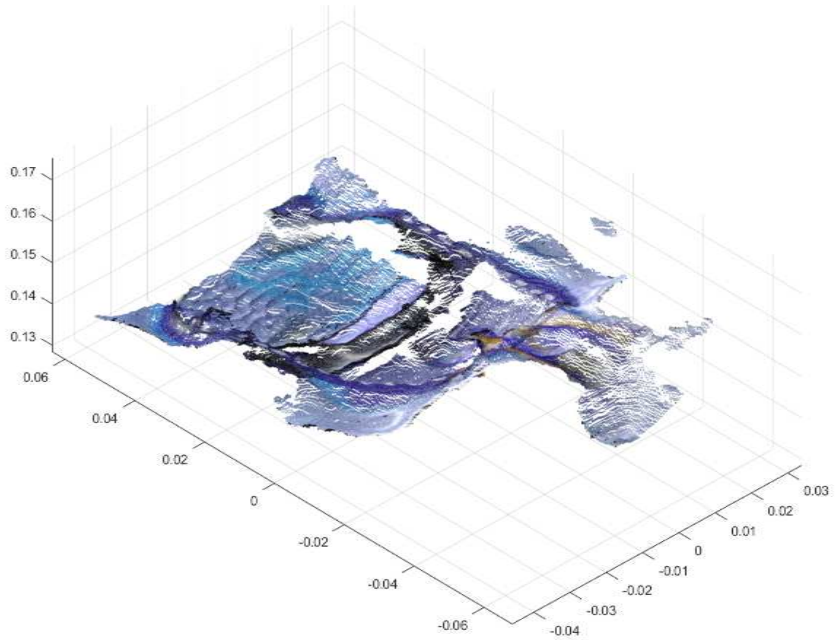

Figure 3. The restored point clouds of an anatomy after the proposed depth denoising algorithm based on cascade with MF and RF filters.

REFERENCES

[1] Echeagaray-Patron, B. A. and Kober, V., "3d face recognition based on matching of facial surfaces," in [*Proc. SPIE*], **9598**, 95980V–8 (2015).

[2] Echeagaray-Patron, B. A. and Kober, V., "Face recognition based on matching of local features on 3d dynamic range sequences," in [*Proc.SPIE*], **9971**, 9971–6 (2016).

[3] Echeagaray-Patrón, B. A., Kober, V. I., Karnaukhov, V. N., and Kuznetsov, V. V., "A method of face recognition using 3d facial surfaces," *Journal of Communications Technology and Electronics* **62**(6), 648–652 (2017).

[4] Ruchay, A., Kober, V., and Yavtushenko, E., "Fast perceptual image hash based on cascade algorithm," in [*Proc.SPIE*], **10396**, 10396–7 (2017).

[5] Ruchay, A., Kober, V., and Gonzalez-Fraga, J. A., "Reliable recognition of partially occluded objects with correlation filters," *Mathematical Problems in Engineering* **2018**, 8284123 (2018).

[6] Tihonkih, D., Makovetskii, A., and Kuznetsov, V., "A modified iterative closest point algorithm for shape registration," in [*Proc. SPIE*], **9971**, 99712D–8 (2016).

[7] Nikolaev, D., Tihonkih, D., Makovetskii, A., and Voronin, S., "An efficient direct method for image registration of flat objects," in [*Proc.SPIE*], **10396**, 10396–8 (2017).

[8] Gonzalez-Fraga, J. A., Kober, V., Diaz-Ramirez, V. H., Gutierrez, E., and Alvarez-Xochihua, O., "Accurate generation of the 3d map of environment with a rgb-d camera," in [*Proc.SPIE*], **10396**, 10396–7 (2017).

[9] Voronin, S. M., Makovetsky, A. Y., Kober, V. I., and Karnaukhov, V. N., "Properties of exact solutions to the problem of regularization of the total variation of one-variable functions," *Journal of Communications Technology and Electronics* **60**(12), 1356–1359 (2015).

