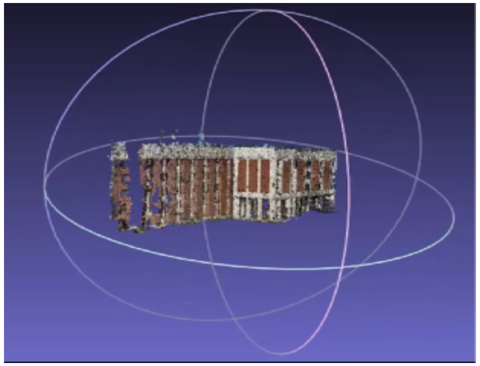
Camera Calibration through Camera Projection Loss

Talha Hanif, Murtaza Taj

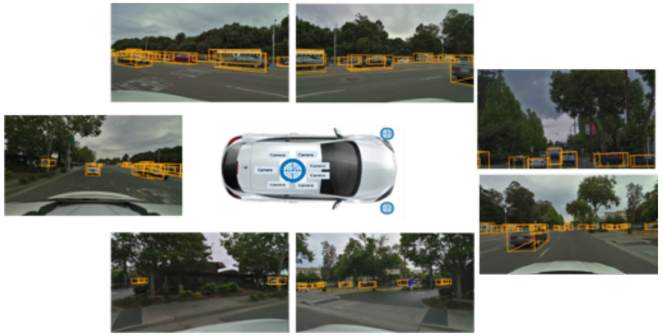
Computer Vision and Graphics Lab (CVGL),
Department of Computer Science,
Lahore University of Management Sciences (LUMS)

Motivation



Photogrammetry

3D Reconstruction

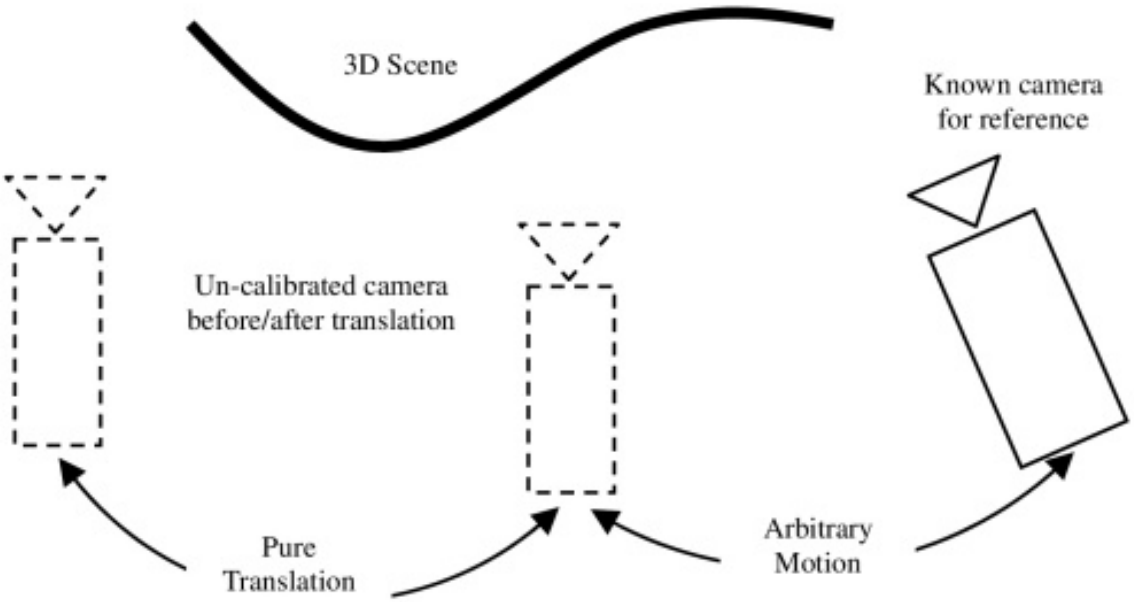


Autonomous Driving

Introduction

Camera Model

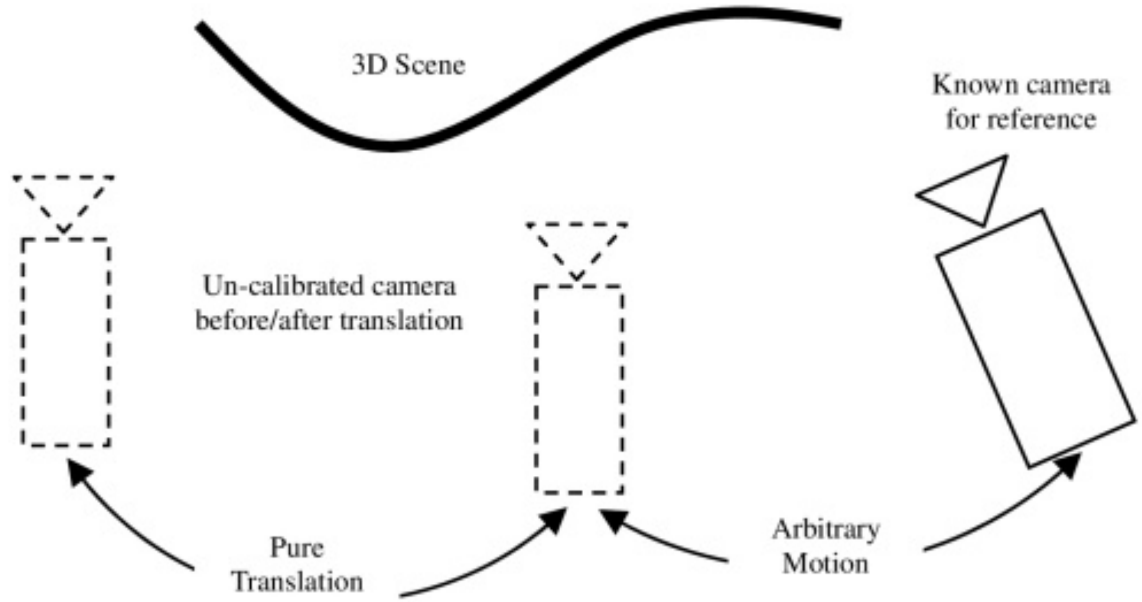
- Extrinsic
- Intrinsic



Introduction

Camera Model

- Extrinsic
- Intrinsic



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

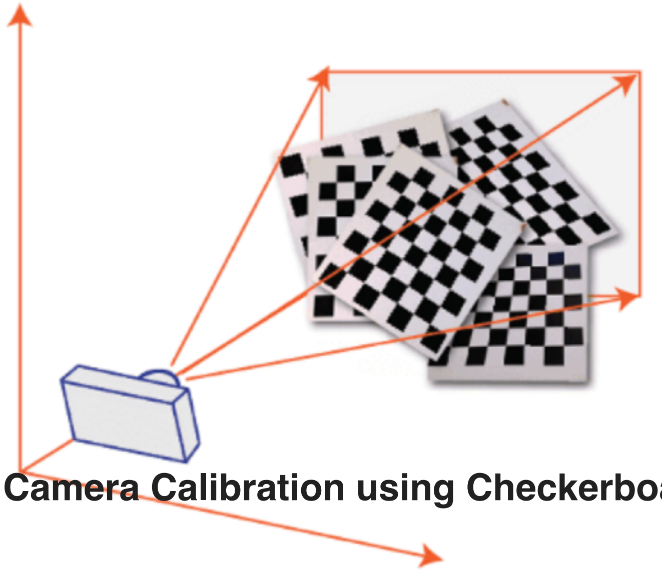
2D Point | Intrinsic | Rotation | Translation | 3D Point

