Project Proposal
Time-of-Flight Coding Function Optimization using Neural Networks

Nicholas Gaudio
Stanford University
Department of Electrical Engineering
nsgaudio@stanford.edu

Jonas Messner
Stanford University
Department of Electrical Engineering
messnerj@stanford.edu

3D imaging has become an important area of research for various applications including autonomous driving, robotics, and gaming. A common 3D imaging technique is Continuous-Wave Time-of-Flight (CW-ToF) imaging. In CW-ToF imaging a modulated light wave is emitted and the object reflected returning wave is measured by a demodulating sensor. The distance of the object can then be determined from the cross-correlation between the emitted and received waves.

The modulation (source) and demodulation (sensor) functions are usually sinusoid or square waves. However, in [3] it is shown that modulation and demodulation functions inherently influence the depth resolution of a CW-ToF system. Further, it is shown that sinusoid and square waves are not the optimal coding functions for Time-of-Flight imaging. Hamiltonian coding functions are introduced, which improve the depth resolution of the system, but are *not provably optimal*. Our goal is to train the coding functions using a neural network and thus find coding functions that outperform Hamiltonian coding functions.

Specialized coding functions were studied by M. Gupta *et al*. in [3]. In [4] it is investigated how Hamiltonian-like coding schemes can be practically implemented in actual imaging hardware. Further works with coding functions other than sinusoid or square waves are [1], [2], [5], and [7].

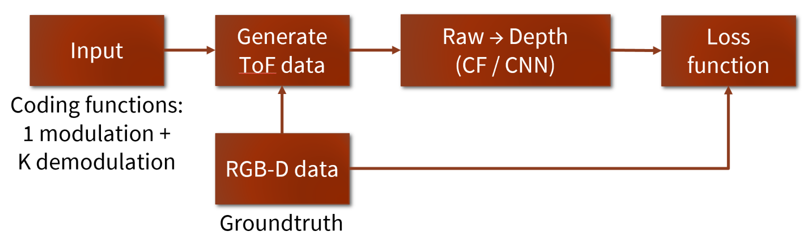

Figure 1. Block diagram of the intended model.

To the best of our knowledge learning CW-ToF coding functions has not been done before. Therefore, there is no existing implementation. Figure 1 shows a block diagram of our model. The inputs are one source modulation function and K demodulation functions, which will all be trained. Raw ToF data is synthetically generated using a simulator that we received from the work of [3]. We will start with a closed-form (CF) solution for the raw-to-depth conversion and compare the obtained and groundtruth depths. In this case, at least three demodulation functions are necessary in order to calculate the depth. We will backpropagate through the model (even though there is no neural network yet), update the coding functions, and investigate the results. In a second step we will replace the raw-to-depth closed-form solution with a convolutional neural network (CNN). In this step we will investigate if CNNs are able to generate good depth maps with less than three demodulation functions and what these functions would look like.

We plan to use the NYU V2 dataset [6], which contains RGB-D images. We also want to show how different scenes might necessitate different coding functions. For instance, CW-ToF imagers used to perform face recognition could have specialized coding functions that work particularly well with face depth data. Therefore, we might add a 3D face dataset to our work later in the project.

We will evaluate our results using root mean square error (RMSE) between the ToF depth map and the 3D groundtruth as in [3]. Our results will be compared to other coding functions such as sinusoid, square and, Hamiltonian coding functions.

References

[1] R. Ferriere, J. Cussey, and J. Dudley. Time-of-flight range detection using low-frequency intensity modulation of a cw laser diode: Application to fiber length measurement. *Optical Engineering - OPT ENG*, 47, 09 2008.

[2] R. Grootjans, W. van der Tempel, D. Van, C. De Tandt, and M. Kuijk. Improved modulation techniques for time-of-flight ranging cameras using pseudo random binary sequences. *Proc. IEEE LEOS Benelux Chapter*, 2006.

[3] M. Gupta, A. Velten, S. K. Nayar, and E. Breitbach. What are optimal coding functions for time-of-flight imaging? *ACM Trans. Graph.*, 37(2):13:1–13:18, Feb. 2018.

[4] F. Gutierrez-Barragan, S. A. Reza, A. Velten, and M. Gupta. Practical coding function design for time-of-flight imaging. *To appear in CVPR 2019*, 2019.

[5] A. Kolb, E. Barth, R. Koch, and R. Larsen. Time-of-flight cameras in computer graphics. *Computer Graphics Forum*, 29(1):141–159, 2010.

[6] P. K. Nathan Silberman, Derek Hoiem and R. Fergus. Indoor segmentation and support inference from rgbd images. In *ECCV*, 2012.

[7] A. Payne, A. A Dorrington, and M. Cree. Illumination waveform optimization for time-of-flight range imaging cameras. *Proceedings of SPIE - The International Society for Optical Engineering*, 8085, 06 2011.