Proceedings of the 2017 IEEE
International Conference on Information and Automation (ICIA)
Macau SAR, China, July 2017

Point-cloud Compression Using Data Independent Method - A 3D Discrete Cosine Transform Approach

Lujia Wang*, Luyu Wang†, Yinting Luo‡, Ming Liu‡
* Center for Cloud Computing, Shenzhen Institutes of Advanced Technology,
Chinese Academy of Sciences, Shenzhen, China
† Department of Endocrine, Huadu District People's Hospital, Southern Medical University, Guangzhou, China
‡ RAM-LAB, Department of Electronics and Computer Engineering,
The Hong Kong University of Science and technology, Hong Kong
lj.wang1@siat.ac.cn, rugga.wang@gmail.com, eelium@ust.hk

***Abstract*—Point-cloud is a widely used representation for objects and scenes. It generally consists of a large amount of 3D coordinates of points describing reflective surfaces. A subtle problem is that the number of points is usually so large that real-time transmission and efficient storage is not feasible. In this work, we propose to use 3D Discrete Cosine Transform (3D-DCT) to compress these two typical categories of data, namely point-cloud data extracted from objects and environments (i.e. 3D maps). Experimental results show that the proposed method leads to high compression ratio and flexible reconstruction behaviors comparing with other related methods.**

I. Introduction

A. Motivation

A wide spectrum of perception applications in robotics are associated with point-cloud-based modeling [1], [2], [3]. Point-cloud is an important format of representation for high level reasoning, such as sensor fusion-based recognition [4], [5]. After the raw points are acquired, the dense information needs to be processed and analyzed. A substantial amount of memory storage is to be occupied because of the high density of points, which makes the real-time retrieval [6], [7] and efficient storage easily intractable without compression.

Several works have studied this problem, among those the Octree encoding [8] is the most prevalently applied in point-cloud compression by recursively dividing the 3D data into eight octants. The similar concept is also adopted in the widely used Point Cloud Library (PCL) [1]. These methods are regardless of the data structure and can be easily generalized. We denote these methods as *structure-independent algorithms*. However, they usually ignore the fact that the inherent structure of different raw data would provide hints for better compression. For example, the point-cloud that describes a flat plane has much potential to be compressed with high compression ratio than that describes a complex object, even though they may contain similar number of points or similar coverage in volume. Therefore, structure-independent algorithms may waste lots of space to store the structured data and usually sensitive to the orientation.

Another type of approaches is relying on the results of structure analysis of the point-cloud, namely *structure-dependent algorithms*. They are usually based on structure analysis. For example, plane extraction [9], [10] and tensor analysis [11], [12], [13] are usually adopted to first analyze the structure of a scene, then descriptions are constructed accordingly. In structured environments or for highly-structured objects, these methods are very efficient, since parametric representations of planes or surfels can greatly reduce the number of points to be stored. However, they generally could not work in unstructured environments where the planes are not detectable or are costly to be efficiently represented. It means these methods are efficient for structured data but not for unstructured data. Recently, there are some other works are proposed [14], [15], [16]. Authors in [14] used depth image of point-cloud for compression, which is not suitable for applications without perceptive views. Authors in [15] can achieve real-time point-cloud compression, but they used 2D point-cloud data which is different from this paper. Authors in [16] used conventional image processing method DCT similar to this paper, but they only use normal image data which is different from data derived from robotic applications.

By analysing the pros and cons of these two typical types of compression techniques, we find the key is to determine the dynamics of the point-clouds along the surface. For unstructured point-clouds, the dynamics is usually fast (high frequency); conversely, the dynamics is slow for structured point-clouds (low frequency). Deriving from this, we consider that an efficient generic compression method must be able to capture the dynamic characteristics of data. A typical variational method with analysis at different scales is by using wavelet to compress the point-cloud in a tree structure [17]. However, this method is prone to low efficient when the point-cloud is sparse, due to the similar reason as structure-dependent methods.