Background Literature

Camera Calibration

via Hand-crafted Features

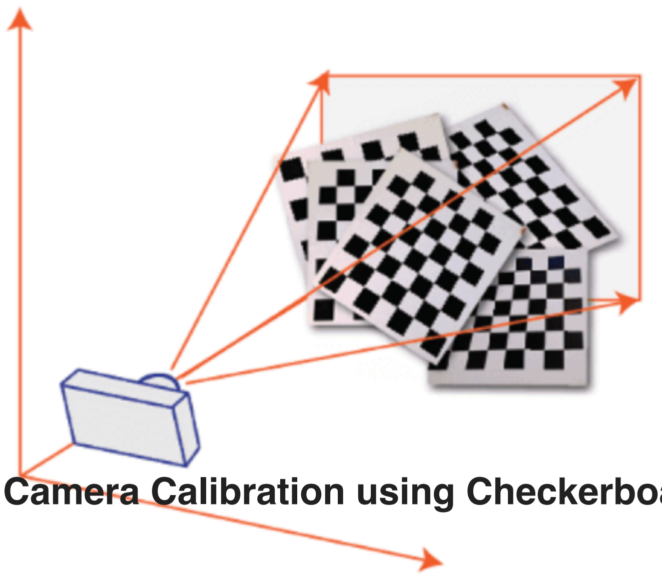
Camera Calibration
via Hand-crafted Features



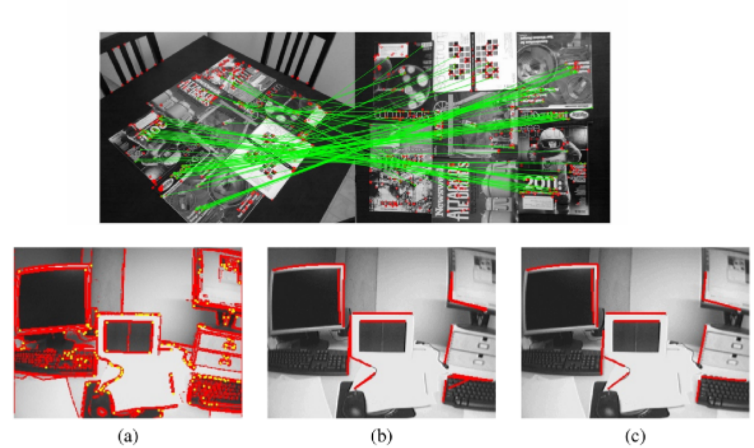
Camera Calibration using Checkerboard Pattern

Camera Calibration

via Hand-crafted Features



Camera Calibration using Checkerboard Pattern

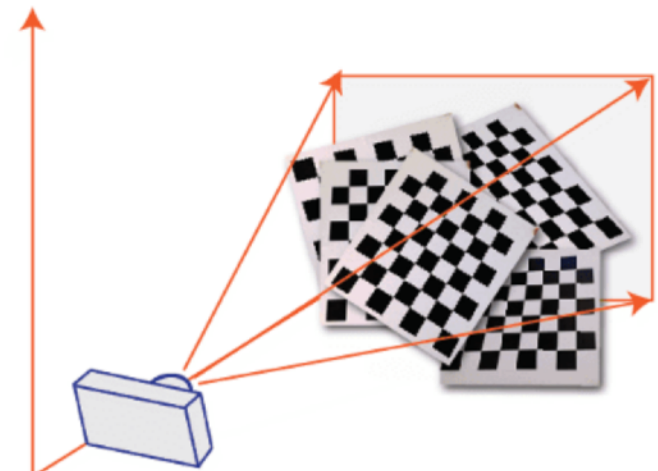


ORB + RANSAC [1]

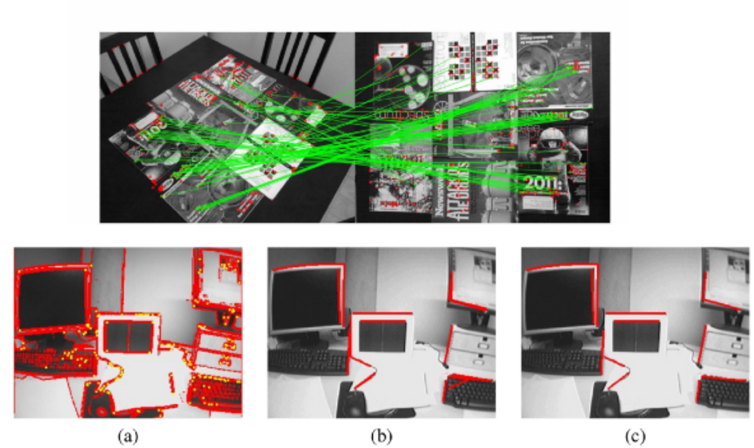
1. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798* (2016).

Camera Calibration

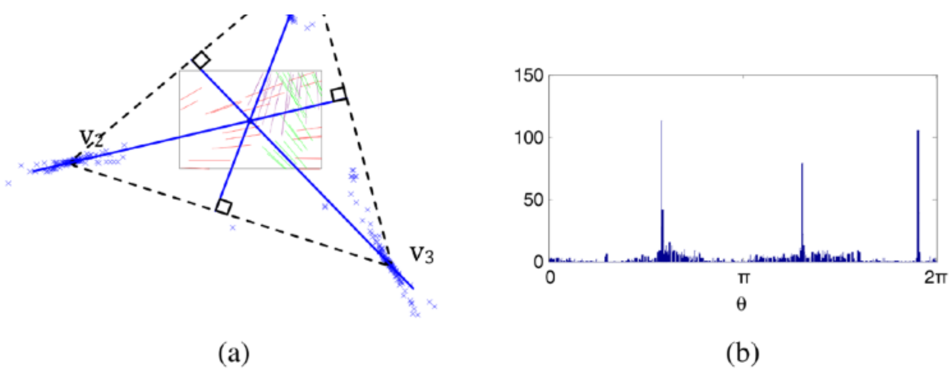
via Hand-crafted Features



Camera Calibration using Checkerboard Pattern



ORB + RANSAC [1]

