Deep Learning based Compensation of Motion and Multi Path Interference in Time-of-Flight Cameras

Veeresh Prakash Dammur
prakas01@ads.uni-passau.de

First supervisor: Prof. Dr. Ralf Krestel
Second supervisor: Prof. Dr. Harald Kosch
External supervisor: Dipl.-Inf. Christoph Wiedemann

June 8, 2020

UNIVERSITÄT PASSAU

Outline

UNIVERSITÄT PASSAU

Introduction to ToF Technology

- What are ToF (Time of Flight) cameras?
- Working principle of ToF camera

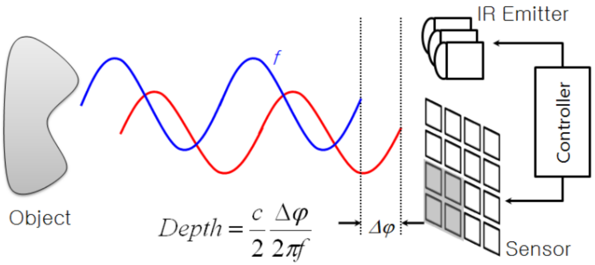

Figure: Working principle of ToF camera [Hansard et al., 2012]

ToF camera artifacts

Classification of ToF camera artifacts

ToF camera artifacts

Classification of ToF camera artifacts

Systematic artifacts

- Shot noise
- Depth wiggling
- Integration error
- Temperature error

ToF camera artifacts

Classification of ToF camera artifacts

Systematic artifacts

- Shot noise
- Depth wiggling
- Integration error
- Temperature error

Non-systematic artifacts

- Multipath interference (MPI)
- Motion artifacts
- Flying pixels

UNIVERSITÄT PASSAU

ToF camera artifacts

What is Multipath interference in ToF cameras?

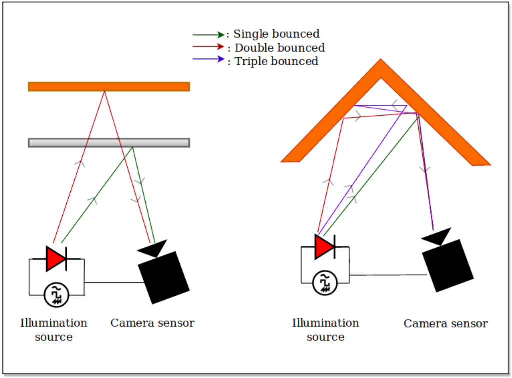

Figure: An illustration of Multipath interference (MPI) artifacts in ToF camera

ToF camera artifacts

What are motion artifacts in ToF cameras?

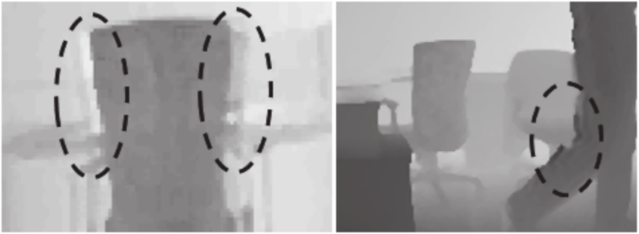

Figure: An illustration of ToF depth motion artifact [Lee et al., 2012]. Blurred regions of the swivel chair is shown in the dashed ellipses.

Research question

Performance evaluation of the approach proposed by [Guo et al., 2018] applied to PMD (Photonic mixer device) camera.

Research question

Performance evaluation of the approach proposed by [Guo et al., 2018] applied to PMD (Photonic mixer device) camera.

- PMD Dataset

Research question

Performance evaluation of the approach proposed by [Guo et al., 2018] applied to PMD (Photonic mixer device) camera.

- PMD Dataset
  - Sensor model
  - Noise model

Research question

Performance evaluation of the approach proposed by [Guo et al., 2018] applied to PMD (Photonic mixer device) camera.

- PMD Dataset
  - Sensor model
  - Noise model
- Number of phases to determine the depth of a scene.

Approach

PMD dataset generation

Dataset Generation:

PMD dataset generation

Dataset Generation:

- Simulated measurements using FLAT [Guo et al., 2018] dataset

PMD dataset generation

Dataset Generation:

- Simulated measurements using FLAT [Guo et al., 2018] dataset
  - What is FLAT?