In this paper, inspired by these observations, we propose to use three-dimensional DCT (Discrete Cosine Transform) to unveil the undiscovered information in the raw data. DCT captures signal properties in frequency domain where the low frequency components provides low dynamic information and high frequency components gives more detailed information. In this case, the coarse shapes of the point-cloud are retained by reconstruction using inversed DCT (IDCT) with coefficients at low frequency. After that, details can be refined by using more coefficients at a higher frequency. An existing solution of compressing 3D data using Discrete Cosine Transform was

978-1-5386-3154-6/17/$31.00 ©2017 IEEE

confined to two dimensions and had the Cartesian coordinates transformed to spherical coordinates [18]. In the spherical coordinates, sampling was performed on two angles and the index associated with angles specified the location of the corresponding depth value. The 2D DCT took the depth value to operate the transform. However, this method only works for depth images, where all the points can be projected to a spherical surface without losing information. This is usually not the case for a complete representation of an object or 3D maps constructed by point-clouds.

Our solution aims to reduce the attenuation of the depth value by introducing one more dimension comparing to [18], by which more precise structures in complicated environments can be preserved. This is implemented by assigning a constant value (potentially can also use the intensity of laser reflection [19]), to the present points in Cartesian coordinates. Sampling and further preprocessing steps are taken in advance to convert the Cartesian coordinates into positive integers before 3D DCT takes place. 3D DCT results in a large amount of coefficients left in the 3D matrix and by quantizing the them properly, a significant amount of data will be compressed and only very few amount of data will be left to transmit.

B. Contributions

In this paper, we implement a robust generic method of data compression subjecting to the information significance of the data, based on 3D-DCT. This approach will be first examined in point-cloud data collected from the scans of individual objects [1] and then it will be extended into the compression of environment scans. The environment data are provided by our previous work [19], where we recorded the data from several scenarios including structured and unstructured instances, e.g. an apartment, a hallway, a mountain plain, stairs, a gazebo in the park and woods. These data sets are representatives of a great variety of environments and they are good sources to be put on trial for the experiments.

C. Organization

This paper consists of four sections. Section I describes the background and our aim in research, continued by Section II explaining the algorithm in data compression. Sections IV presents the result in the experiment and discusses the reasons behind. At the end, Section IV will draw conclusions.

II. Proposed Algorithm for Point-cloud Compression

We assume the raw point-cloud data were provided as an $M \times 3$ matrix, including a large amount of real 3D row vectors, each of which represents a single point of the cloud. In order to reduce the computational complexity and help the designation of proper parameters for 3D-DCT, we first perform several preprocessing steps to convert the raw data into appropriate positive integers. After the refined data were obtained, a 3D matrix was generated to house all the grid points whose Cartesian coordinates were specified in the refined 2D matrix. The 3D matrix was then processed by 3D DCT and it holds all the DCT coefficients acquired. Then a quantization procedure using a constant quantization table is required to reduce the valid DCT coefficients. This process is inspired by the standard JPEG compression process [20]. It reduces a significant amount of storage required, as many coefficients regarding the components of high frequency would be quantized to 0. In this experiment, we further boost up the compression ratio by iteratively reducing the amount of quantized coefficients to transfer. Dequantization and 3D IDCT were then followed. Finally, the compressed data was reconstructed and errors were measured. An illustration of the mentioned algorithm is described in the flow chart below.

[1] http://www.csse.uwa.edu.au/∼ajmal/recognition.html

A. Sampling and Preprocessing of Data Points

Given the original data set, sampling was first done to reduce the amount of points by factor of 10. The mean value of the coordinates were taken within the sampling frame. The sampled coordinates were rounded to closest integers and shifted by the range of the data set to prepare for the positive integer-based 3D DCT operation. The unique data points were selected to be contained in a three dimensional matrix, where the range of the Cartesian X, Y, Z coordinates represented each dimension of the matrix respectively. The intensity value 255 was assigned to indicate the presence of a data point whereas 0 denoted the absence.