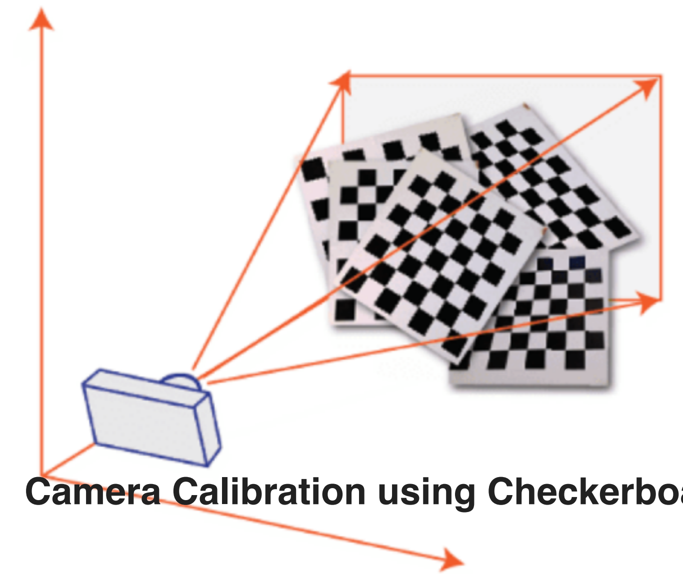


Li et al. [2]

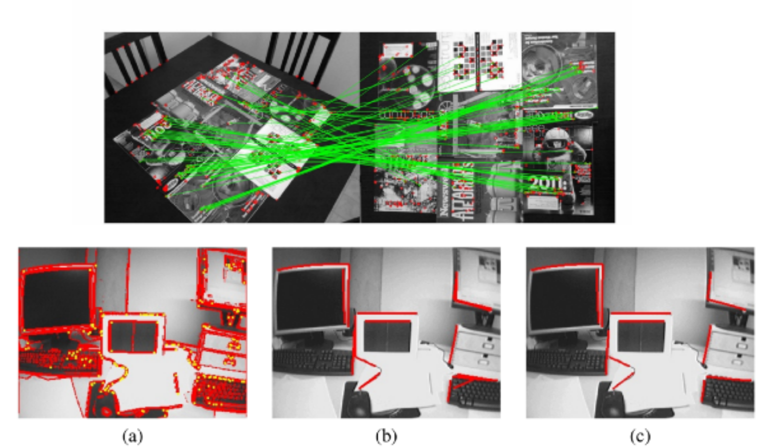
1. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798* (2016).
2. Bo Li, Kun Peng, Xianghua Ying, and Hongbin Zha, “Simultaneous vanishing point detection and camera calibration from single images,” in ISVC, 2010.

Camera Calibration

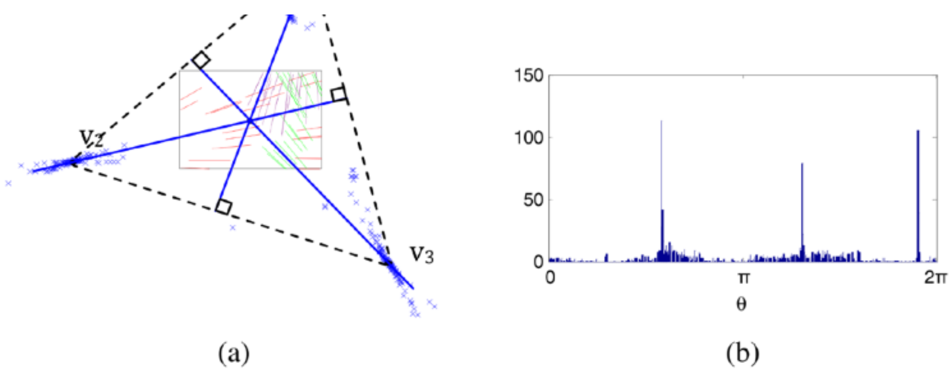
via Hand-crafted Features



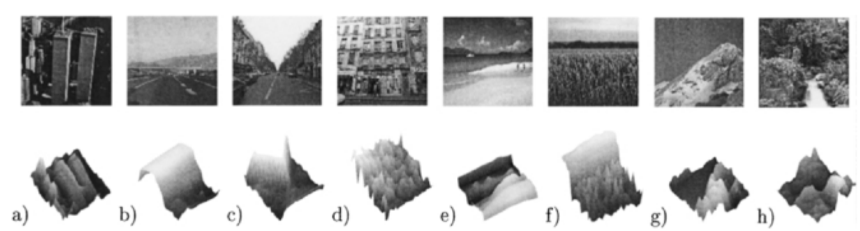
Camera Calibration using Checkerboard Pattern



ORB + RANSAC [1]



Li et al. [2]



GIST + Random Forest [3]

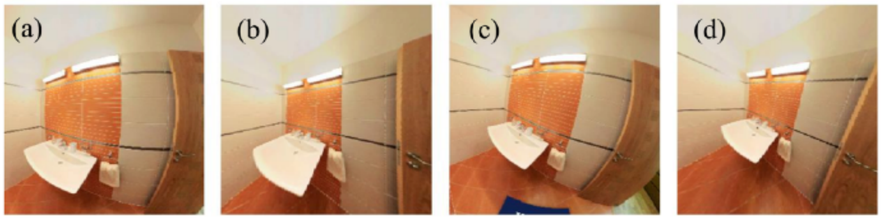
1. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798* (2016).
2. Bo Li, Kun Peng, Xianghua Ying, and Hongbin Zha, "Simultaneous vanishing point detection and camera calibration from single images," in ISVC, 2010.
3. Workman, Scott, et al. "Deepfocal: A method for direct focal length estimation." *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015.

Camera Calibration

End-to-end learning

Camera Calibration

End-to-end learning

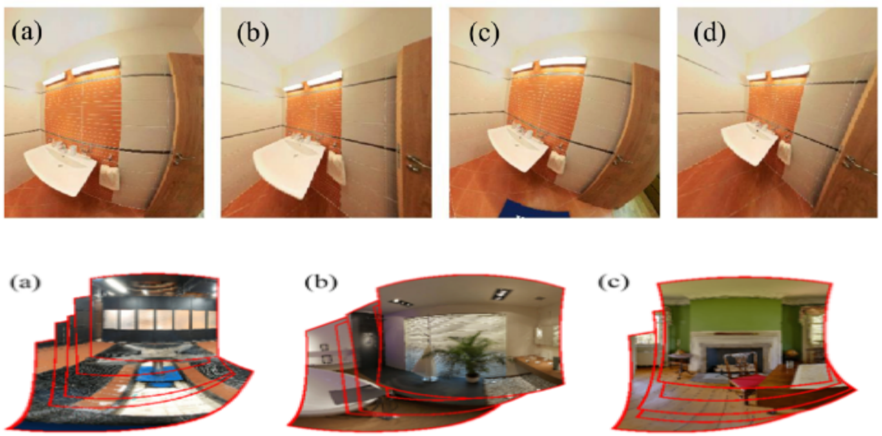


Deep-PTZ [1]
(focal length, distortion, rotation)

1. Zhang, Chaoning, et al. "Deepptz: Deep self-calibration for PTZ cameras." *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.