PMD dataset generation

Dataset Generation:

- Simulated measurements using FLAT [Guo et al., 2018] dataset
  - What is FLAT?
    - Flexible, Large, Augmentable, ToF dataset

PMD dataset generation

Dataset Generation:

- Simulated measurements using FLAT [Guo et al., 2018] dataset
  - What is FLAT?
    - Flexible, Large, Augmentable, ToF dataset
  - Why FLAT dataset?

PMD dataset generation

Dataset Generation:

- Simulated measurements using FLAT [Guo et al., 2018] dataset
  - What is FLAT?
    - Flexible, Large, Augmentable, ToF dataset
  - Why FLAT dataset?
    - Scenes are rendered by Transient rendering

PMD dataset generation

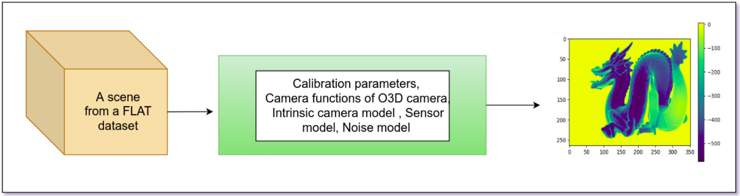

Figure: A block diagram that depicts the sequence of operations in generating the PMD measurements.

Sensor model

- Why?

Sensor model

- Why?
  - To generate sensor model for two phase operational mode.

Sensor model

- Why?
  - To generate sensor model for two phase operational mode.
- How?

Sensor model

- Why?
  - To generate sensor model for two phase operational mode.
- How?
  - By solving a least square minimization problem

Sensor model

- Why?
  - To generate sensor model for two phase operational mode.
- How?
  - By solving a least square minimization problem

$$\sum_i e^2 =$$

$$\left(f_{cal}(A^P, B^P) - \frac{1}{2}[(A^P - B^P) - (A^{P+180^\circ} - B^{P+180^\circ})] \right)^2$$

and $f_{cal}(A^P, B^P)$ is defined as,

$$f_{cal}(A^P, B^P) = m_1.A^P - n_1.B^P + t_1$$

where A, B are sensor readout values. $P = [0^\circ, 90^\circ]$.

UNIVERSITÄT PASSAU

Noise model

- Estimate the noise associated with A-B measurements.
- Performed acquisition of raw measurements.
- Noise model - Normal distribution with standard deviation depending linearly on amount of light integrated.

Noise model

- Estimate the noise associated with A-B measurements.
- Performed acquisition of raw measurements.
- Noise model - Normal distribution with standard deviation depending linearly on amount of light integrated.

The standard deviation of A, B can be represented as,

$$\sigma_A = n_1 \cdot \mu_A + t_1$$
$$\sigma_B = n_2 \cdot \mu_B + t_2$$

where μ_A, μ_B indicate mean of A, B respectively.

UNIVERSITÄT PASSAU

Noise model

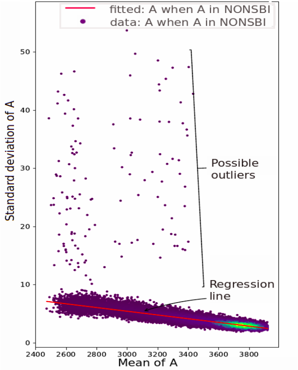

Figure: Plot of standard deviation of A vs. mean of A.

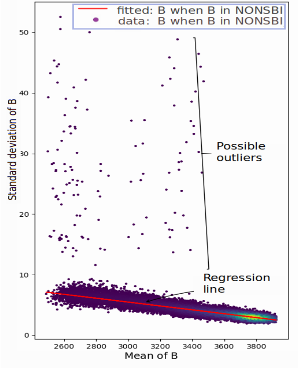

Figure: Plot of standard deviation of B vs. mean of B.