B. 3D-DCT Operation and Quantization

The three dimensional matrix constructed above was passed to perform 3D-DCT. The 3D DCT formula is given by,

$$\begin{aligned} F(u,v,w) = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}\sum_{k=0}^{n-1} f(i,j,k) a\left(u\right) a\left(v\right) a\left(w\right) \\ \cos\frac{\left(2i+1\right)u\pi}{2n} \cos\frac{\left(2j+1\right)v\pi}{2n} \cos\frac{\left(2k+1\right)w\pi}{2n} \end{aligned} \tag{1}$$

where

$$a(u) = \begin{cases} \sqrt{\dfrac{1}{n}} \; for \; u = 0 \\ \sqrt{\dfrac{2}{n}} \; for \; u = 1, 2, ..n-1 \end{cases} \tag{2}$$

Quantization of coefficients was done at a uniform level of 255. Several iterations to select reduced amount of the quantized coefficients will be run to inspect how error rate varies with compression ratio. Compression ratio is calculated by:

$$\text{Compression Ratio (CR)} = \frac{\#\text{points after preprocessing}}{\#\text{coefficients for storage or transfer}} \tag{3}$$

C. Dequantization and 3D-IDCT

To reconstruct the original data, dequantization was first performed and the 3D-IDCT operation next. A table marking the location of data points was stored before 3D-DCT operation. It could be used to select the recovered data points after IDCT operation.

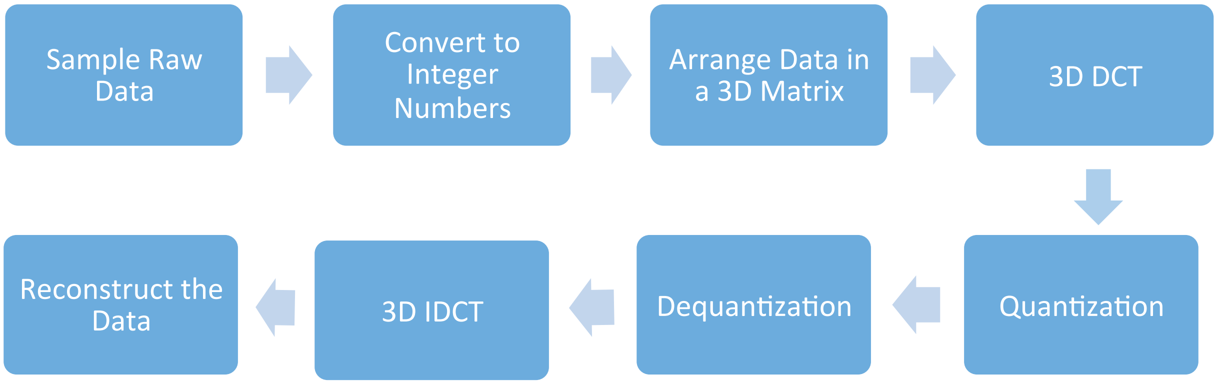


Fig. 1. The work flow of the proposed algorithm using 3D-DCT

$$f'(i,j,k) = \sum_{u=0}^{n-1}\sum_{v=0}^{n-1}\sum_{w=0}^{n-1} F(u,v,w)\cos\frac{(2i+1)u\pi}{2n}\cos\frac{(2j+1)v\pi}{2n}\cos\frac{(2k+1)w\pi}{2n} \tag{4}$$

Method	Octree [8], [1]	Proposed
Compression	Octree construction	3D-DCT
Decompression	Using nodes of the tree	3D-IDCT
Complexity	$O\left(Nlog(N)\right)$	$O\left(N^2\right)$
Compression Ratio	Low	High
Sensitive to data orientation	Yes	No

TABLE I
COMPARISON BETWEEN OCTREE-BASED AND THE PROPOSED METHODS

Object	Chicken	Trex	Kangaroo	Chef
Compression Ratio	23.42	37.86	51.97	48.10
Error Rate (%)	2.26	2.45	4.32	2.08

TABLE II
COMPARISON OF COMPRESSION RATIO AND ERROR RATE OF POINT-CLOUD COMPRESSION ON OBJECT

All the selected coefficients in the 3D matrix was used to perform IDCT operation. The IDCT formula is given by equation (4) on the next page.

