ICPR 2020

Orchestrating a brighter world

NEC

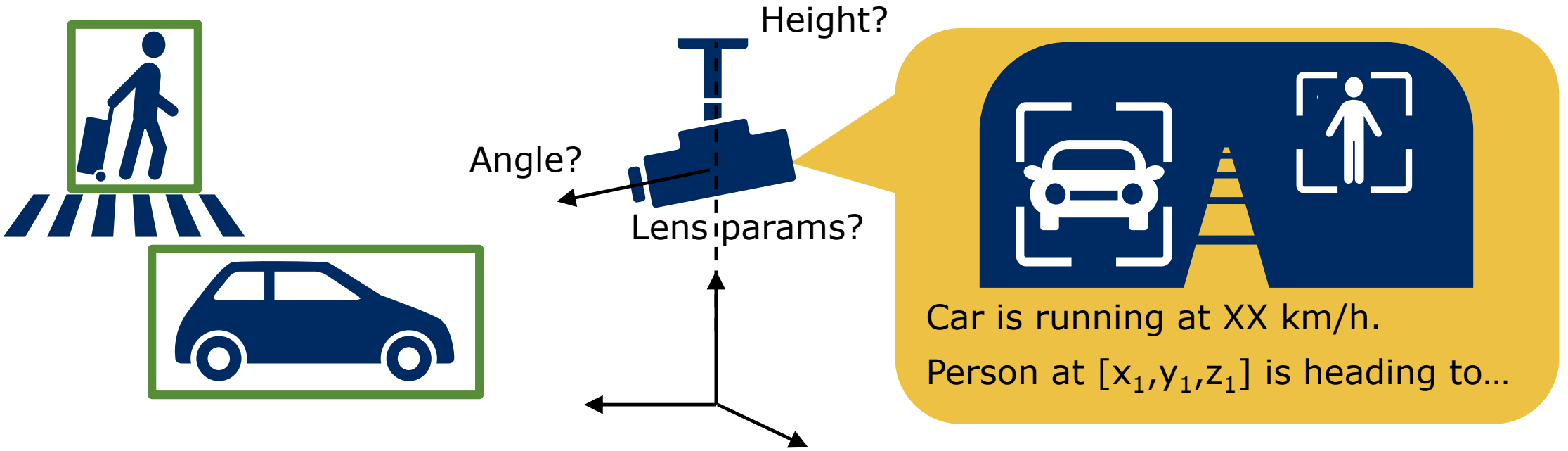
# Camera Calibration using Parallel Line Segments

**Gaku Nakano**
g-nakano@nec.com
Central Research Labs, NEC Corporation, Japan

# Introduction

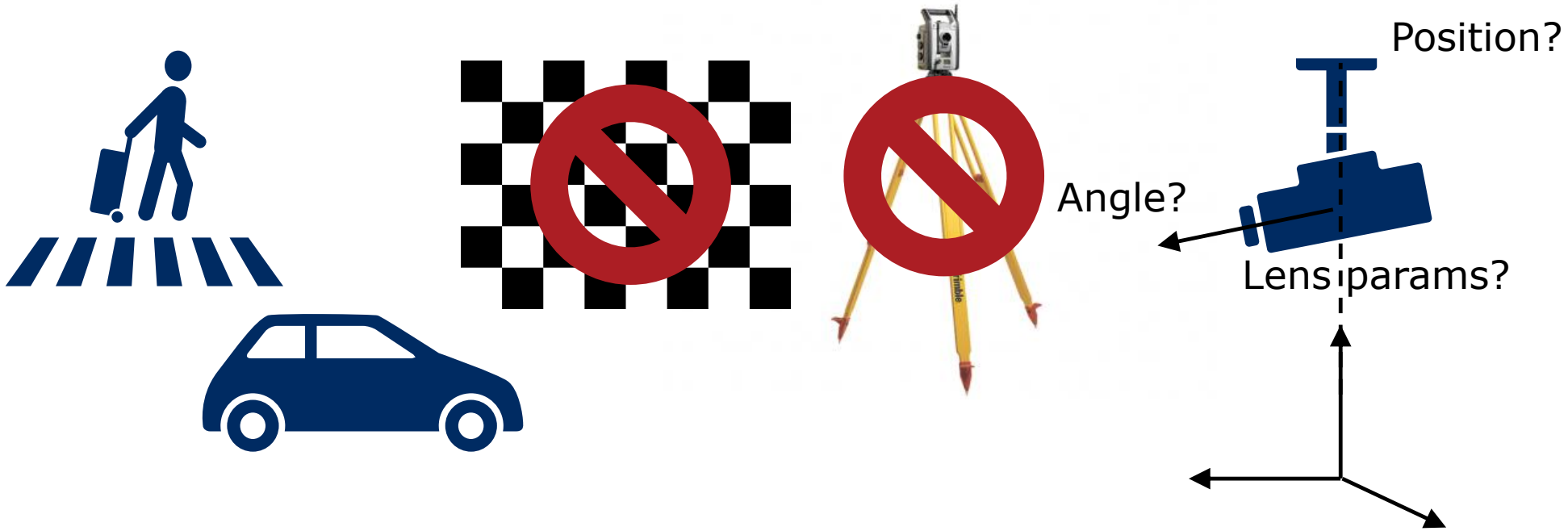
# 3D analysis from 2D video of surveillance cameras

- Traffic estimation, pedestrian monitoring,…
- Requires camera parameters in 3D scene



