Fig. 9 to 12 show images of the two environments examined in the experiment.

Fig. 9 shows the results obtained when the road surface has many complex edge components, including the shadows of buildings and trees. While camera sensors encounter multiple problems, Fig. 10 shows that the proposed method can estimate the road surface without difficulty.

Fig. 12 shows that the proposed method could estimate the road surface even when there were various markings on the road surface, such as speed limits, cross walks, and lane demarcations.

IV. CONCLUSION

Our proposed method involves two steps. First, it uses a single camera and deep learning to produce a depth map, with PyD-Net code and a deep neural network with a pyramid form employed for rapid calculation. Since PyD-Net has significantly fewer parameters than a convolutional neural network, high-speed calculation is possible. Then, using the obtained depth map, a normal vector is obtained at each point in the image; these are then quantized and clustered for real-time operation, which is much faster. Our method requires very few calculations regardless of the environmental conditions and changes in the image. The method enables real-time estimation with no drop in performance. We plan to conduct additional experiments to identify other applications using a single camera, such as augmented reality applications.

ACKNOWLEDGMENT

REFERENCES

[1] B. Jähne and H. Horst, *Computer vision and applications*, pp. 111-151, 2000.

[2] J. Brownlee, *Deep learning for computer vision: image classification, object detection, and face recognition in python*, 2019.

[3] M. Poggi, F. Aleotti, F. Tosi, and S. Mattoccia, Towards real-time unsupervised monocular depth estimation on cpu, *In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5848-5854, Oct. 2018.

[4] K. Y. Lin and H. M. Hang, Depth map enhancement on rgb-d video captured by kinect v2, *In 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1530-1535, Nov. 2018.

[5] Y. W. Kuan, N. O. Ee, and L. S. Wei, Comparative study of intel R200, Kinect v2, and primesense RGB-D sensors performance outdoors, *IEEE Sensors Journal, 19(19)*, pp. 8741-8750, 2019.

[6] D. Pohl, S. Dorodnicov, and M. Achtelik, Depth map improvements for stereo-based depth cameras on drones, *In 2019 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 341-348, Sep. 2019.

[7] G. Yang, X. Song, C. Huang, Z. Deng, J. Shi, and B. Zhou, Drivingstereo: A large-scale dataset for stereo matching in autonomous driving scenarios, *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 899-908, 2019.

[8] O. Ozyesil, V. Voroninski, R. Basri, and A. Singer, A survey of structure from motion, *arXiv preprint arXiv:1701.08493*, 2017.

[9] M. R. U. Saputra, A. Markham, and N. Trigoni, Visual SLAM and structure from motion in dynamic environments: A survey, *ACM Computing Surveys (CSUR), 51(2)*, pp. 1-36, 2018.

[10] A. Shalaby, M. Elmogy, and A. A. El-Fetouh, Algorithms and applications of structure from motion (SFM): *A survey, Algorithms, 6(06)*, 2017.

[11] H. Zhan, R. Garg, C. S. Weerasekera, K. Li, H. Agarwal, and I. Reid, Unsupervised learning of monocular depth estimation and visual odometry with deep feature reconstruction, *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 340-349, 2018.

[12] J. W. Bian, Z. Li, N. Wang, H. Zhan, C. Shen, M. M. Cheng, and I. Reid, Unsupervised scale-consistent depth and ego-motion learning from monocular video, *arXiv preprint arXiv:1908.10553*, 2019.

[13] C. Godard, O. M. Aodha, M. Firman, and G. J. Brostow, Digging into self-supervised monocular depth estimation, *In Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3828-3838, 2019.

[14] A. Ardakani, C. Condo, M. Ahmadi, and W. J. Gross, An architecture to accelerate convolution in deep neural networks, *IEEE Transactions on Circuits and Systems I: Regular Papers, 65(4)*, 1349-1362, 2017.

[15] S. Karen and A. Zisserman, Deep Convolutional networks for large-scale image recognition, *arXiv preprint arXiv:1409.1556*, 2014.

[16] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, Vision meets robotics: The kitti dataset, *The International Journal of Robotics Research, 32(11)*, pp. 1231-1237, 2013.

[17] A. Geiger, P. Lenz, and R. Urtasun, Are we ready for autonomous driving? the kitti vision benchmark suite, *In 2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3354-3361, June 2012.

[18] M. Poggi, PyDnet, 2018. [online]. Available: https://github.com/mattpoggi/pydnet

[19] D. Holz, S. Holzer, R. B. Rusu, and S. Behnke, Real-time plane segmentation using RGB-D cameras, *In Robot Soccer World Cup*, pp. 306-317, July 2011.

[20] J. Oh, K. S. Kim, M. Park, and S. Kim, A comparative study on camera-radar calibration methods, *In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 1057-1062, Nov. 2018.

[21] Z. Zhang, R. Zhao, E. Liu, K. Yan, and Y. Ma, A single-image linear calibration method for camera, *Measurement, 130*, pp. 298-305, 2018.

[22] O. Bouafif, B. Khomutenko, and M. Daoudi, Monocular 3D Head Reconstruction via Prediction and Integration of Normal Vector Field, *In 15th International Conference on Computer Vision, Theory and Applications*, Feb. 2020.