D. Reconstruction of Data and Evaluation of Errors

Hamming distance was introduced to evaluate the reconstructed errors. If a sample cannot be reconstructed successfully, the hamming distance will increment by 1. It calculated the total variation from the data points obtained before 3D DCT operation to after IDCT operation. Error rate is calculated by:

$$\text{Error Rate (ER)} = \frac{\text{Summed Hamming distance}}{\#\text{Points after reconstruction}} \tag{5}$$

E. Summary

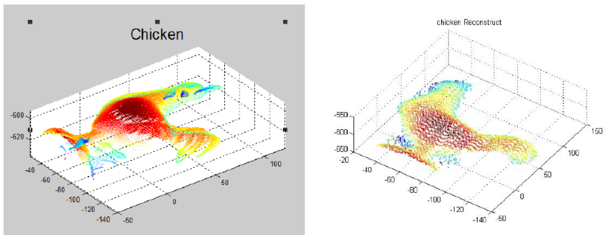
Given the proposed approach, we can have an intuitive comparison to the octree-based methods [8], [1] as shown in table I.

III. EXPERIMENT RESULTS

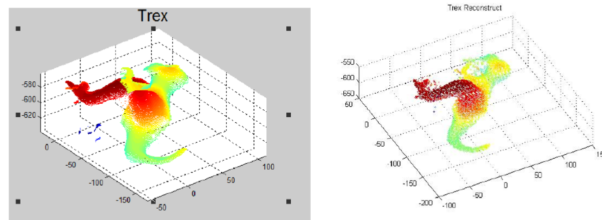
While preserving all of the 3D DCT coefficients, we observed a low error rate in reconstruction with the compression ratio resulted. The reconstruction of the objects can be easily identified by human eye perception and thus validated the 3D DCT-based encoding in data point-cloud compression of individual objects. The table below provides the statistics of object compression ratio and error rate.

When the amount of 3D DCT coefficients to store was further reduced, by which was implemented by iteratively selecting fewer coefficients, we observed that a relatively stable error rate was maintained while the compression ratio improved. As DCT decorrelates signal best when the signal information concentrates at the low frequency, an object outlined by a smooth contour and covered by a continuous surface fulfills this property. The error rate associated with the reconstructed object demonstrated the application of 3D DCT in object point-cloud compression achieve a very high compression ratio by compromising little errors. Comparisons between the objects constructed from raw data and compressed data are displayed as follows. A table presenting varying compression ratio vs error rate of four different objects is also provided.

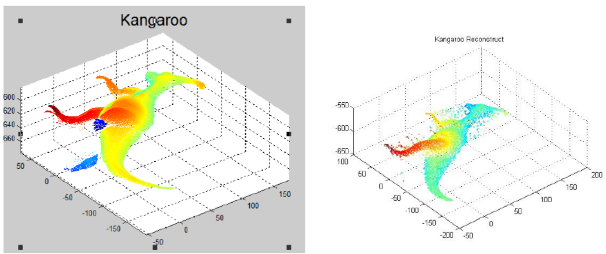
With these results, the experiment then progressed to explore the possibilities of the applying 3D DCT in environment settings to achieve compression. Multiple objects existed in an environment and some of them were dynamically moving while some were not. As each environment was structured in a different way, not all environments contain a "compact energy" property and thus 3D DCT operation may not be effective to all environments. Static environments such as apartment, stairs and mountain plain tend to have low error rates whereas