# Requirements for surveillance camera calibration

- No large calibration objects
- No traffic restrictions



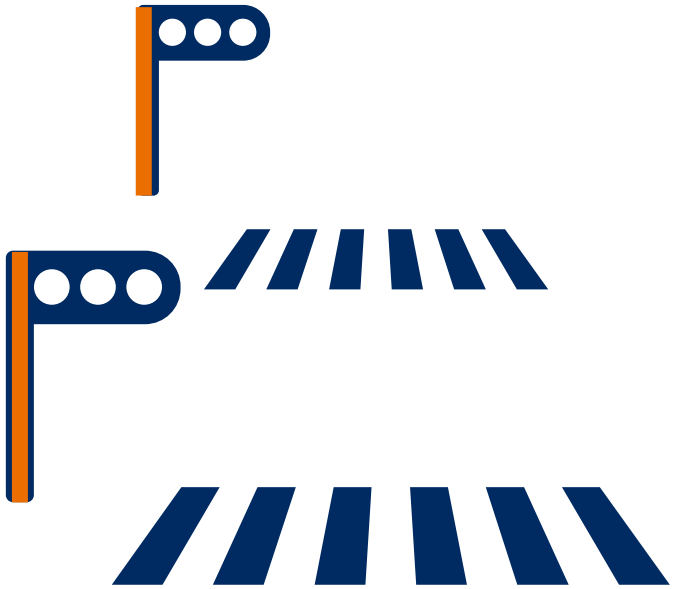

**We propose a practical method for calibrating surveillance cameras.**

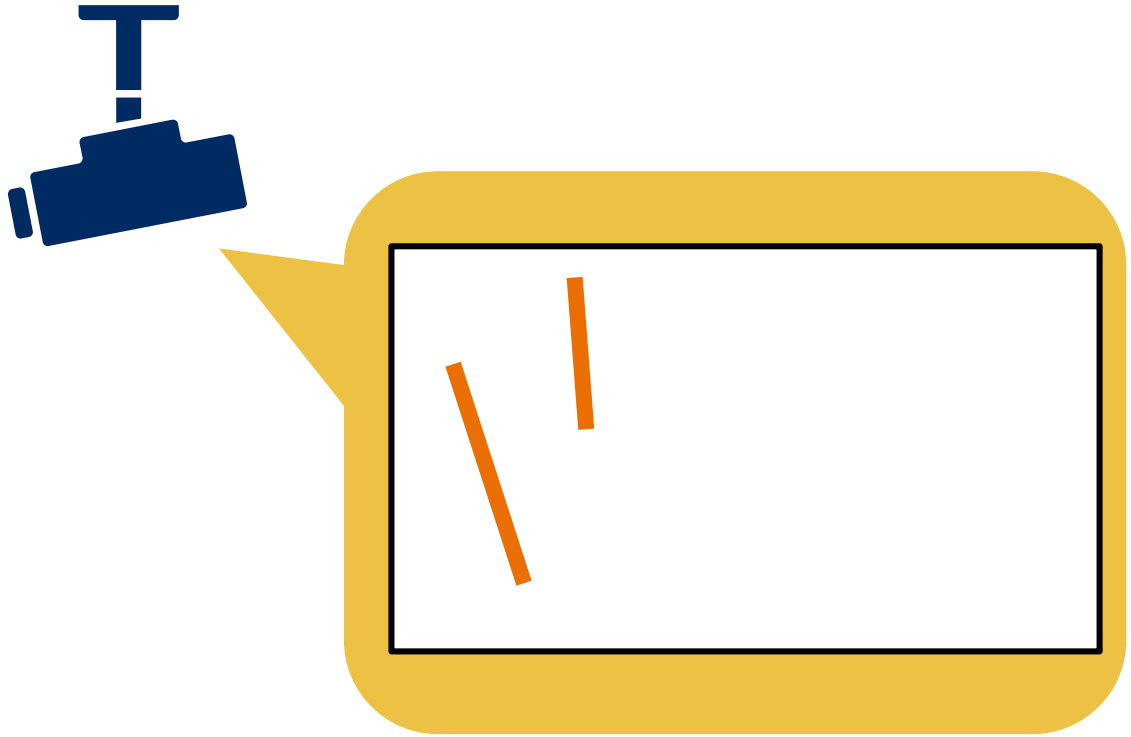
# Key Idea

# Key Idea

## Parallel line segments of the same length as calibration object

- Traffic light





# Key Idea

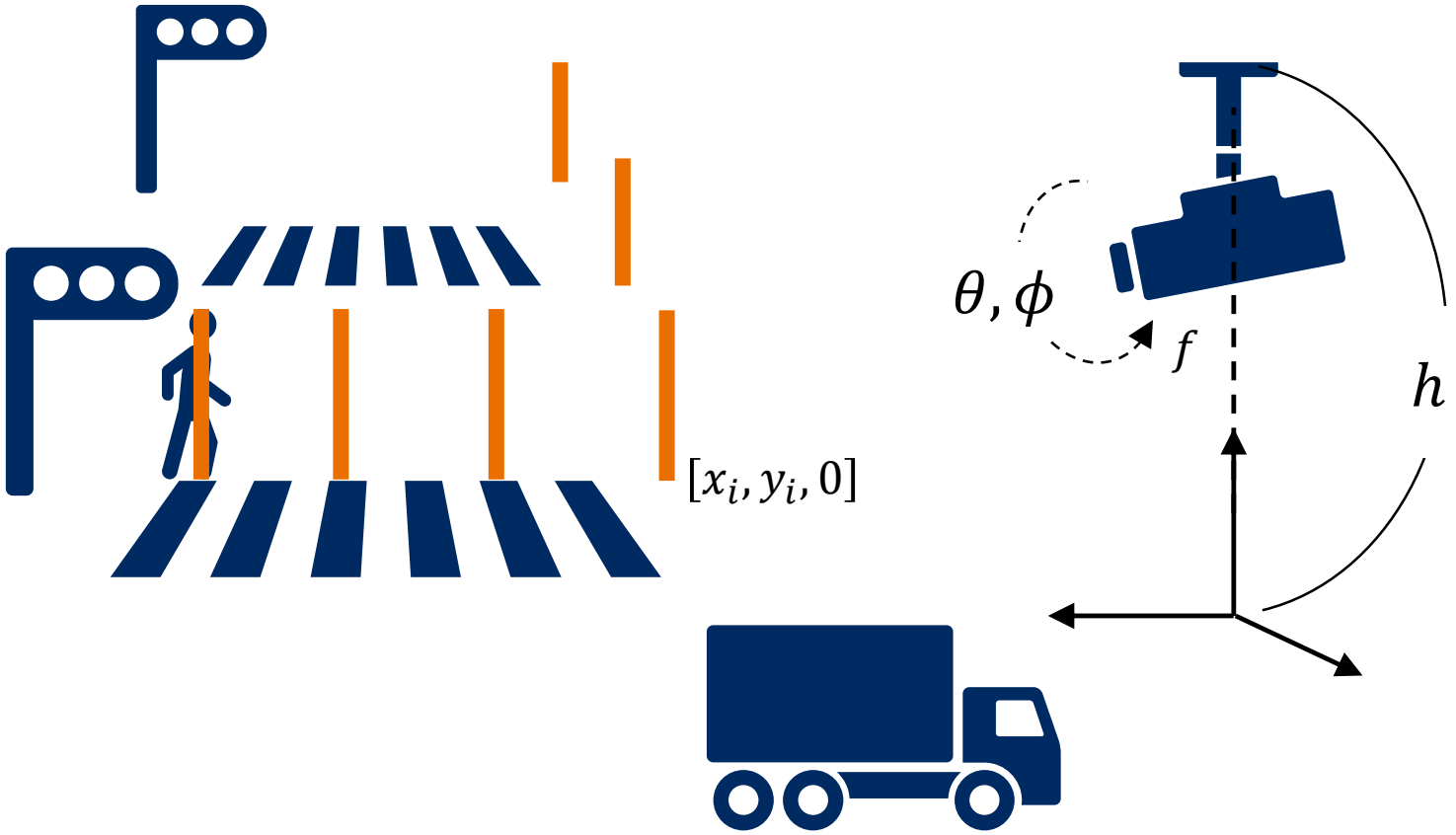
## Parallel line segments of the same length as calibration object