Camera Calibration

End-to-end learning



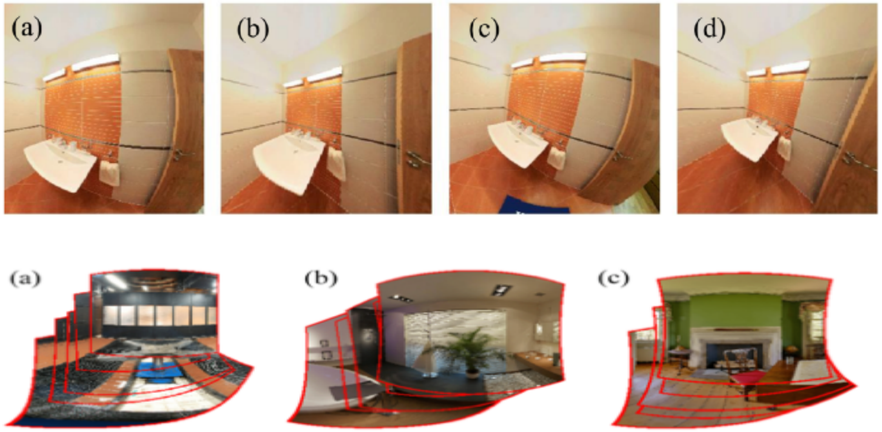
Deep-PTZ [1]
(focal length, distortion, rotation)

Deep-Focal [2]
(focal length)

1. Zhang, Chaoning, et al. "Deepptz: Deep self-calibration for PTZ cameras." *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.
2. Workman, Scott, et al. "Deepfocal: A method for direct focal length estimation." *2015 IEEE International Conference on Image Processing*. IEEE, 2015.

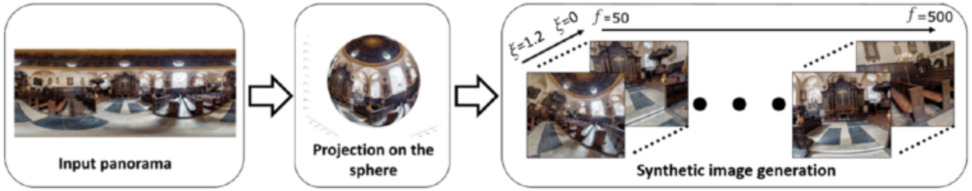
Camera Calibration

End-to-end learning



Deep-PTZ [1]
(focal length, distortion, rotation)

Deep-Focal [2]
(focal length)

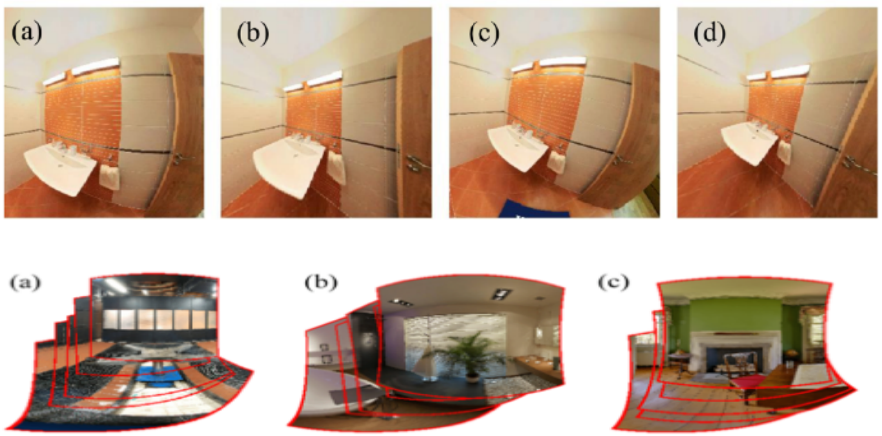


Deep-Calib [3]
(focal length, distortion)

1. Zhang, Chaoning, et al. "Deepptz: Deep self-calibration for PTZ cameras." *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.
2. Workman, Scott, et al. "Deepfocal: A method for direct focal length estimation." *2015 IEEE International Conference on Image Processing*. IEEE, 2015.
3. Bogdan, Oleksandr, et al. "DeepCalib: a deep learning approach for automatic intrinsic calibration of wide field-of-view cameras." *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*. 2018.

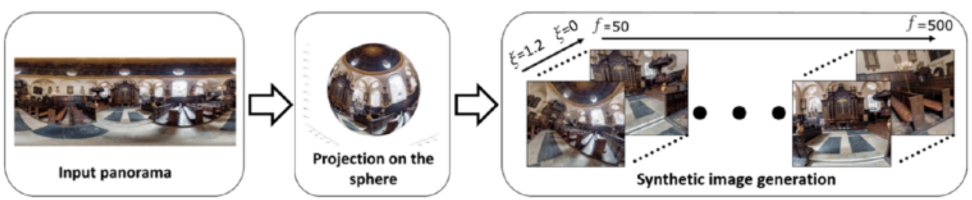
Camera Calibration

End-to-end learning

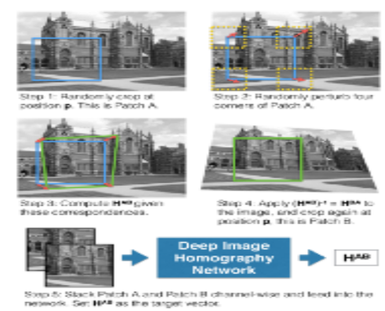


Deep-PTZ [1]
(focal length, distortion, rotation)

Deep-Focal [2]
(focal length)



Deep-Calib [3]
(focal length, distortion)



Deep-Homo [4]
(homography)

1. Zhang, Chaoning, et al. "Deepptz: Deep self-calibration for ptz cameras." *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2020.
2. Workman, Scott, et al. "Deepfocal: A method for direct focal length estimation." *2015 IEEE International Conference on Image Processing*. IEEE, 2015.
3. Bogdan, Oleksandr, et al. "DeepCalib: a deep learning approach for automatic intrinsic calibration of wide field-of-view cameras." *Proceedings of the 15th ACM SIGGRAPH European Conference on Visual Media Production*. 2018.
4. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Deep image homography estimation." *arXiv preprint arXiv:1606.03798*, 2016.