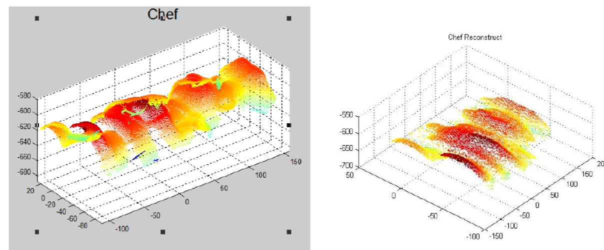
(a) Chicken Data-set



(b) Trex Data-set



(c) Kangaroo Data-set



(d) Chef Data-set

Fig. 2. Compression on point-cloud from objects

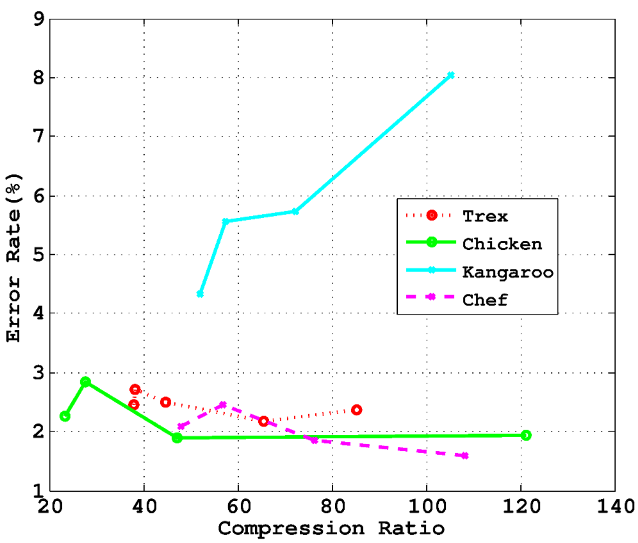


Fig. 3. Error Rate of the four objects

the gazebo and woods environment with dynamic objects, e.g. people travelling around, seem to bring in more difficulties in accurate reconstructions. The gazebo and wood scenarios were representations of semi-structured and unstructured environments respectively. The complicated wood structures in these two settings could be the source triggering the slightly low precision in recovery. The apartment data set recorded was at a low complexity level, by which the apartment was bounded by walls, ceilings and floors. These indoor structures did not change rapidly and made the apartment scenario appropriate for 3D DCT operation, which can be verified from the low error rate in reconstructions. Unlike the highly constrained setting of apartment, the mountain plain was considered low constrained as it was in a large open space where the major element was dry vegetation. The continuity and the simple setting of the plain created low complexity of this environment, thus the error rate of mountain plain data was also low.

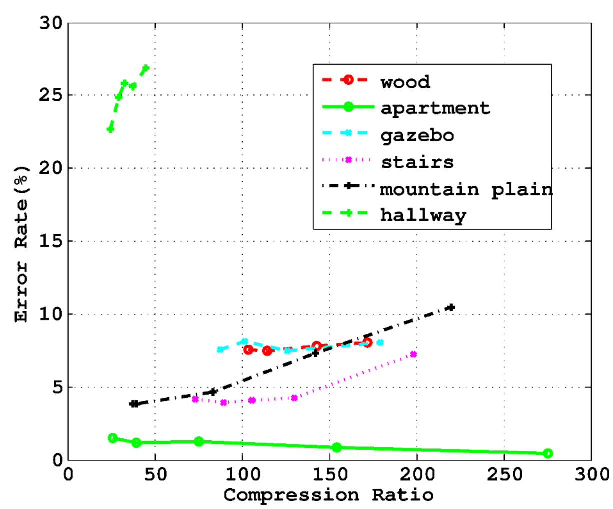


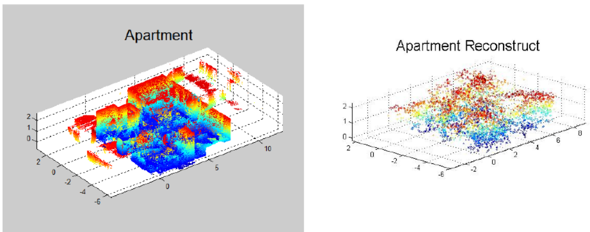
Fig. 4. Error Rate of 3D maps

The response of hallway scenario to 3D DCT operation was less satisfactory with an error rate as high as 22.71%. One of the reason can be the structure of hallway. The hallway was constituted with many alternating pillars and arches and thus contain more elements of high frequency. The other factor contributed to the high error rate was the design of the quantizer. In both the experiment of object and environment data testing, the quantizer used was a uniform one, quantizing at a level of 255. The error rate of hallway data compression was over 40% under this level and it diminished to around