- Traffic light, Truck box, Pedestrian's height

# Novelty: Calibration from >=2 line segments

## Closed-form solution to

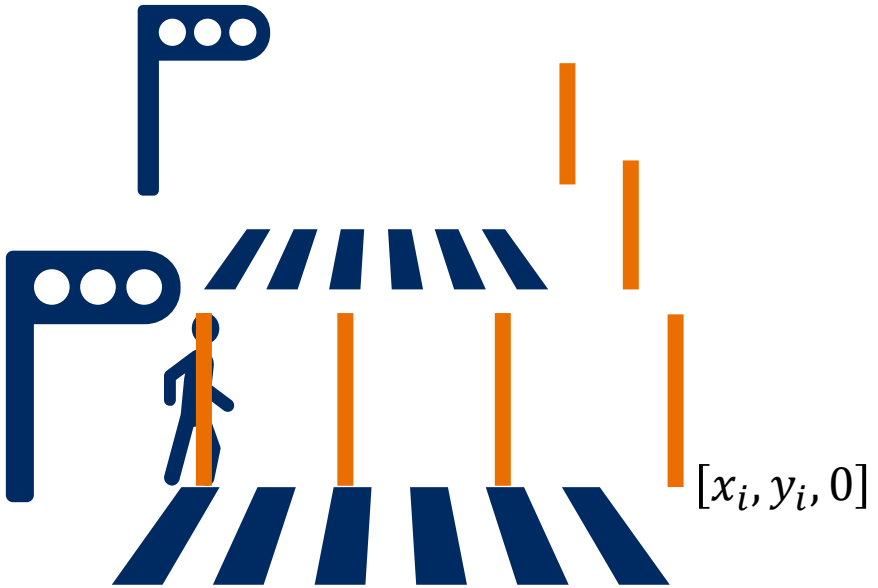
- Camera height, angles, focal length, 3D position of lines



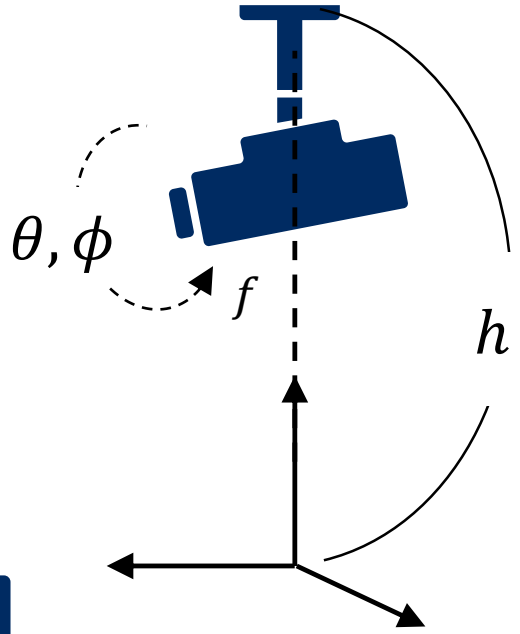
# Novelty: Calibration from >=2 line segments

## Non-linear optimization including lens distortion

- Refinement by minimizing reprojection error



$[x_i, y_i, 0]$



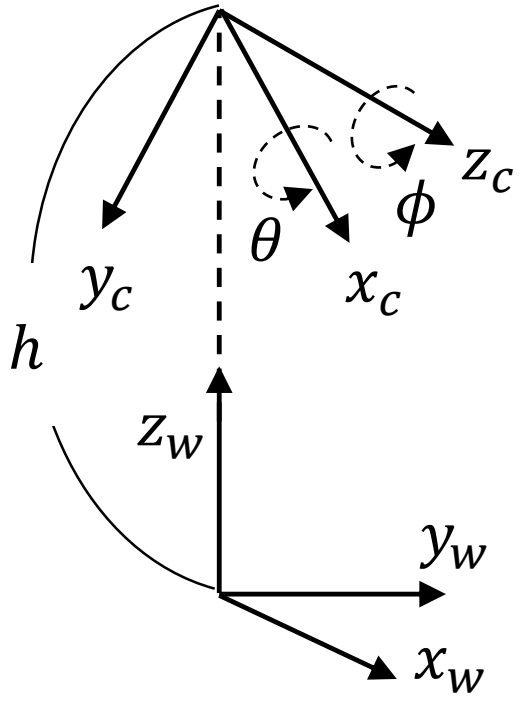

**Bundle Adjustment with lens distortion**

# Problem Formulation

# Problem Formulation: Coordinate system

## Camera

- Located at height $h$ on $z_w$-axis.
- $z_c$-axis is aligned with $y_w$-axis.