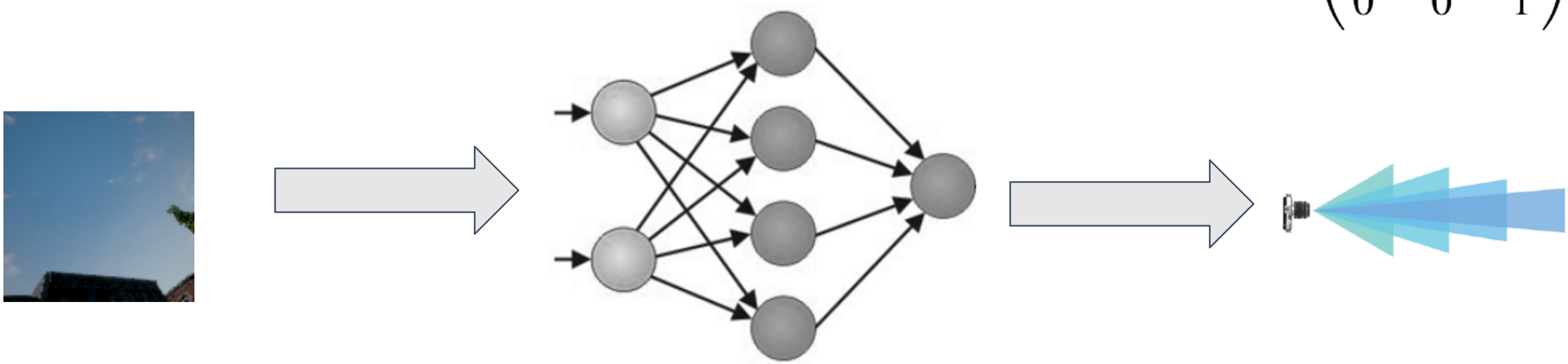
Methodology

Camera Calibration

Intrinsic

- via Single Image

$$\begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

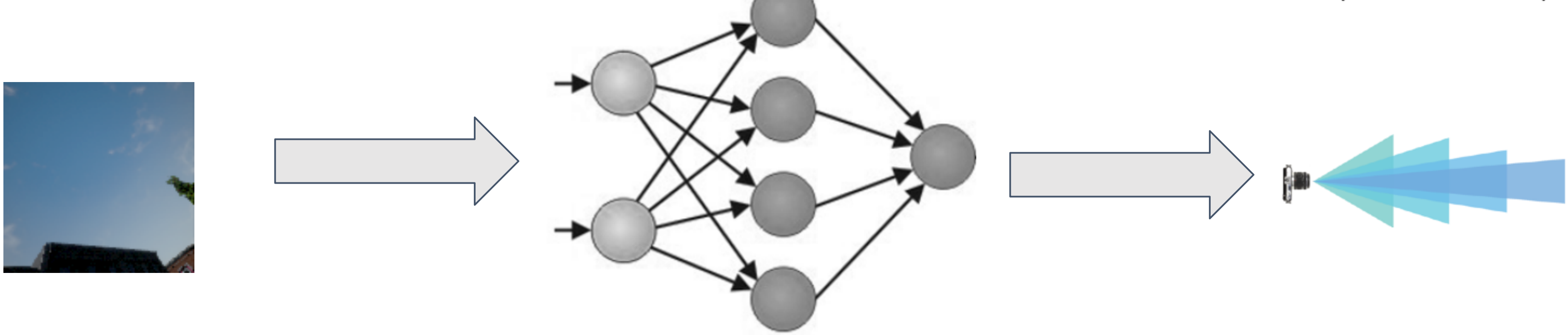


Camera Calibration

Intrinsic

- via Single Image

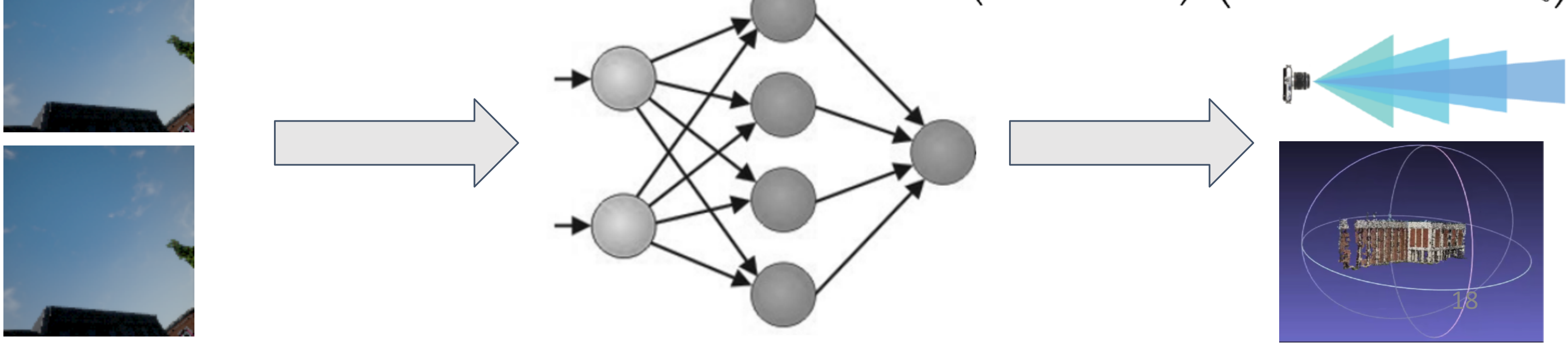
$$\begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$



