UNIVERSITY OF TARTU
Institute of Computer Science

NetCalib: A Novel Approach for LiDAR-Camera Auto-Calibration based on Deep Learning

Shan Wu[1], Amnir Hadachi[1], Damien Vivet[2], Yadu Prabhakar[3]

[1] ITS Lab, University of Tartu; [2] ISAE-SUPAERO, Universite´ de Toulouse ; [3] IAM-CED, Northern Alberta Institute of Technology

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Abstract

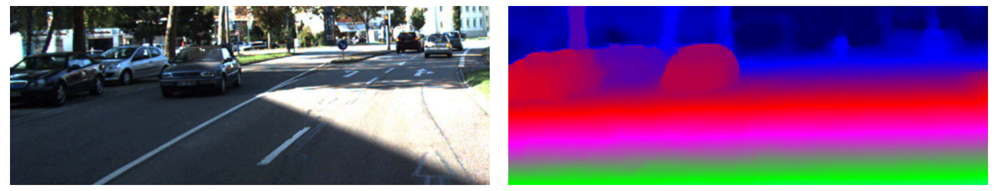
A fusion of LiDAR and cameras have been widely used in many robotics applications such as classification, segmentation, object detection, and autonomous driving. It is essential that the LiDAR sensor can measure distances accurately, which is a good complement to the cameras. Hence, calibrating sensors before deployment is a mandatory step. The main purpose of this research work is to build a deep neural network that is capable of automatically finding the geometric transformation between LiDAR and cameras. The results show that our model manages to find the transformations from randomly sampled artificial errors. Besides, our work is open-sourced for the community to fully utilize the advances of the methodology for developing more the approach, initiating collaboration, and innovation in the topic.

Methodology

Our methodology consists of four components:

- Stereo Matching
- LiDAR Projection
- Artificial Error Generation
- Deep Neural Network

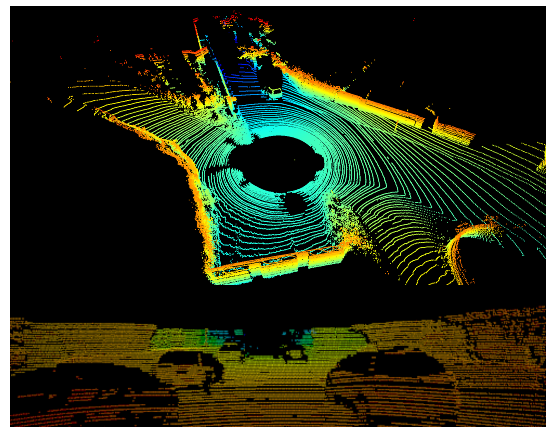
Stereo Matching



- Semi Global Block Matching
- ML methods based on RGB images

$$depth = \frac{B \cdot f}{disparity}$$

LiDAR Projection



$$T = \begin{bmatrix} R(r_x, r_y, r_z) & [t_x, t_y, t_z]^T \\ 0 & 1 \end{bmatrix}$$

$$P = \begin{bmatrix} f_u & 0 & c_u & 0 \\ 0 & f_v & c_v & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$[u, v, 1]^T = PT[x, y, z, 1]^T$$

where R is rotation matrix, t is translation errors, T is transformation matrix, P is projection matrix.

- Transform LiDAR point cloud to the origin of the camera
- Project the transformed point cloud to a 2D frame

Error Generation

- Rotation Error Range: ± 2 degrees
- Translation Error Range: ± 20 cm
- With a uniform distribution
- Encoded in a vector:

$$e = [r_x, r_y, r_z, a, b, c]$$

- Could be transformed to calibration matrix:

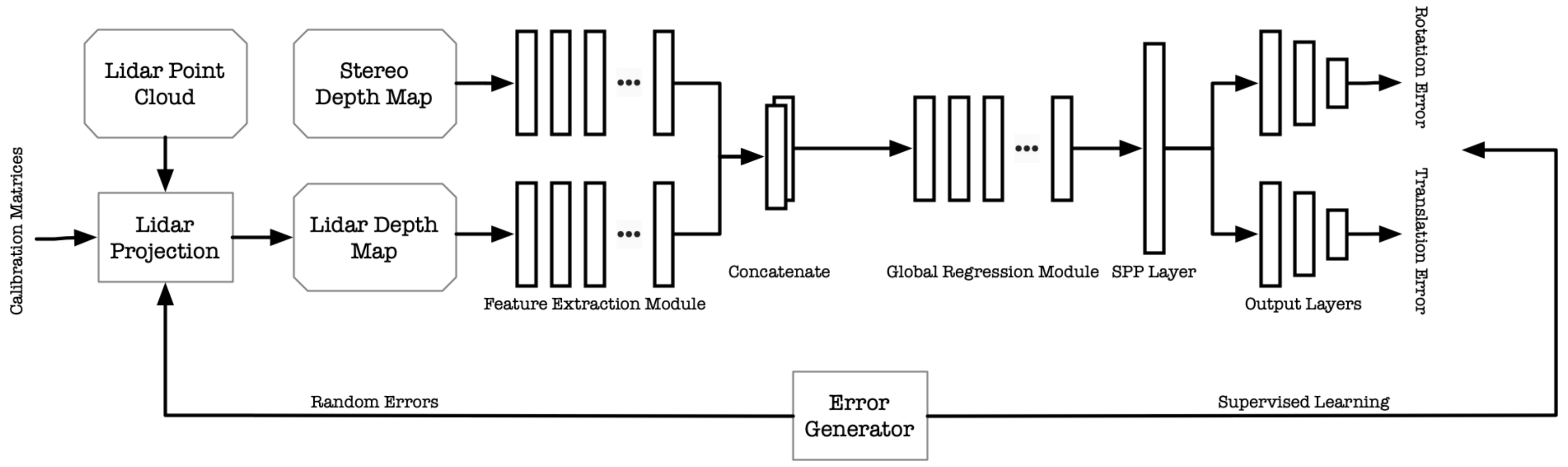
$$T_e = \begin{bmatrix} R(r_x, r_y, r_z) & [a, b, c]^T \\ 0 & 1 \end{bmatrix}$$

- Embedded into LiDAR projection:

$$[u', v', 1]^T = PT_eT[x, y, z, 1]^T$$

$$[u, v, 1]^T = PT^{-1}T_eT[x, y, z, 1]^T$$

CalibNet

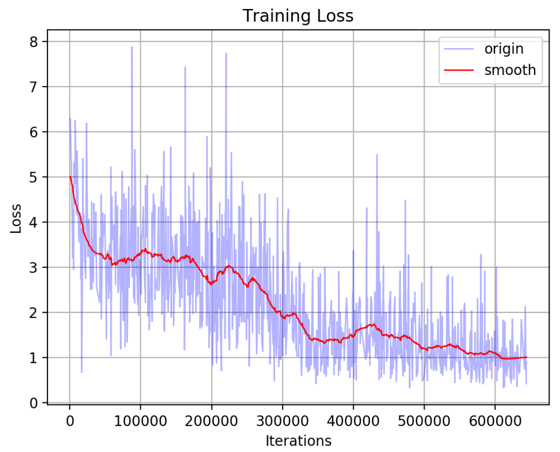


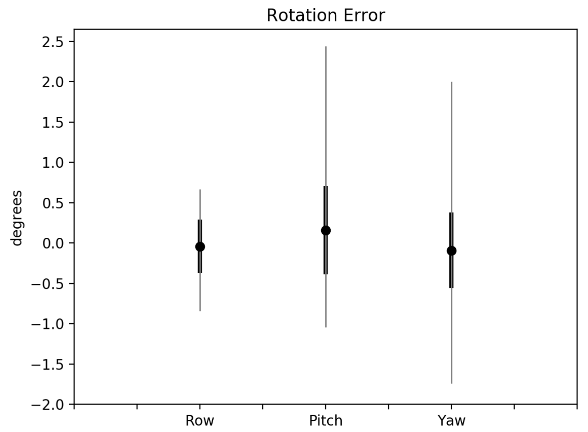
- Feature extraction networks for stereo depth and LiDAR depth respectively
- Output features are concatenated in channels
- Global regression network for the features from both sensors
- Spatial pyramid pooling layer to unify the length of the output features
- Two sets of output layers for rotation error estimation and translation error estimation.