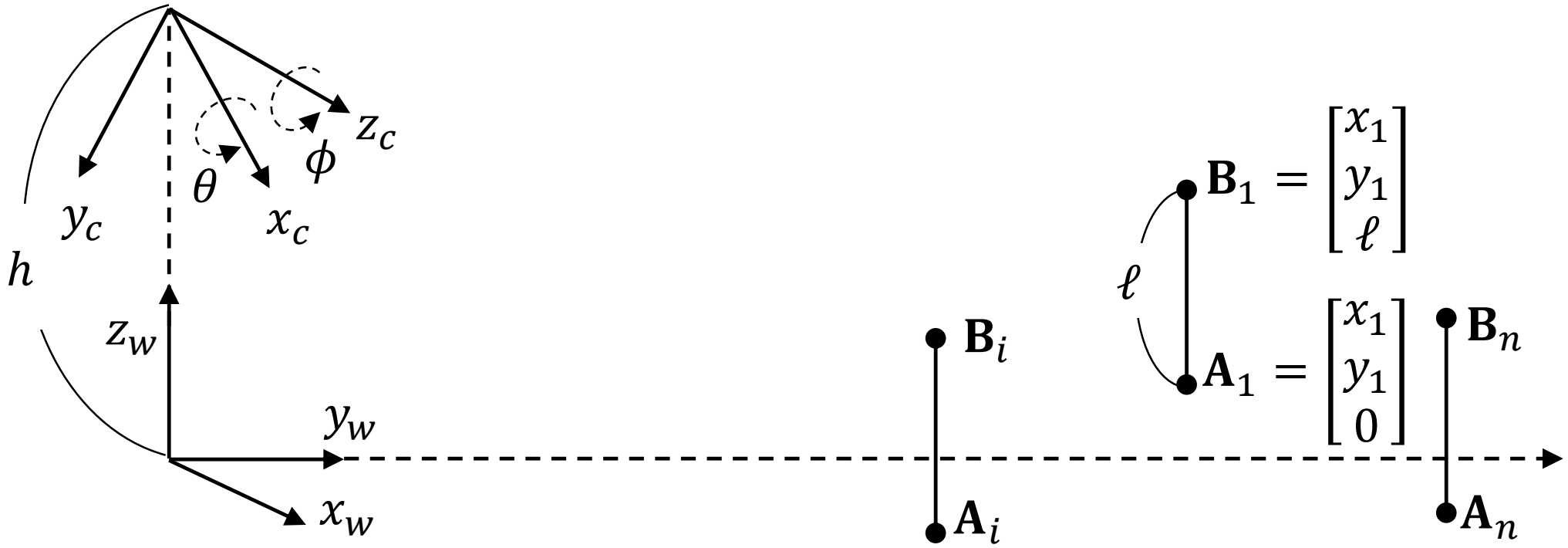
# Problem Formulation: Coordinate system

## Camera

- Located at height $h$ on $z_w$-axis.
- $z_c$-axis is aligned with $y_w$-axis.

## Line segments

- on the ground plane $z = 0$.
- have the *same unknown* length $\ell$.



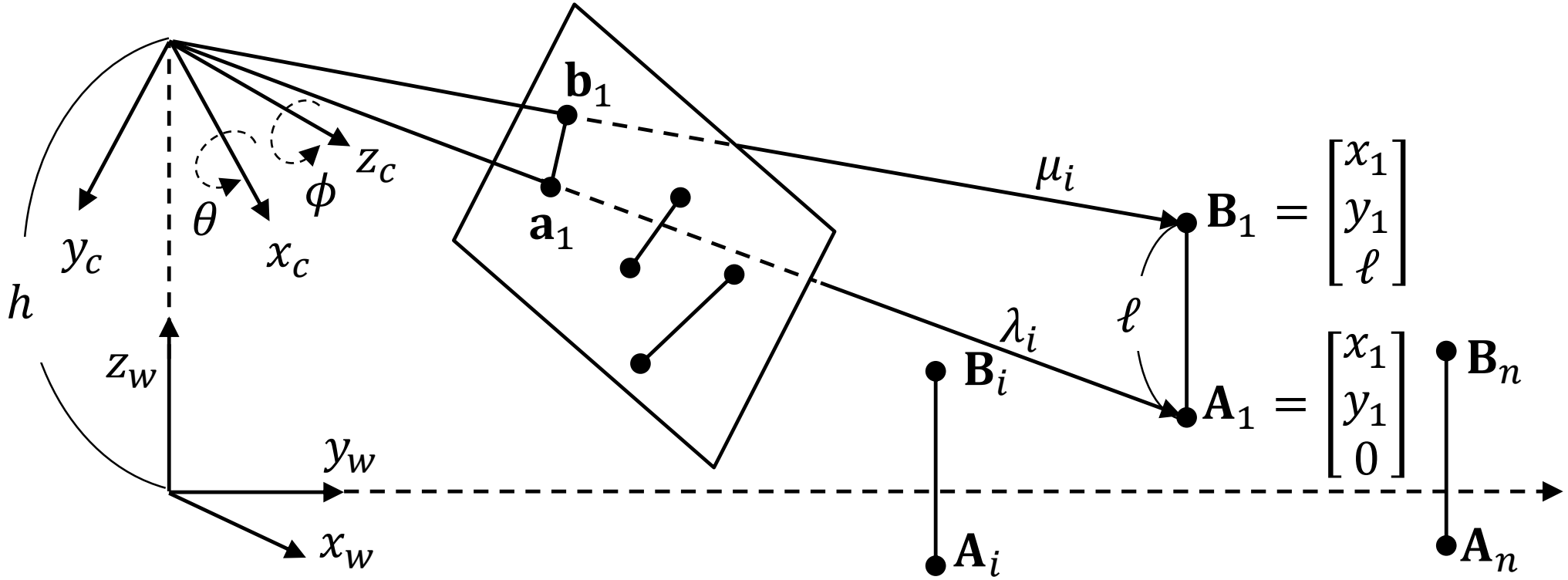
# Problem Formulation: Coordinate system

## Camera

- Located at height $h$ on $z_w$-axis.
- $z_c$-axis is aligned with $y_w$-axis.

## Line segments

- on the ground plane $z = 0$.
- have the *same unknown* length $\ell$.