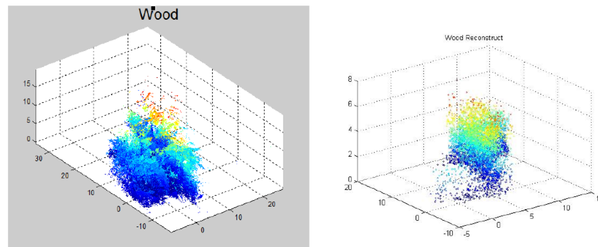
Environment	Apartment	Wood	Hallway	Stairs	Gazebo	Mountain Plain
Compression Ratio	25.87	104.02	24.42	73.45	87.41	37.23
Error Rate (%)	1.53	7.59	22.71	4.15	7.55	3.86

TABLE III

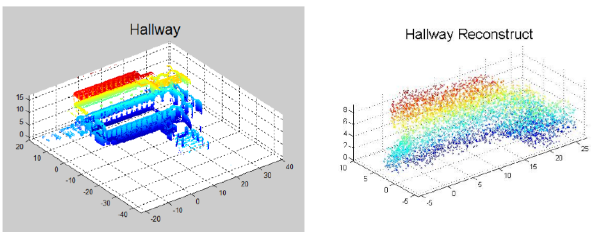
COMPARISON OF COMPRESSION RATIO AND ERROR RATE OF POINT-CLOUD COMPRESSION ON 3D MAPS



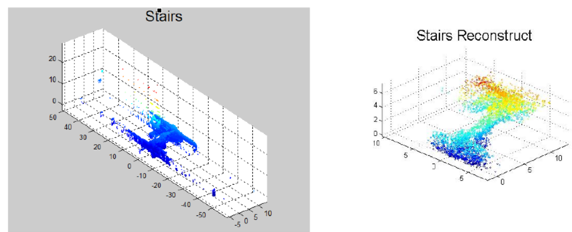
(a) Apartment Data-set



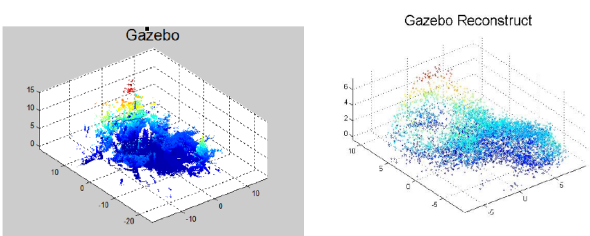
(b) Wood Data-set



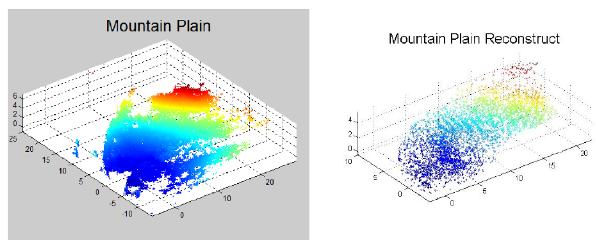
(c) Hallway Data-set



(d) Stair Data-set



(e) Gazebo Data-set



(f) Mountain Plain Data-set

Fig. 5. Compression on 3D point-cloud maps

20% when the quantization level was reduced from 255 to 40. Therefore, the error rate in this experiment was subject to the design of quantizer. If a more sophisticated quantizer that gives the 3D DCT coefficients different quantization levels considering the frequency of components, error rate may be reduced. As can be seen from the reconstructed images, a major outline of every environment can be preserved well but regions with abrupt and vast changes suffer from much information loss. The corners in the apartment, the steps of the staircase and each individual arch and pillar in the hallway failed to recover when this transform applied.

IV. ACKNOWLEDGES

This project is supported by Shenzhen Science, Technology and Innovation Comission (SZSTI) JCYJ20160428154842603 and JCYJ20160401100022706; partially supported by the Research Grant Council of Hong Kong SAR Government, China, under project No. 16206014 and No. 16212815; National Natural Science Foundation of China No. 6140021318, awarded to Prof. Ming Liu; National Natural Science Foundation of China No. 61603376, awarded to Dr. Lujia Wang.