[10] Makovetskii, A. and Kober, V., "Analysis of the gradient descent method in problems of the signals and images restoration," *Pattern Recognition and Image Analysis* **25**(1), 53–59 (2015).

[11] Tihonkih, D., Makovetskii, A., and Voronin, A., "A modified iterative closest point algorithm for noisy data," in [*Proc.SPIE*], **10396**, 10396–7 (2017).
[12] Makovetskii, A., Voronin, S., and Kober, V., "An efficient algorithm for total variation denoising," in [*Analysis of Images, Social Networks and Texts*], 326–337, Springer International Publishing, Cham (2017).
[13] Ruchay, A. and Kober, V., "Clustered impulse noise removal from color images with spatially connected rank filtering," in [*Proc. SPIE*], **9971**, 99712Y–99712Y–10 (2016).
[14] Ruchay, A. and Kober, V., "Removal of impulse noise clusters from color images with local order statistics," in [*Proc.SPIE*], **10396**, 10396–10 (2017).
[15] Ruchay, A. and Kober, V., "Impulsive noise removal from color video with morphological filtering," in [*Proc.SPIE*], **10396**, 10396–9 (2017).
[16] Ruchay, A. and Kober, V., "Impulsive noise removal from color images with morphological filtering," in [*Analysis of Images, Social Networks and Texts*], 280–291, Springer International Publishing, Cham (2018).
[17] Boubou, S., Narikiyo, T., and Kawanishi, M., "Adaptive filter for denoising 3d data captured by depth sensors," in [*2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*], 1–4 (2017).
[18] Fu, J., Wang, S., Lu, Y., Li, S., and Zeng, W., "Kinect-like depth denoising," in [*2012 IEEE International Symposium on Circuits and Systems*], 512–515 (2012).
[19] Lin, B. S., Chou, W. R., Yu, C., Cheng, P. H., Tseng, P. J., and Chen, S. J., "An effective spatial-temporal denoising approach for depth images," in [*2015 IEEE International Conference on Digital Signal Processing (DSP)*], 647–651 (2015).
[20] Milani, S. and Calvagno, G., "Correction and interpolation of depth maps from structured light infrared sensors," *Signal Processing: Image Communication* **41**, 28–39 (2016).
[21] Zhang, X. and Wu, R., "Fast depth image denoising and enhancement using a deep convolutional network," in [*2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*], 2499–2503 (2016).
[22] Frank, M., Plaue, M., and Hamprecht, F. A., "Denoising of continuous-wave time-of-flight depth images using confidence measures," *Optical Engineering* **48**(7), 48–13 (2009).
[23] Moser, B., Bauer, F., Elbau, P., Heise, B., and Schoner, H., "Denoising techniques for raw 3d data of tof cameras based on clustering and wavelets," in [*Proc.SPIE*], **6805**, 6805–12 (2008).
[24] Kopf, J., Cohen, M. F., Lischinski, D., and Uyttendaele, M., "Joint bilateral upsampling," *ACM Trans. Graph.* **26**(3) (2007).
[25] Bhattacharya, S., Venkatesh, K. S., and Gupta, S., "Depth filtering using total variation based video decomposition," in [*2015 Third International Conference on Image Information Processing (ICIIP)*], 23–26 (2015).
[26] Georgiev, M., Gotchev, A., and Hannuksela, M., "Real-time denoising of tof measurements by spatio-temporal non-local mean filtering," in [*2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*], 1–6 (2013).
[27] Chen, R., Liu, X., Zhai, D., and Zhao, D., "Depth image denoising via collaborative graph fourier transform," in [*Digital TV and Wireless Multimedia Communication*], 128–137, Springer Singapore (2018).
[28] Lei, J., Zhang, C., Wu, M., You, L., Fan, K., and Hou, C., "A divide-and-conquer hole-filling method for handling disocclusion in single-view rendering," *Multimedia Tools and Applications* **76**(6), 7661–7676 (2017).
[29] Liu, S., Chen, C., and Kehtarnavaz, N., "A computationally efficient denoising and hole-filling method for depth image enhancement," in [*Proc.SPIE*], **9897**, 9897–9 (2016).
[30] Pourazad, M. T., Zhou, D., Lee, K., Karimifard, S., Ganelin, I., and Nasiopoulos, P., "Improving depth map compression using a 3-phase depth map correction approach," in [*2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*], 1–6 (2015).
[31] Zhang, Q., Chen, M., Zhu, H., Wang, X., and Gan, Y., "An efficient depth map filtering based on spatial and texture features for 3d video coding," *Neurocomputing* **188**, 82–89 (2016).
[32] Fu, M. and Zhou, W., "Depth map super-resolution via extended weighted mode filtering," in [*2016 Visual Communications and Image Processing (VCIP)*], 1–4 (2016).