UNIVERSITÄT PASSAU

PMD dataset generation

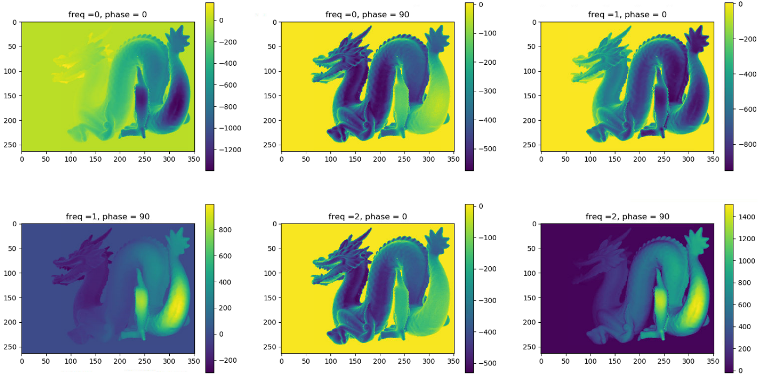

Figure: Sample full PMD raw measurements of a static scene. Frequency and phase values are displayed on each subplot. These six raw measurements are used to measure the depth of a scene.

Deep learning models

Deep learning models

- Experiments
  - Multi reflection experiment
  - Motion experiment
- Both experiments are structured as follow
  - System architecture
  - Setup
    - Dataset statistics
    - Loss function
    - Baseline
  - Loss curves
  - Results

Multi reflection experiment

System architecture

- Multi reflection model based on Kernel predictive network (KPN) [Bako et al., 2017]

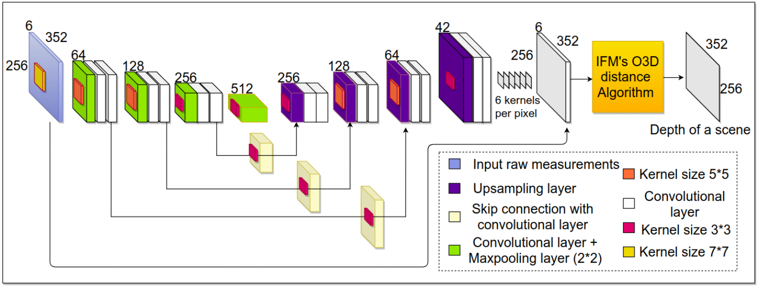

Figure: Architecture of the Multi reflection model (MRM).

Setup

- Dataset statistics

Table: Details of the dataset used in the multi reflection experiment

Scene type	Train	Test	Total
Foreground scenes	648	0	648
Background scenes	202	117	319
Total	850	117	967

Setup

- Dataset statistics

Table: Details of the dataset used in the multi reflection experiment

Scene type	Train	Test	Total
Foreground scenes	648	0	648
Background scenes	202	117	319
Total	850	117	967

- Loss function

$$loss_{scene} = \sqrt{\frac{\sum(y_{out} - y')^2}{N_{validPixels}}}$$

UNIVERSITÄT PASSAU

Setup

- Dataset statistics

Table: Details of the dataset used in the multi reflection experiment

Scene type	Train	Test	Total
Foreground scenes	648	0	648
Background scenes	202	117	319
Total	850	117	967

- Loss function

$$loss_{scene} = \sqrt{\frac{\sum(y_{out} - y')^2}{N_{validPixels}}}$$

- Baseline
  Ifm's distance algorithm

UNIVERSITÄT PASSAU

Multi reflection experiment

Loss curves:

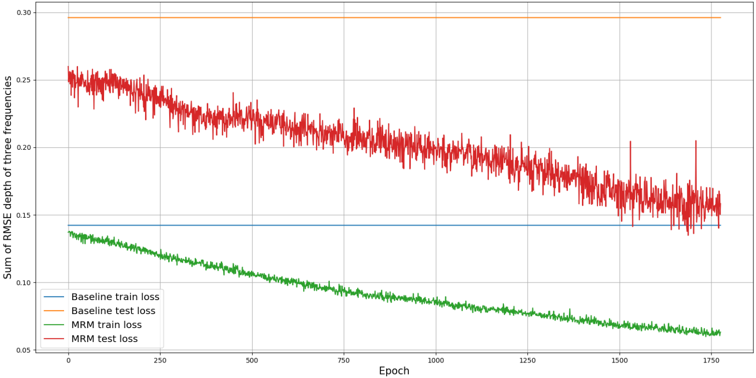

Figure: Loss curves of the the multi reflection experiment. The unit of the vertical axis is meters

Multi reflection experiment

Results:

Table: Results of the multi reflection experiment. All statistics are calculated on test data set.