Intrinsic & Extrinsic Both

- via Image Pair

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

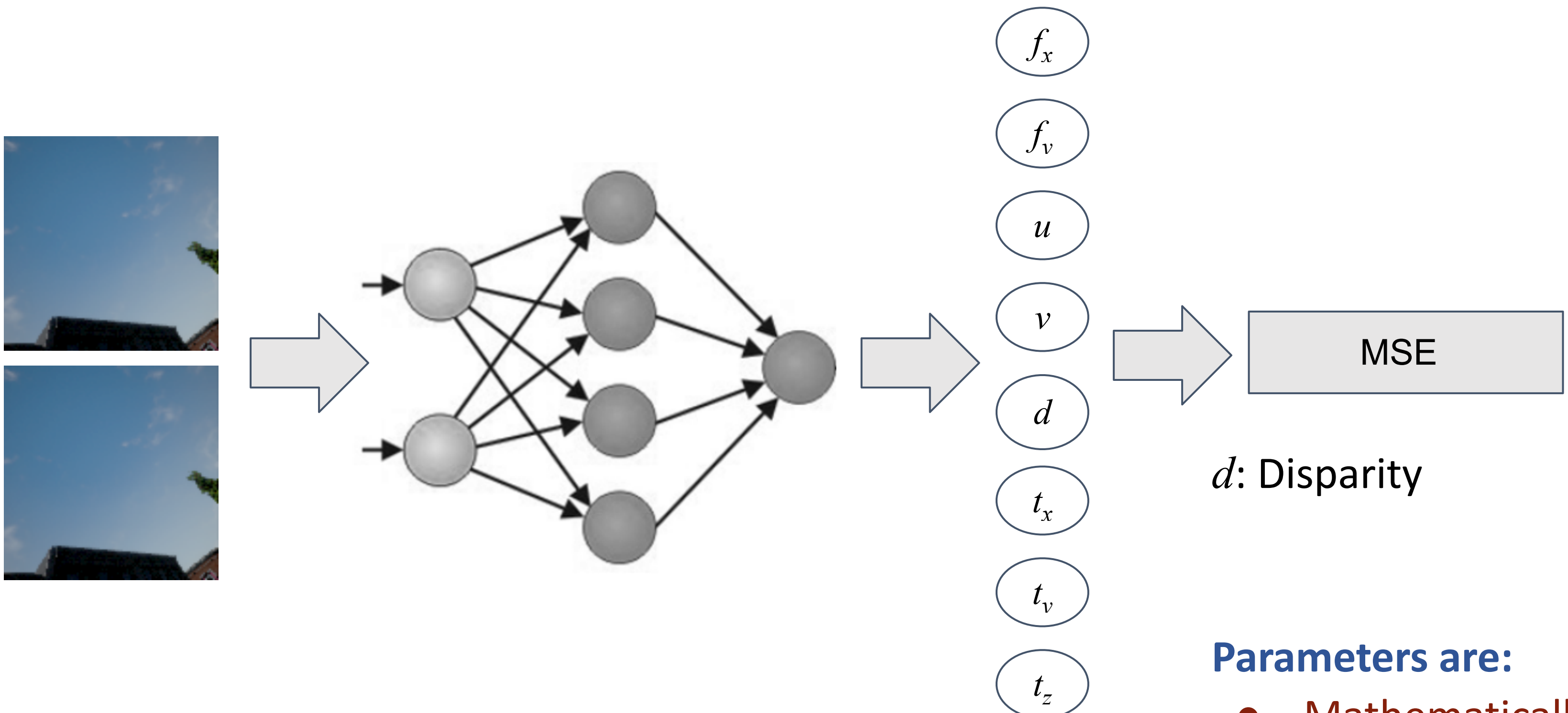


Camera Calibration

via End-to-end learning

Intrinsic & Extrinsic Both

- via Image Pair



d: Disparity

Parameters are:

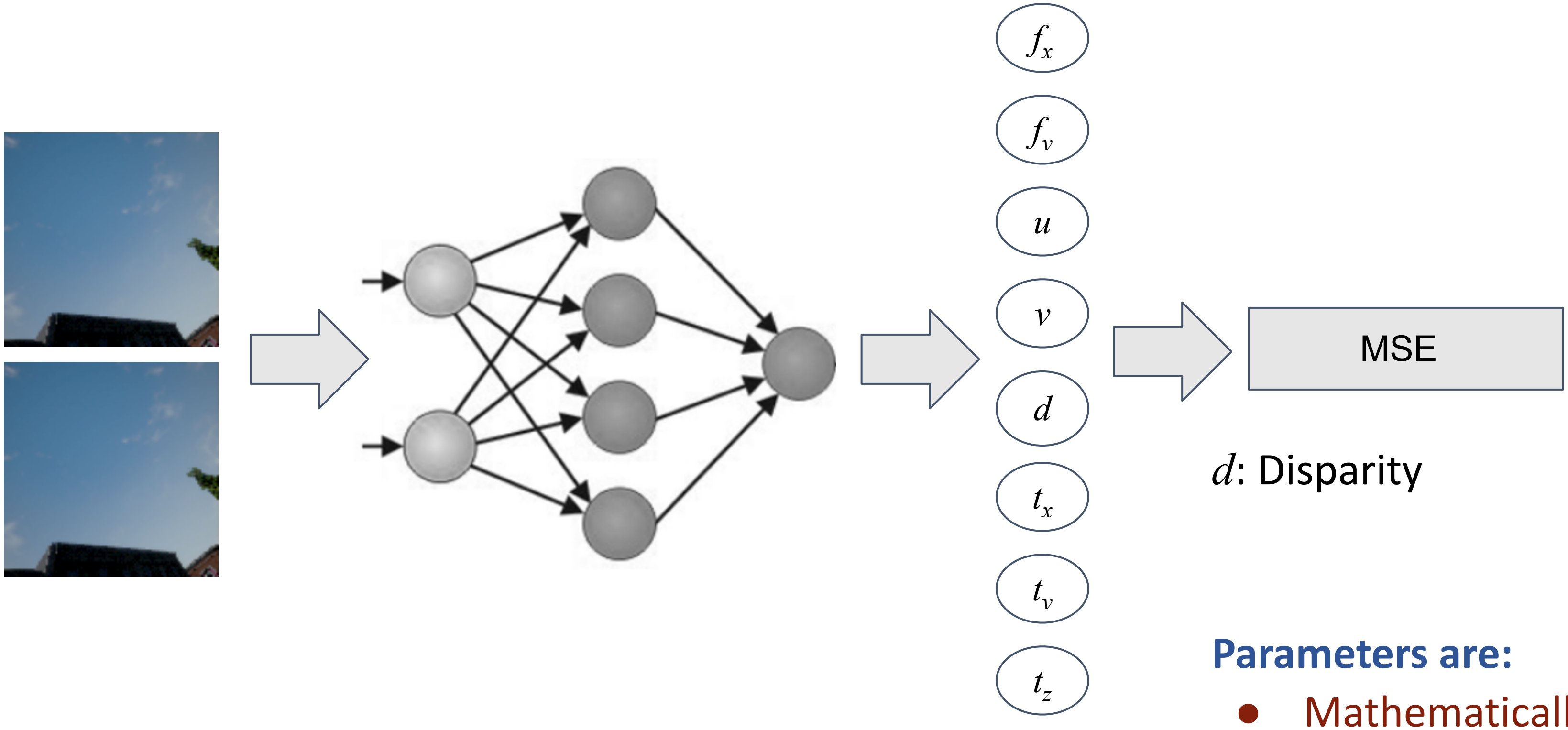
- Mathematically unrelated
- Totally data-driven

Camera Calibration

via End-to-end learning

Intrinsic & Extrinsic Both

- via Image Pair



d: Disparity

Parameters are:

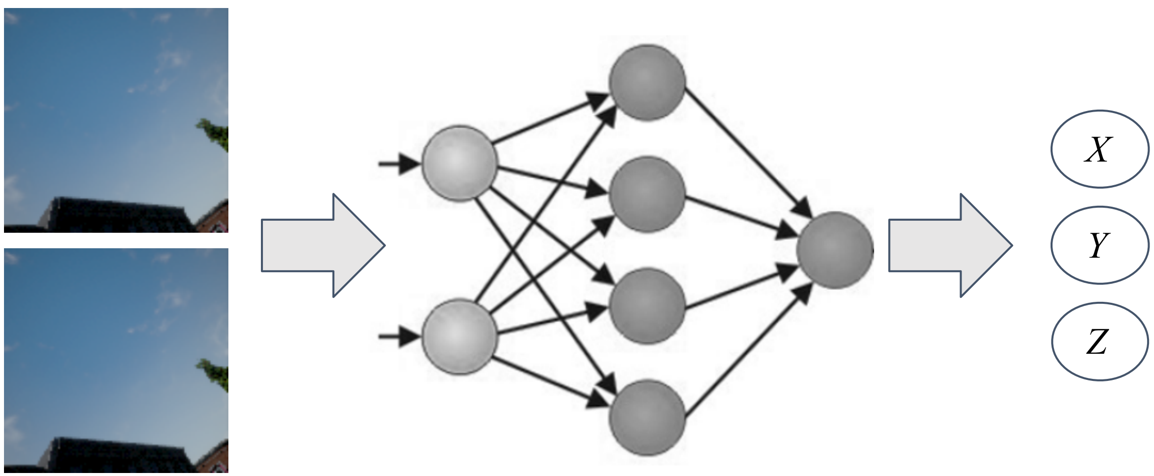
- Mathematically unrelated
- Totally data-driven

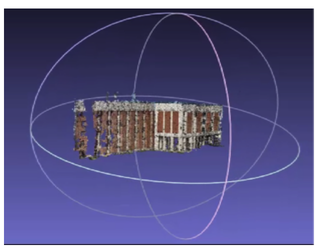
3D Projection

via End-to-end learning

Intrinsic & Extrinsic Both

- via Image Pair





3D Projection

Parameters are:

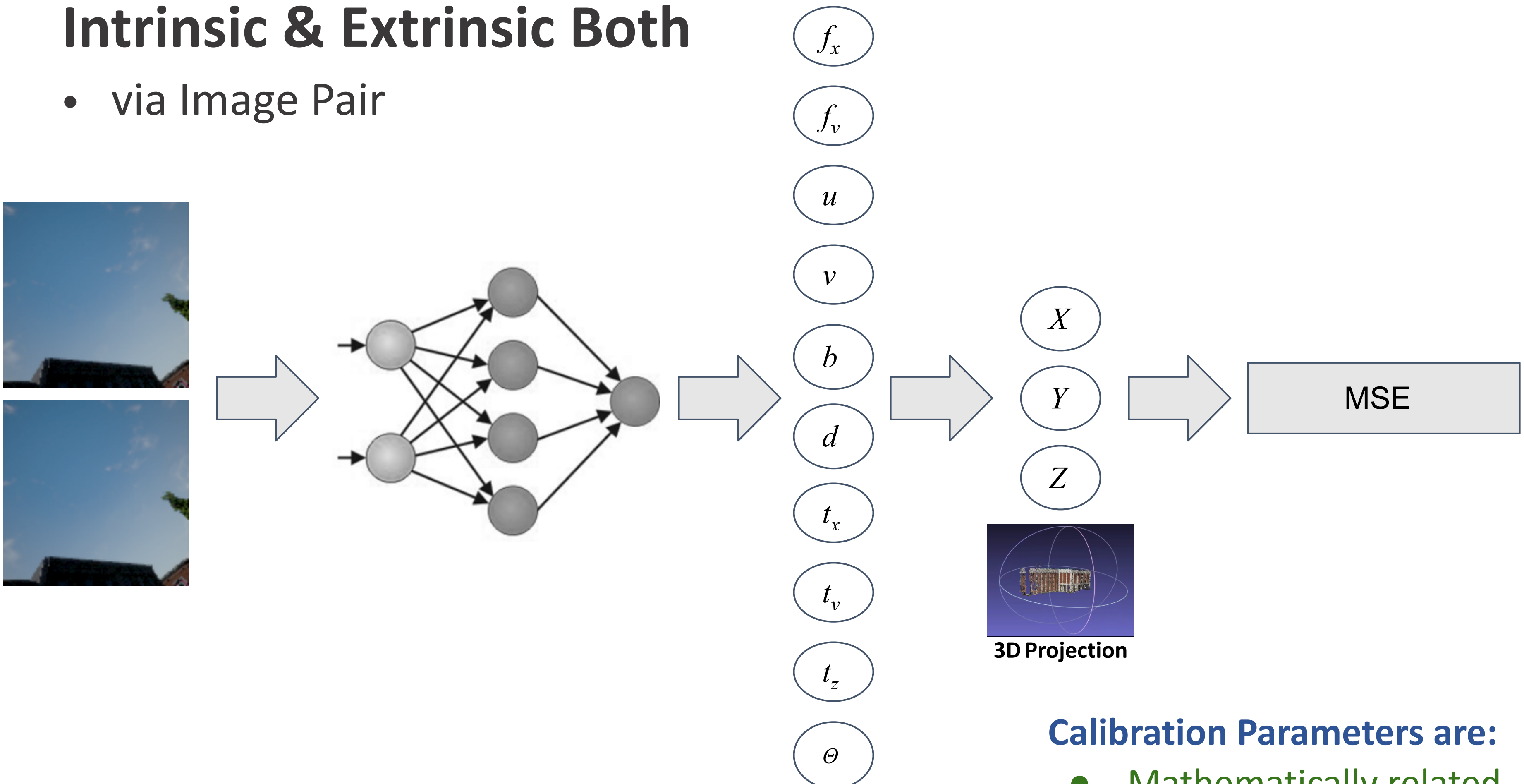
- Mathematically unrelated
- Totally data-drive

Camera Calibration

via Camera Projection Loss

Intrinsic & Extrinsic Both

- via Image Pair



Calibration Parameters are:

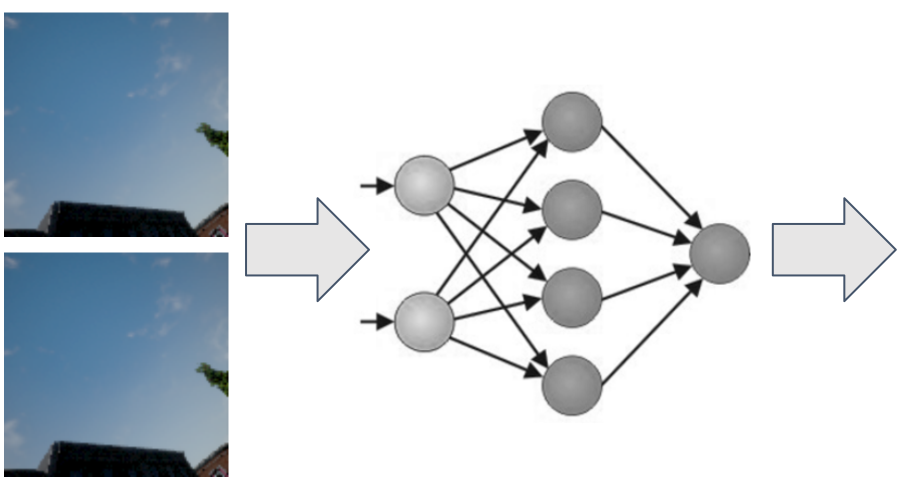
- Mathematically related
- data-drive with maths

