Camera Calibration Through Camera Projection Loss

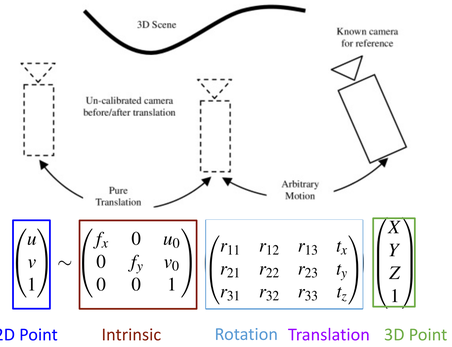
Talha Hanif, Murtaza Taj

Computer Vision and Graphics Lab (CVGL),
Department of Computer Science, Lahore University of Management Sciences (LUMS)

icassp 2022 Singapore
Poster ID: 3448

1. Motivation

Camera Calibration



$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \sim \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

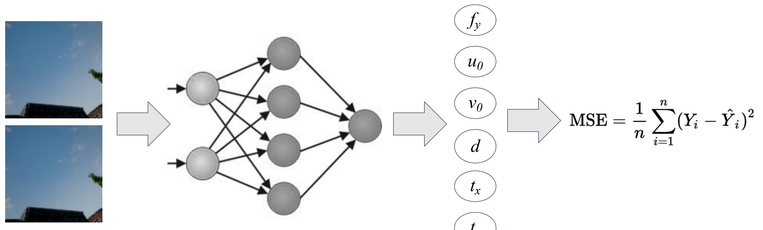
2D Point | Intrinsic | Rotation | Translation | 3D Point

Applications

3D Reconstruction

Autonomous Driving

2. Camera Calibration via End-to-end Learning

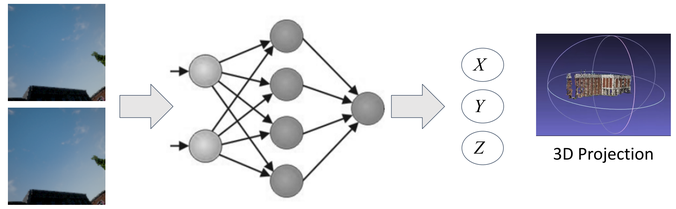


$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

d: Disparity

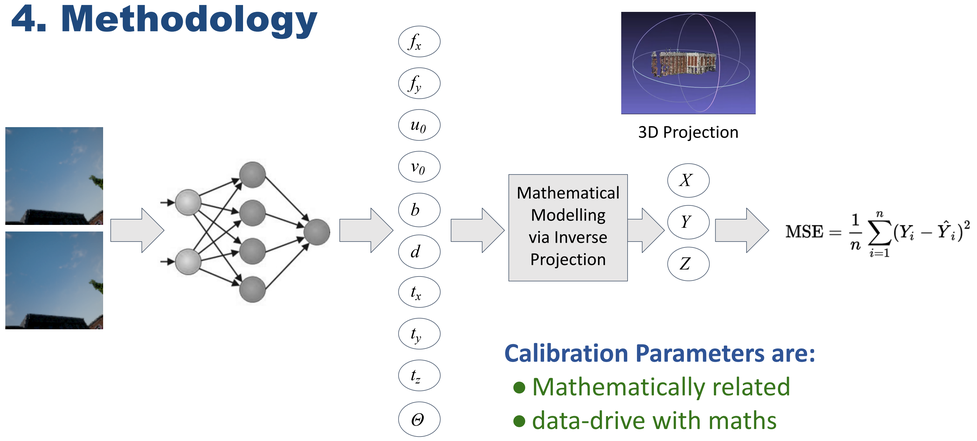
Parameters are:
- Mathematically unrelated
- Totally data-driven

3. 3D Reconstruction via End-to-end Learning



Parameters are:
- Mathematically unrelated
- Totally data-driven

4. Methodology



$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$$

Calibration Parameters are:
- Mathematically related
- data-drive with maths

Inverse Projection

Image->Camera->World

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_{3\times1}^T & 1 \end{pmatrix}^{-1} \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

Image->Camera

$$x_{cam} = f_x * b/d$$

$$y_{cam} = -(x_{cam}/f_x) * (u - u_0)$$

$$z_{cam} = (x_{cam}/f_y) * (v_0 - v)$$

Camera->World

$$X = x_{cam} * \cos\theta + z_{cam} * \sin\theta + t_x$$

$$Y = y_{cam} + t_y$$

$$Z = -x_{cam} * \sin\theta + z_{cam} * \cos\theta + t_z$$

5. Datasets

CVGL Camera Calibration Dataset
- Synthetic
- via CARLA Simulator
- 2 Towns
- 49 Camera Configurations
- 79,320 image pairs

Tsinghua-Daimler Cyclist Detection
- 2,914 images comprising of the test set used for evaluation

6. Results & Evaluation

Evaluation on CVGL Camera Calibration Dataset
- via Normalised Mean Absolute Error

Method	f_x	f_y	u_0	v_0	b	d	t_x	t_y	t_z	Θ
Average [1]	0.840	0.786	0.432	0.542	6.552	3.607	6.552	9.372	5.361	0.74
Deep-Homo [2]	0.062	0.062	0.008	0.008	0.156	0.065	0.156	0.161	0.155	0.05
MTL-CPL-U	0.935	0.685	0.892	0.737	0.938	0.432	0.400	0.329	0.432	1.06
MTL-Baseline	0.030	0.029	0.017	0.007	0.057	0.013	0.064	0.076	0.071	0.02
MTL-CPL-A	0.022	0.022	0.004	0.006	0.093	0.007	0.097	0.116	0.098	0.02

Evaluation on Tsinghua-Daimler Cyclist Detection Benchmark (without any training or transfer learning)
- via Normalised Mean Absolute Error

Method	f_x	f_y	u_0	v_0	b	d	t_x	t_y	t_z	Θ
Average [1]	0.994	0.991	0.969	0.951	112.438	0.492	10.843	271.935	13.798	982.41
Deep-Homo [2]	0.958	0.958	0.946	0.895	9.985	1.233	0.166	27.141	0.862	2746.99
MTL-CPL-U	0.872	0.888	0.782	0.795	0.081	1.271	0.147	23.836	0.635	7700.97
MTL-Baseline	0.957	0.958	0.944	0.893	18.323	1.258	1.035	32.946	0.999	2418.25
MTL-CPL-A	0.938	0.938	0.946	0.895	14.182	1.259	0.727	30.640	1.418	1995.35

1. Workman, Scott, et al. "Deepfocal: A method for direct focal length estimation." *2015 IEEE International Conference on Image Processing*. IEEE, 2015.
2. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798*, 2016.

7. Summary & Conclusions

- A new **dataset** for Camera Calibration.
- A **new representation** to incorporate camera model equations in a neural network in a multi-task learning framework.
- A new **loss utilising camera model** neural network to reconstruct 3D projection and uses the reconstruction loss to estimate the camera parameters.
- The proposed method **performs better** than both traditional and learning based methods.

Project Page: https://cvlab.lums.edu.pk/cpl
Email: murtaza.taj@lums.edu.pk
l181864@lhr.nu.edu.pk
GitHub: https://github.com/thanif/CPL