# Problem Formulation: # of lines needed to solve

## Known variables

$\mathbf{a}_i, \mathbf{b}_i = [u_i, v_i, 1]^{\mathrm{T}}$

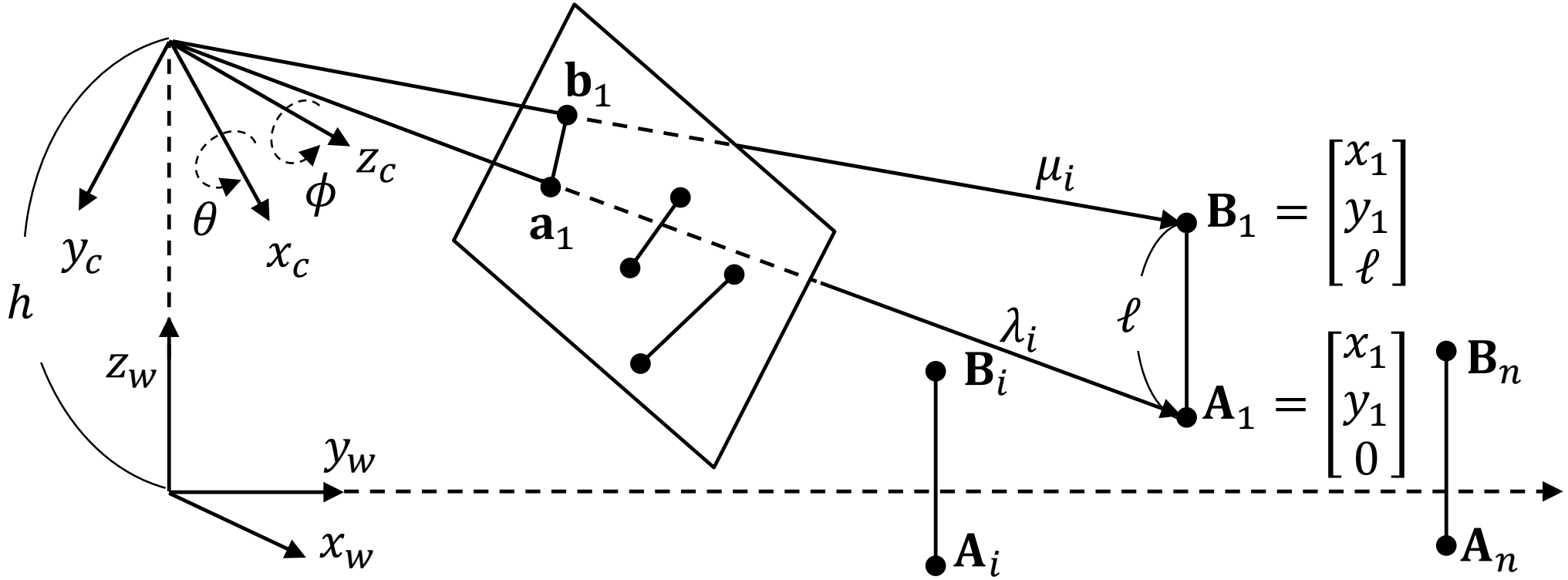
→ $6n$

## Unknown variables

$\theta, \phi, h, \ell, x_i, y_i, \lambda_i, \mu_i, f$

scale ambiguity

→ $4n + 4$



# Problem Formulation: # of lines needed to solve

## Known variables

$\mathbf{a}_i, \mathbf{b}_i = [u_i, v_i, 1]^{\mathrm{T}}$

→ $6n$

## Unknown variables

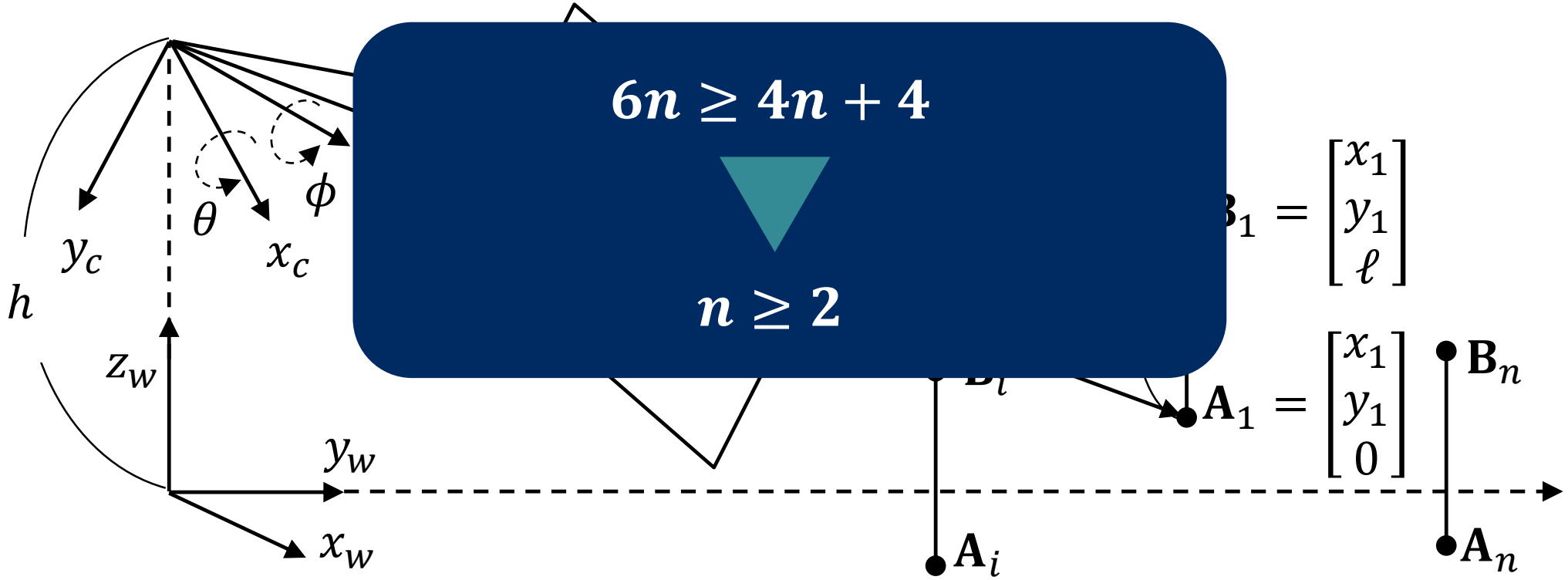
$\theta, \phi, h, \ell, x_i, y_i, \lambda_i, \mu_i, f$