	Median depth error	Percentile (70^{th}) depth error
PMD pipeline (baseline)	10.72 cm	13.92 cm
MRM + PMD pipeline	4.68 cm	6.57 cm

Multi reflection experiment

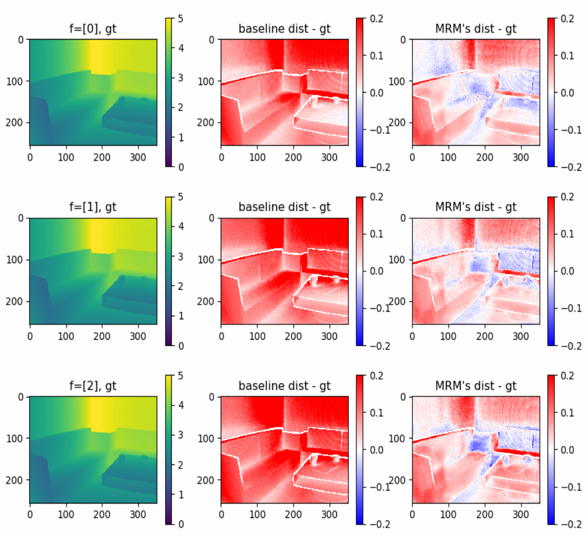

Figure: A sample scene result of the MRM experiment. The unit of colorbar is meters.

Motion experiment

System architectures

- Motion model base on Flownet architecture [Fischer et al., 2015]

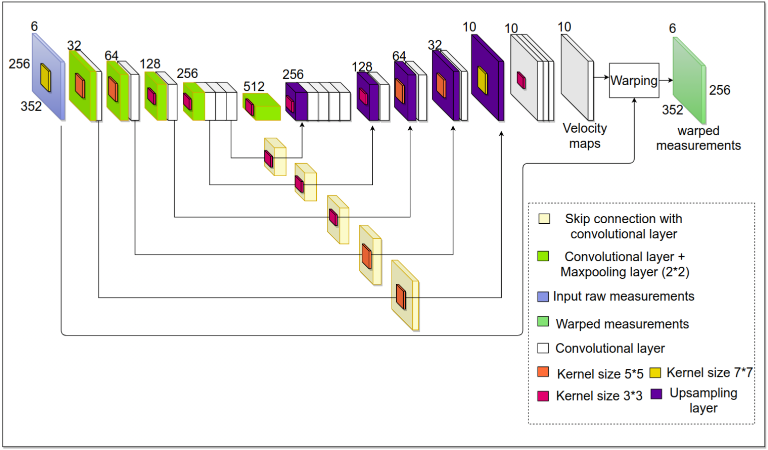

Figure: Architecture of the motion model (MoM)

Setup

- Dataset statistics

Table: Details of the dataset used in the motion experiment

Dataset	Foreground scenes	Background scenes	Augmented scenes
Train set	572	202	6400
Test set	76	117	1600
Total	648	319	8000

UNIVERSITÄT PASSAU

Setup

- Dataset statistics

Table: Details of the dataset used in the motion experiment

Dataset	Foreground scenes	Background scenes	Augmented scenes
Train set	572	202	6400
Test set	76	117	1600
Total	648	319	8000

- Loss function

$$loss_{scene} = \sqrt{\frac{\sum (v_{yx} - v'_{yx})^2}{size(v_{yx})}}$$

Setup

- Dataset statistics

Table: Details of the dataset used in the motion experiment

Dataset	Foreground scenes	Background scenes	Augmented scenes
Train set	572	202	6400
Test set	76	117	1600
Total	648	319	8000

- Loss function

$$loss_{scene} = \sqrt{\frac{\sum(v_{yx} - v'_{yx})^2}{size(v_{yx})}}$$

- Baseline : RMSE of gt velocity maps.

Loss curves

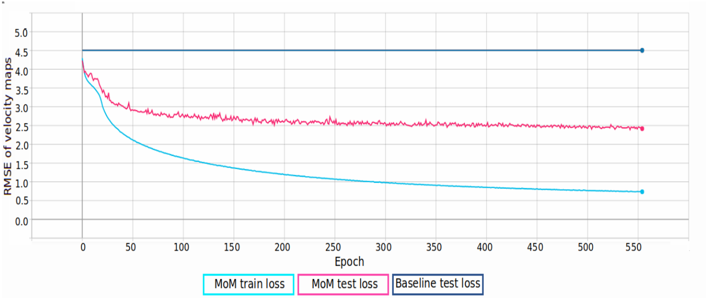

Figure: Train and test losses of motion experiment.