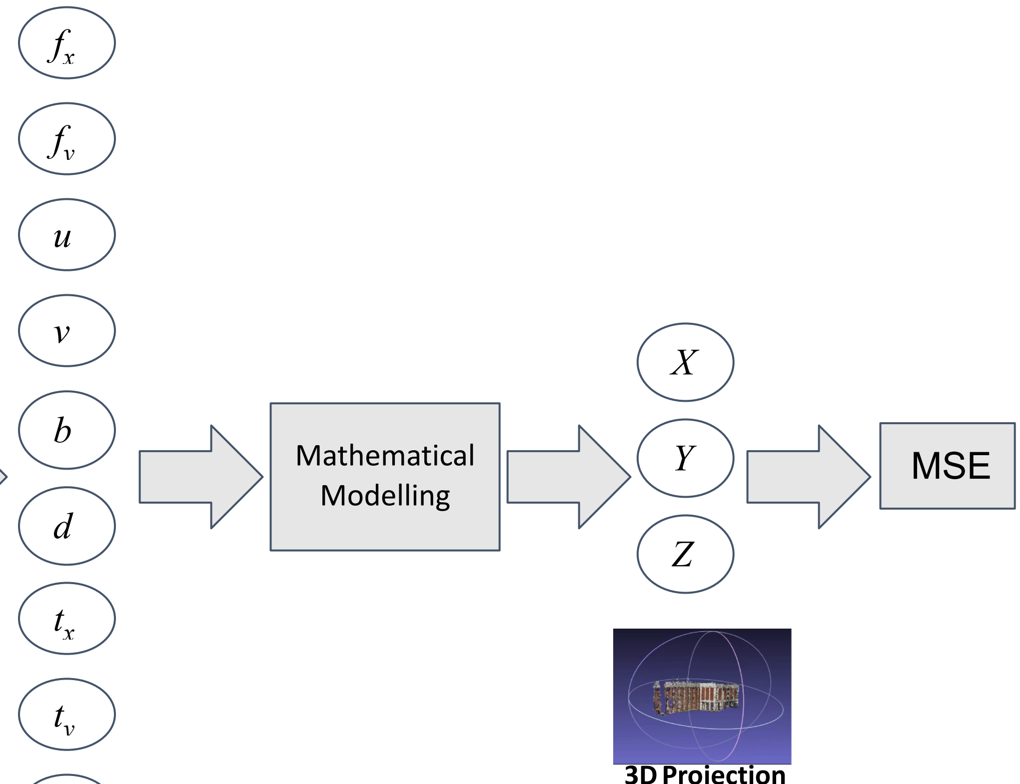
Camera Calibration

via Camera Projection Loss

3D Projection

- via Mathematical Modelling





Calibration Parameters are:

- Mathematically related
- data-drive with maths

Camera Calibration

via Inverse Projection

Mathematical Modelling

- via Image to Camera to World

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

Camera Calibration

via Inverse Projection

Mathematical Modelling

- via Image to Camera to World

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

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

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

Camera Calibration

via Inverse Projection

Mathematical Modelling

- via Image to 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 Calibration

via Inverse Projection

Mathematical Modelling

- via Camera to World

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \sim \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_{3\times 1}^T & 1 \end{pmatrix} \begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \sim \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \begin{pmatrix} x_{cam} \\ y_{cam} \\ z_{cam} \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

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

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

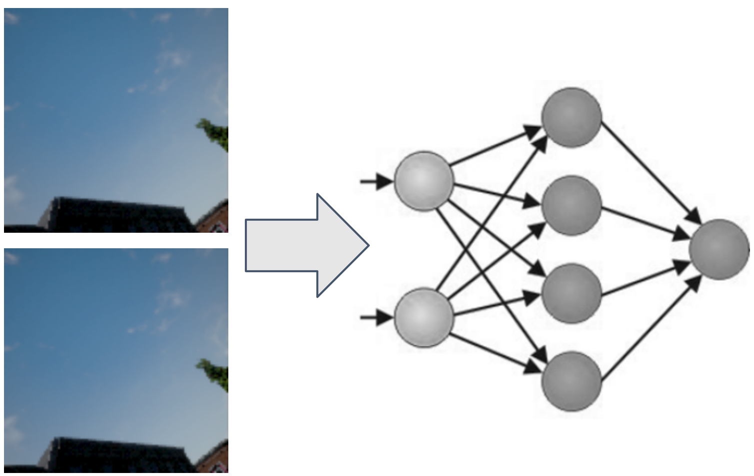
$$Z = -x_{cam} * \sin\theta + z_{cam} * \cos\theta + t_z$$

Camera Calibration

via Camera Projection Loss

Mathematical Modelling

- via Lambda Layers



f_x, f_y, u_0, v_0, b, d, t_x, t_y, t_z, Θ

Mathematical Modelling

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

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

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

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

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

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

X, Y, Z

Results and Evaluation

Datasets

CVGL Camera Calibration Dataset

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

Datasets

CVGL Camera Calibration Dataset

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

Tsinghua-Daimler Cyclist Detection Benchmark

- 2,914 images comprising of the test set used for evaluation

Quantitative 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	0.744
Deep-Homo [2]	0.062	0.062	0.008	0.008	0.156	0.065	0.156	0.161	0.155	0.045
MTL-CPL-U	0.935	0.685	0.892	0.737	0.938	0.432	0.400	0.329	0.432	1.060
MTL-Baseline	0.030	0.029	0.017	0.007	0.057	0.013	0.064	0.076	0.071	0.024
MTL-CPL-A	0.022	0.022	0.004	0.006	0.093	0.007	0.097	0.116	0.098	0.017

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.

Quantitative Evaluation

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	**982.413**
Deep-Homo [2]	0.958	0.958	0.946	0.895	9.985	1.233	0.166	27.141	0.862	2746.994
MTL-CPL-U	**0.872**	**0.888**	**0.782**	**0.795**	**0.081**	1.271	**0.147**	**23.836**	**0.635**	7700.968
MTL-Baseline	0.957	0.958	0.944	0.893	18.323	1.258	1.035	32.946	0.999	2418.250
MTL-CPL-A	0.938	0.938	0.946	0.895	14.182	1.259	0.727	30.640	1.418	1995.353

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.

Summary & Conclusions

Summary

- 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.

Please come to the poster for further details!

Poster Session: IVMSP-36: Camera Calibration & Human Pose

Time: Thurs, 12 May, 21:00-21:45 (Singapore Time)

Email: murtaza.taj@lums.edu.pk
l181864@lhr.nu.edu.pk

Project Page: https://cvlab.lums.edu.pk/cpl

GitHub: https://github.com/thanif/CPL

CVG Lab Website: https://cvlab.lums.edu.pk