scale ambiguity

→ $4n + 4$

$$6n \geq 4n + 4$$

$$n \geq 2$$



# Proposed Method

# Proposed Method: Overview

$$\mathbf{a}_i, \mathbf{b}_i \ (2 \leq i \leq n)$$

**Direct Linear Solution**

$$\theta, \phi, h, \ell, x_i, y_i, \lambda_i, \mu_i, f$$

**Bundle Adjustment with radial lens distortion $k$**

$$\theta, \phi, h, \ell, x_i, y_i, f, k$$

# Proposed Method: Direct Linear Solution

$\mathbf{a}_i, \mathbf{b}_i \ (2 \leq i \leq n)$

**Direct Linear Solution**

$\theta, \phi, h, \ell, x_i, y_i, \lambda_i, \mu_i, f$

**Bundle Adjustment with radial lens distortion $k$**

$\theta, \phi, h, \ell, x_i, y_i, f, k$

**Projective depth $\lambda_i, \mu_i$**

**Focal length $f$**

**Rotational angles $\theta, \phi$**

**Line length $\ell$**

**Line positions $x_i, y_i$**
**Camera height $h$**

**All steps are closed-form**

# Proposed Method: Non-linear Refinement

$\mathbf{a}_i, \mathbf{b}_i \ (2 \leq i \leq n)$

**Direct Linear Solution**

$\theta, \phi, h, \ell, x_i, y_i, \lambda_i, \mu_i, f$

**Bundle Adjustment with radial lens distortion $k$**

$\theta, \phi, h, \ell, x_i, y_i, f, k$

**Reprojection error minimization**

$$\min \sum d(\mathbf{a}_i, \mathbf{A}_i)^2 + d(\mathbf{b}_i, \mathbf{B}_i)^2$$

# Real Data Experiments

# V.S. checkerboard method on OpenCV

## Accurate line segments from checkerboard corners

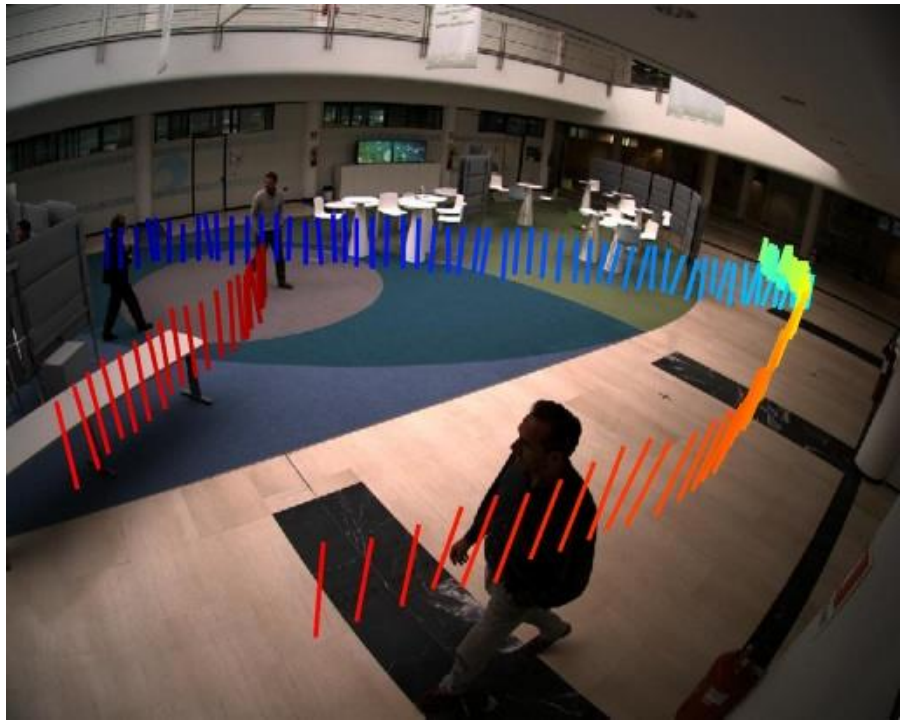
6x9 Checkerboard

**24 images**

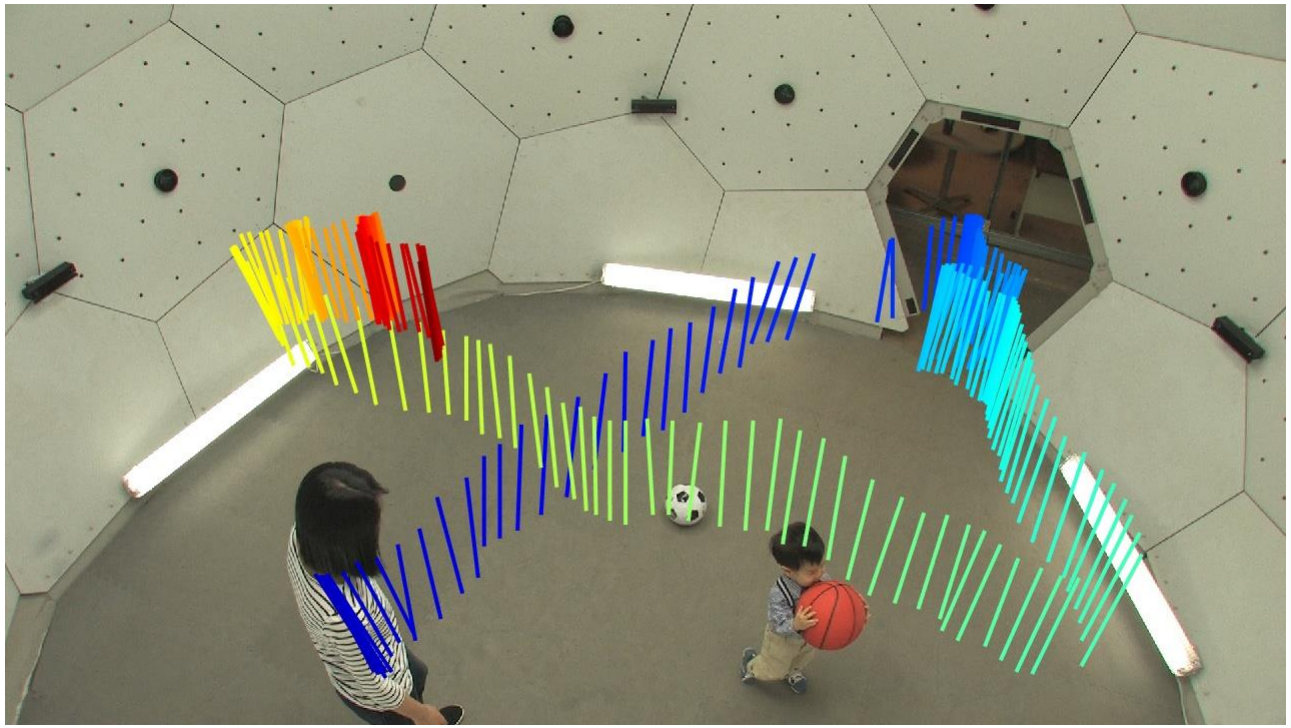
| | Focal length | Radial distortion |
|---|---|---|
| OpenCV (54 points/image) | 974.2 | -0.0485 |
| Proposed (9 lines/image) | 980.7 | -0.0470 |
| Relative error | 0.6% | 3.1% |