V. CONCLUSION

In this research, we have preprocessed the acquired 3D data and make them applicable to 3D DCT operation. We have applied 3D DCT on both the object data as well as the environment data and devised a quantizer to compress the data to transmit. The result of the experiment demonstrated very effective compression outcomes on objects and several scenarios of environments. The performance of the compression on data can be bettered by more advanced quantizer designs. The conversion of decimal coordinates to positive integers is another important factor affecting the reconstruction effect of the original point-cloud. In the future, we will look into other representations of data before taking them to 3D DCT and improve the design of quantizers to further enhance the results in the experiment.

References

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

[2] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann, "6d slam-3d mapping outdoor environments," *Journal of Field Robotics*, vol. 24, no. 8-9, pp. 699–722, 2007.

[3] F. Colas, S. Mahesh, F. Pomerleau, M. Liu, and R. Siegwart, "3d path planning and execution for search and rescue ground robots," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, 2013, pp. 722–727.

[4] M. Liu, L. Wang, and R. Siegwart, "DP-Fusion: A generic framework for online multi sensor recognition," in *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2012.

[5] M. Gianni, P. Papadakis, F. Pirri, M. Liu, F. Pomerleau, F. Colas *et al.*, "A unified framework for planning and execution-monitoring of mobile robots," in *Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence*, 2011.

[6] L. Wang, M. Liu, M. Q.-H. Meng, and R. Siegwart, "Towards Real-Time Multi-Sensor Information Retrieval in Cloud Robotic System," in *Proceedings of the IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2012.

[7] L. Wang, M. Liu, and M. Meng, "An auction-based resource allocation strategy for joint-surveillance using networked multi-robot systems," in *IEEE International Conference on Information and Automation (ICIA)*, 2013, pp. 424–429.

[8] K. M. Wurm, A. Hornung, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: A probabilistic, flexible, and compact 3d map representation for robotic systems," in *Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation*, vol. 2, 2010.

[9] G. Vosselman, B. G. Gorte, G. Sithole, and T. Rabbani, "Recognising structure in laser scanner point clouds," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 46, no. 8, pp. 33–38, 2004.

[10] M. Y. Yang and W. Förstner, "Plane detection in point cloud data," in *Proceedings of the 2nd int conf on machine control guidance, Bonn*, vol. 1, 2010, pp. 95–104.

[11] J. Jia and C.-K. Tang, "Inference of segmented color and texture description by tensor voting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 6, pp. 771–786, 2004.

[12] M. Liu, F. Pomerleau, F. Colas, and R. Siegwart, "Normal estimation for pointcloud using gpu based sparse tensor voting," in *Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference on*. IEEE, 2012, pp. 91–96.

[13] M. Liu and R. Siegwart, "Navigation on point-cloud - a riemannian metric approach," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014.

[14] H. Houshiar and A. Nüchter, "3d point cloud compression using conventional image compression for efficient data transmission," in *2015 XXV International Conference on Information, Communication and Automation Technologies (ICAT)*, Oct 2015, pp. 1–8.

[15] T. Golla and R. Klein, "Real-time point cloud compression," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, to appear, Conference 28 September - 3 October 2015.

[16] X. Zhang, W. Wan, and X. An, "Clustering and dct based color point cloud compression," *Journal of Signal Processing Systems*, vol. 86, no. 1, pp. 41–49, 2017. [Online]. Available: http://dx.doi.org/10.1007/s11265-015-1095-0

[17] C.-C. Yang, "The compression of 3d point cloud data using wavelet transform," Master's thesis, I-shou University, 2007.

[18] C.-F. Chen and C.-Y. Chen, "Tech report: Compression of 3d point data using discrete cosine transform."

[19] F. Pomerleau, A. Breitenmoser, M. Liu, F. Colas, and R. Siegwart, "Noise characterization of depth sensors for surface inspections," in *Applied Robotics for the Power Industry (CARPI), 2012 2nd International Conference on*. IEEE, 2012, pp. 16–21.

[20] G. Beretta, V. Bhaskaran, K. Konstantinides, and B. K. Natarajan, "Method for selecting jpeg quantization tables for low bandwidth applications," March 1999, uS Patent 5,883,979.