Motion experiment

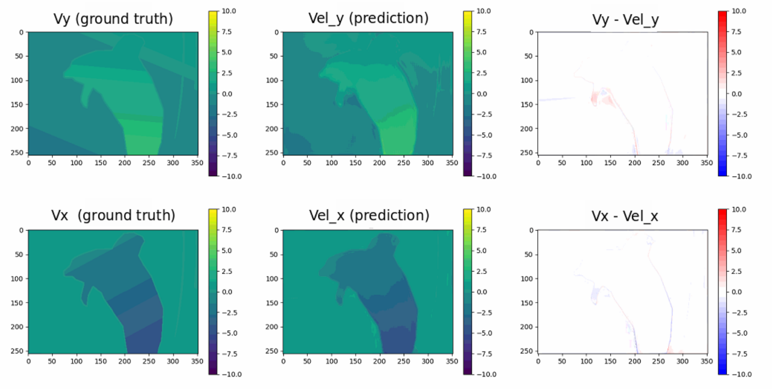

Figure: Velocity vectors difference between model's predicted and gt (from train set)

Motion experiment

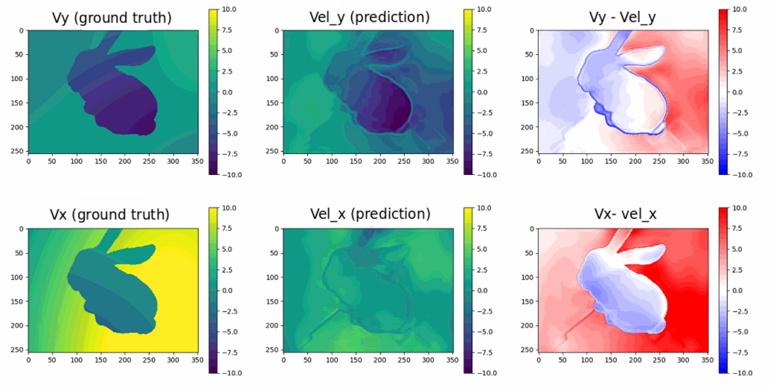

Figure: Velocity vectors difference between model's predicted and gt (from test set)

Motion experiment

Possible reasons for high velocity vector errors on the testset:

- Occlusion pixels.
- Trainset was not sufficient for the model to generalize on test set.

Conclusion and future work

Conclusion:

- Used FLAT dataset for generating the PMD measurements.
- Demonstrated end to end training of the MRM with two phases per frequency. Median depth error got reduced by 50%.
- The motion model did not work as expected.

Future work:

- With more data and time, motion model can be generalized on the test set.
- Motion and multi reflection models can be trained together.
- Training and testing of the models with real camera data.

References I

Bako, S., Vogels, T., McWilliams, B., Meyer, M., Novák, J., Harvill, A., Sen, P., Derose, T., and Rousselle, F. (2017).
Kernel-predicting convolutional networks for denoising monte carlo renderings.
ACM Trans. Graph., 36(4):97–1.

Fischer, P., Dosovitskiy, A., Ilg, E., Häusser, P., Hazırbaş, C., Golkov, V., Van der Smagt, P., Cremers, D., and Brox, T. (2015).
Flownet: Learning optical flow with convolutional networks.
arXiv preprint arXiv:1504.06852.

References II

Guo, Q., Frosio, I., Gallo, O., Zickler, T., and Kautz, J. (2018).
Tackling 3d tof artifacts through learning and the flat dataset.
In *The European Conference on Computer Vision (ECCV)*.

Hansard, M., Lee, S., Choi, O., and Horaud, R. P. (2012).
Time-of-flight cameras: principles, methods and applications.
Springer Science & Business Media.

References III

Lee, S., Kang, B., Kim, J. D., and Kim, C. Y. ((2012)). Motion blur-free time-of-flight range sensor. In *Sensors, Cameras, and Systems for Industrial and Scientific Applications XIII*, volume 8298, page 82980U. International Society for Optics and Photonics.