**The proposed method can estimate the camera parameters close to that of OpenCV.**

# Evaluation on Open Dataset



FBK Task Decomposition

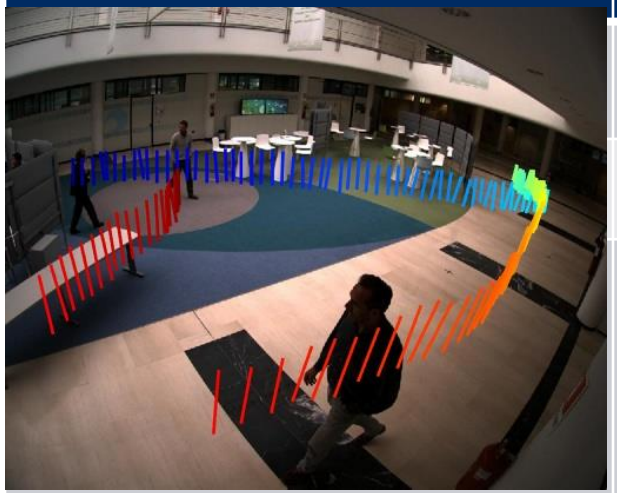


CMU Panoptic Studio

**Human pose are detected by OpenPose**

**Neck-midhip connections as line segments for calibration.**

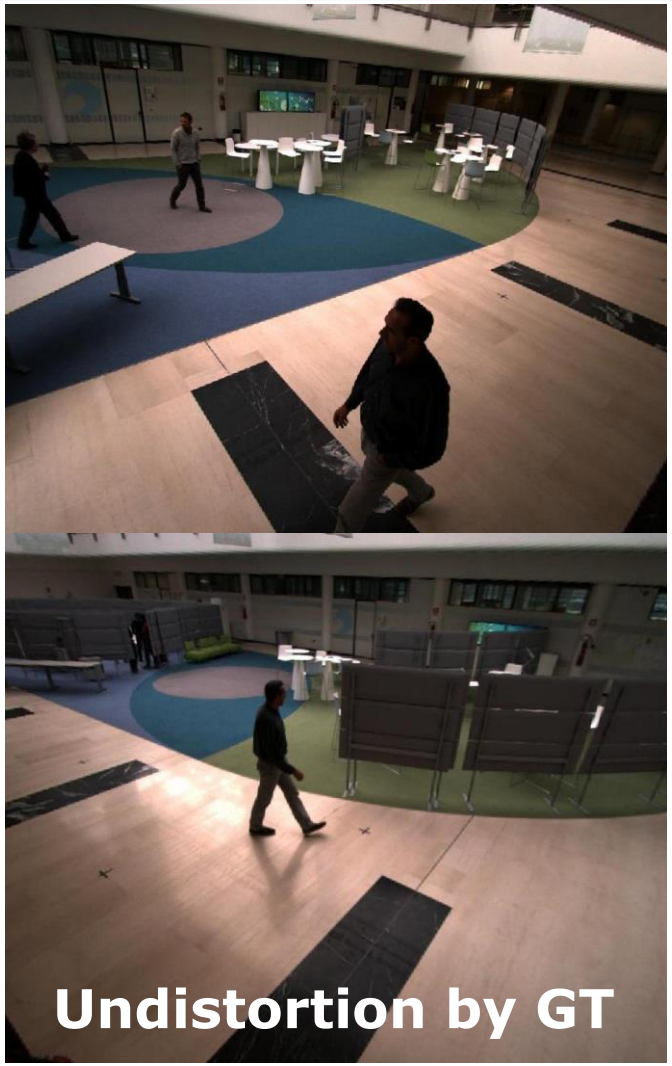
# Results: FBK Task Decomposition

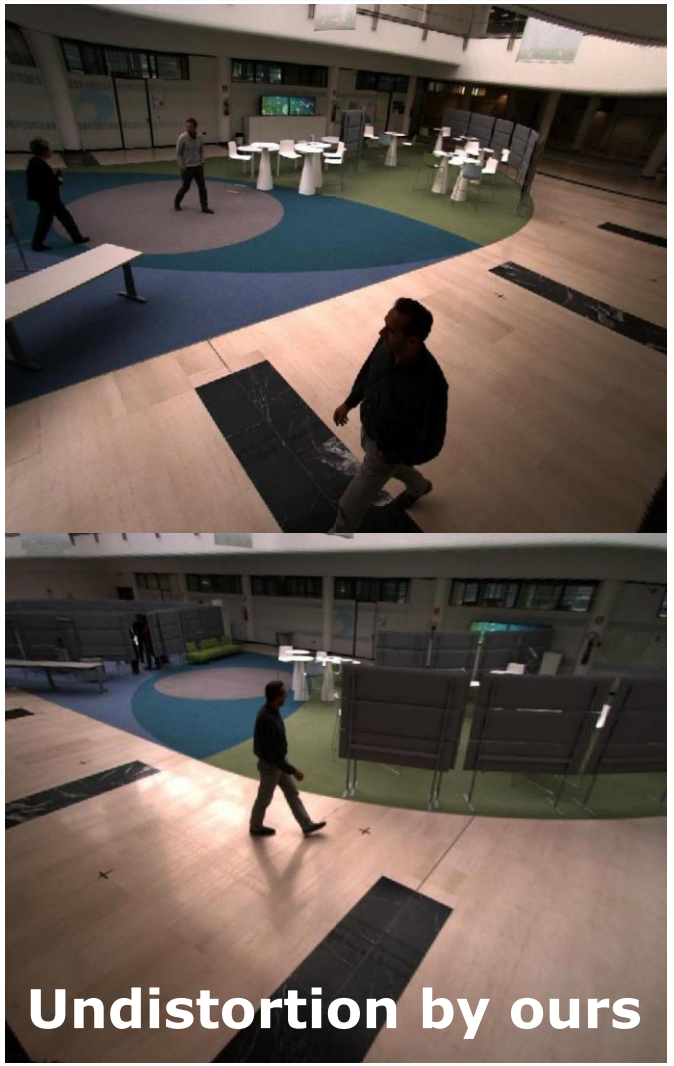


| | Focal length | Lens distortion | $\theta$ [deg] | $\phi$ [deg] | **Height** [m] |
|---|---|---|---|---|---|
| Ground Truth | 747 | -0.356 | 27.7 | 6.9 | 2.78 |
| Proposed | 715 | -0.217 | 28.2 | 1.7 | 2.73 |
| Relative error | 4.3% | 39% | 1.8% | 75% | 1.8% |
| Ground Truth | 731 | -0.361 | 30.4 | 1.4 | 2.78 |
| Proposed | 713 | -0.204 | 27.7 | 1.7 | 2.66 |
| Relative error | 2.5% | 43.5% | 8.9% | 21% | 4.3% |