[33] Boubou, S., Narikiyo, T., and Kawanishi, M., "Adaptive filter for denoising 3d data captured by depth sensors," in [*2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*], 1–4 (2017).

[34] Petschnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M., and Toyama, K., "Digital photography with flash and no-flash image pairs," *ACM Transactions on Graphics* **23**(3), 664–672 (2004).

[35] Chan, D., Buisman, H., Theobalt, C., and Thrun, S., "A noise-aware filter for real-time depth upsampling," in [*In workshop on multi-camera and multi-modal sensor fusion algorithms and applications*], (2008).

[36] Min, D., Lu, J., and Do, M. N., "Depth video enhancement based on weighted mode filtering," *IEEE Transactions on Image Processing* **21**(3), 1176–1190 (2012).

[37] Liu, J. and Gong, X., "Guided depth enhancement via anisotropic diffusion," in [*Proceedings of the 14th Pacific-Rim Conference on Advances in Multimedia Information Processing — PCM 2013 - Volume 8294*], 408–417 (2013).

[38] Diebel, J. and Thrun, S., "An application of markov random fields to range sensing," in [*Proceedings of the 18th International Conference on Neural Information Processing Systems*], *NIPS'05*, 291–298 (2005).

[39] Harrison, A. and Newman, P., "Image and sparse laser fusion for dense scene reconstruction," in [*Field and Service Robotics*], 219–228, Springer Berlin Heidelberg, Berlin, Heidelberg (2010).

[40] Yang, Q., Yang, R., Davis, J., and Nister, D., "Spatial-depth super resolution for range images," in [*2007 IEEE Conference on Computer Vision and Pattern Recognition*], 1–8 (2007).

[41] Newcombe, R. A., Izadi, S., Hilliges, O., Kim, D., Davison, A. J., Kohli, P., Shotton, J., Hodges, S., and Fitzgibbon, A., "Kinectfusion: Real-time dense surface mapping and tracking," in [*IEEE ISMAR*], (2011).

[42] Fuhrmann, S. and Goesele, M., "Fusion of depth maps with multiple scales," *ACM Trans. Graph.* **30**(6), 148:1–148:8 (2011).

[43] Okada, M., Ishikawa, T., and Ikegaya, Y., "A computationally efficient filter for reducing shot noise in low s/n data," *PLoS ONE* **11**(6), e0157595 (2016).

[44] Rusu, R. B. and Cousins, S., "3d is here: Point cloud library (pcl)," in [*2011 IEEE International Conference on Robotics and Automation*], 1–4 (2011).

[45] Cai, Z., Han, J., Liu, L., and Shao, L., "Rgb-d datasets using microsoft kinect or similar sensors: a survey," *Multimedia Tools and Applications* **76**(3), 4313–4355 (2017).

[46] Lee, K. and Nguyen, T. Q., "Realistic surface geometry reconstruction using a hand-held RGB-D camera," *Mach. Vis. Appl.* **27**(3), 377–385 (2016).