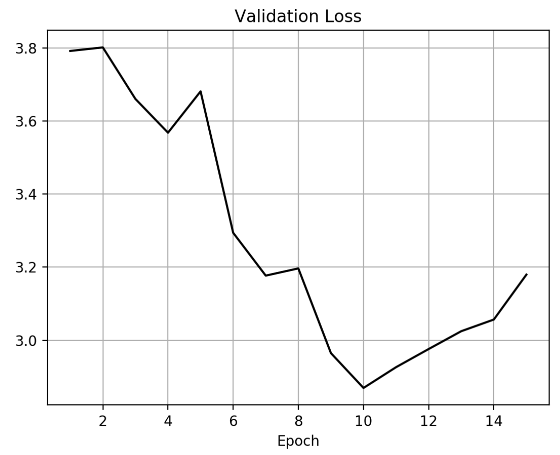
Experiments

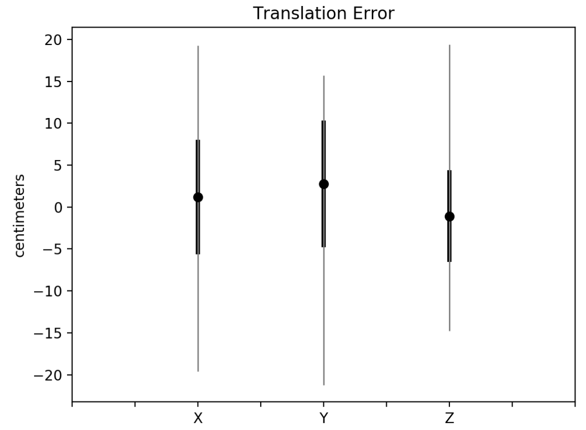
Dataset	Number of Frames
Training	42949
Validation	3168
Testing	258

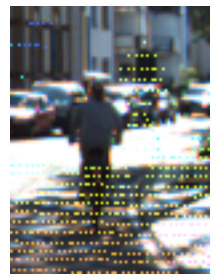
KITTI Dataset



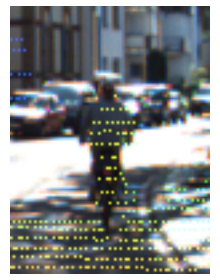




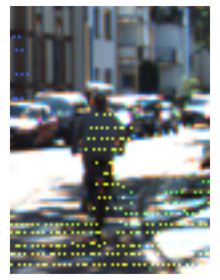




Projection with errors



Projection with model estimation



Ground truth projection

- Training speed: 0.055s/it, Testing speed: 0.016s/it
- Model converged at 10 epochs
- Testing mean L1 loss: 1.0064
- Mean errors:
  - -0.04°, 0.16°, - 0.09° for roll, pitch, yaw.
  - 1.20 cm, 2.77 cm, and -1.10 cm for X, Y, and Z axis
- Standard deviations:
  - 0.33°, 0.55°, and 0.47° for roll, pitch, yaw
  - 6.83 cm, 7.57 cm, 5.47 cm for X, Y, and Z axis

Conclusions & Future Perspectives

- Proposed model can be used for auto-calibration, it shrinks the error ranges
- Model can be adapted for arbitrary input dimensions
- Open sourced at: https://github.com/simonwu53/NetCalib-Lidar-Camera-Auto-calibration
- Improve the quality of input depths could lead to a better performance

References

- N. Schneider, F. Piewak, C. Stiller, and U. Franke, “Regnet: Multimodal sensor registration using deep neural networks,” in 2017 IEEE intelligent vehicles symposium (IV), pp. 1803–1810, IEEE, 2017.
- K. Park, S. Kim, and K. Sohn, “High-precision depth estimation using uncalibrated lidar and stereo fusion,” IEEE Transactions on Intelligent Transportation Systems, 2019.