**The camera parameters are accurately estimated only from human joints.**

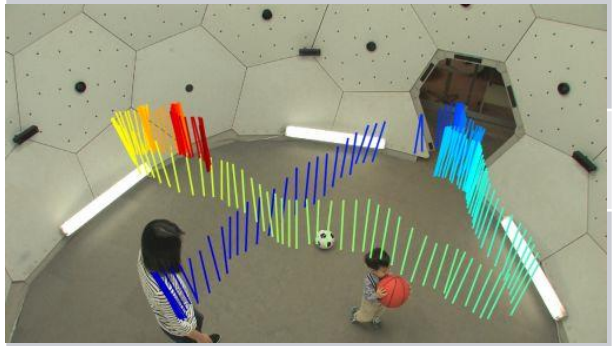
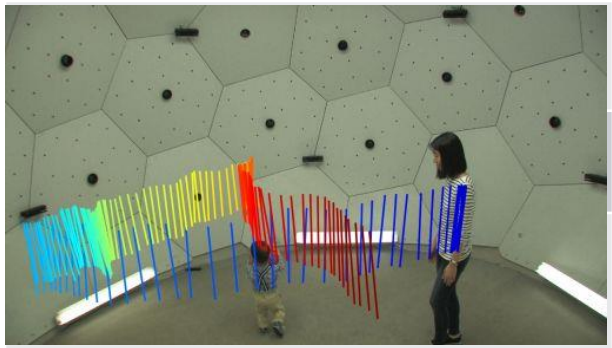
# Results: FBK Task Decomposition


**Lens distortion seems to be bad, however, the visual quality of undistortion is almost comparable.**

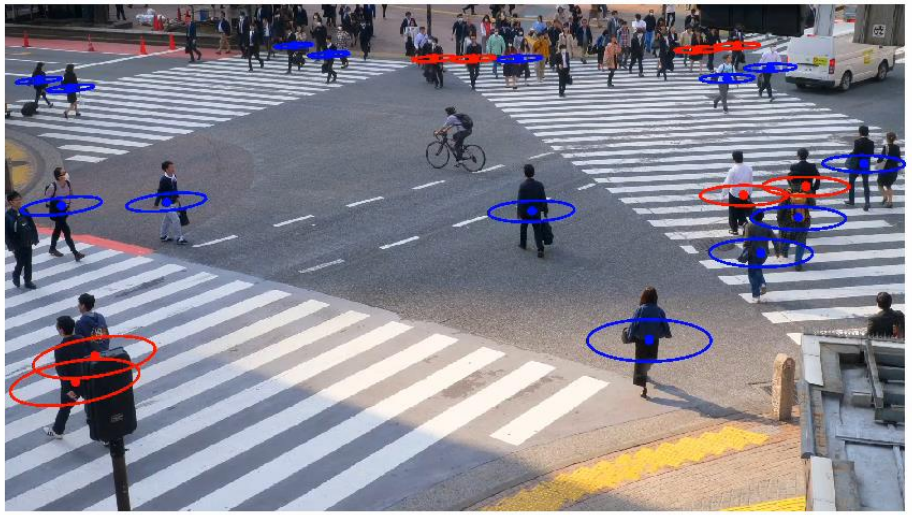
# Results: CMU Panoptic Studio

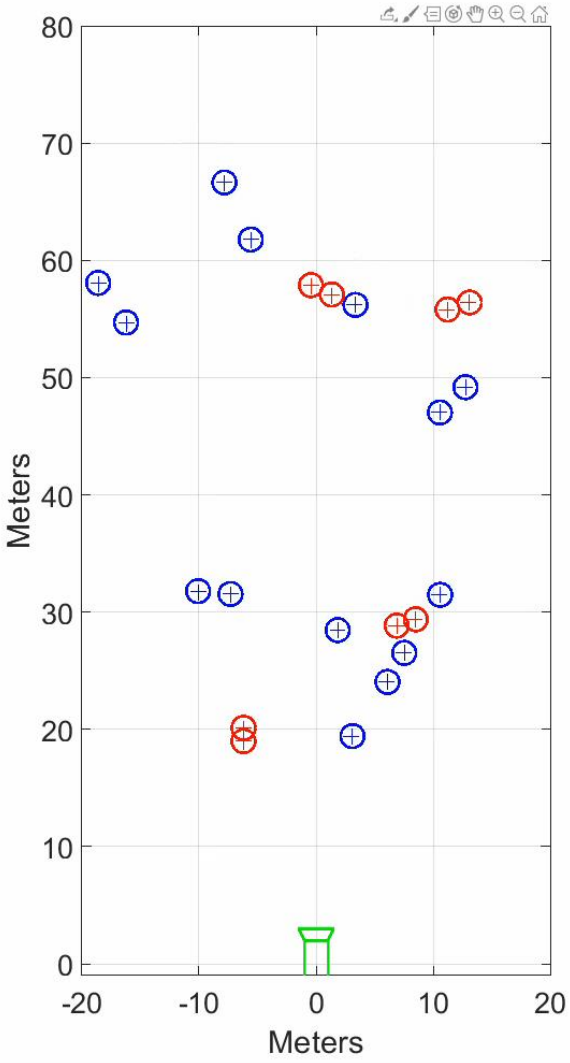
| | | **Focal length** | **Lens distortion** | $\theta$ [deg] | $\phi$ [deg] | **Height** [m] |
|---|---|---|---|---|---|---|
|  | Ground Truth | 1396 | -0.286 | 40.0 | 2.2 | 3.12 |
| | Proposed | 1386 | -0.203 | 41.7 | -0.7 | 3.15 |
| | Relative error | 0.7% | 29% | 4.3% | 131% | 1.0% |
|  | Ground Truth | 1407 | -0.283 | 12.7 | -3.6 | 1.97 |
| | Proposed | 1330 | -0.246 | 11.9 | -3.1 | 1.94 |
| | Relative error | 5.5% | 13% | 6.3% | 14% | 1.5% |

**Surveillance cameras can be calibrated by observing a pedestrian.**

# Applications

# Social Distancing Detection






Too close

Good distance

**Using the proposed method, we can calibrate surveillance cameras which were already installed in town to measure physical distance between pedestrians.**

# Planar Image Rectification


**Frontal images can be recovered by specifying two edges of a rectangle.**

# Conclusions

# Conclusions

- Proposed a camera calibration method using line segments of the same length
  - Can be solved by a closed-form.
  - Radial distortion estimation by bundle adjustment.

- Experimental results demonstrate
  - Human joints (neck-midhip connection) are available for calibration including lens distortion.
  - Calibration can be conducted without a large pattern nor traffic restrictions.

# Orchestrating a brighter world

NEC creates the social values of
safety, security, fairness and efficiency
to promote a more sustainable world
where everyone has the chance to reach
their full potential.