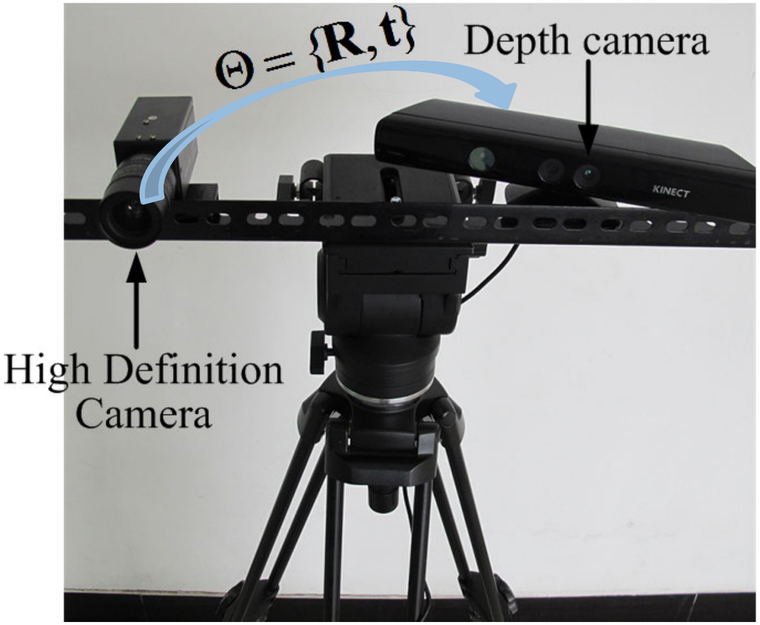
Single-shot Extrinsic Calibration of Generically Configured RGB-D Camera Rig from Scene Constraints

Jiaolong Yang*, Yuchao Dai^,
Hongdong Li^, Henry Gardner^ and Yunde Jia*

*Beijing Institute of Technology (BIT),
^The Australian National University (ANU)

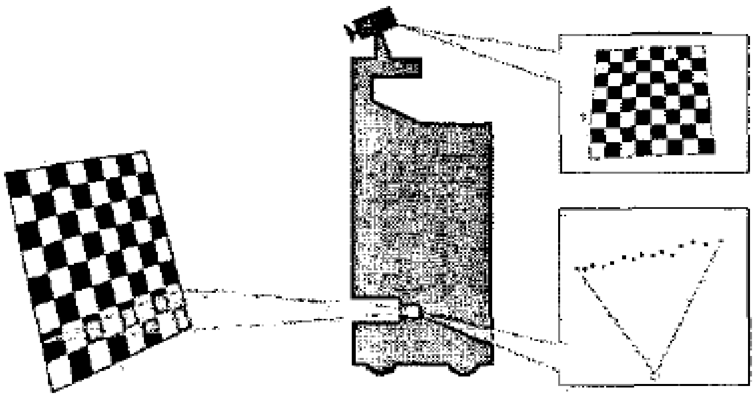
Introduction

- Calibration parameters are required to fuse the color and depth images for AR/MR Application
- Intrinsic parameters are often fixed while extrinsic parameters can be subject to change, see *e.g.*

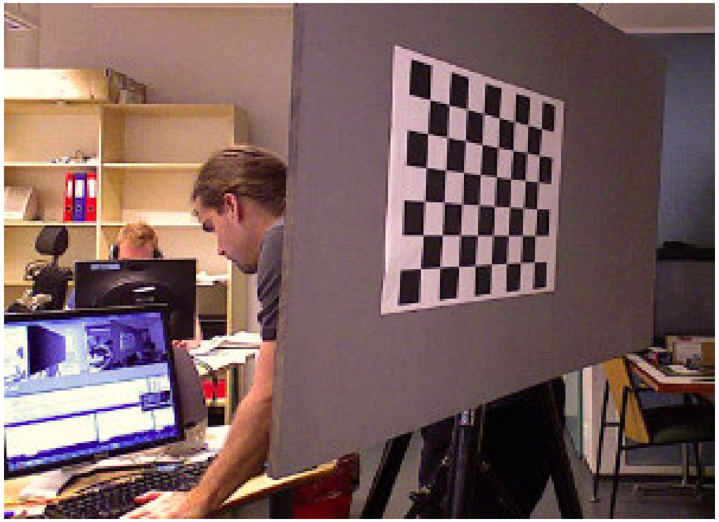


Related Work

- Calibration with checkerboard pattern



Q. Zhang and R. Pless *IROS*'04

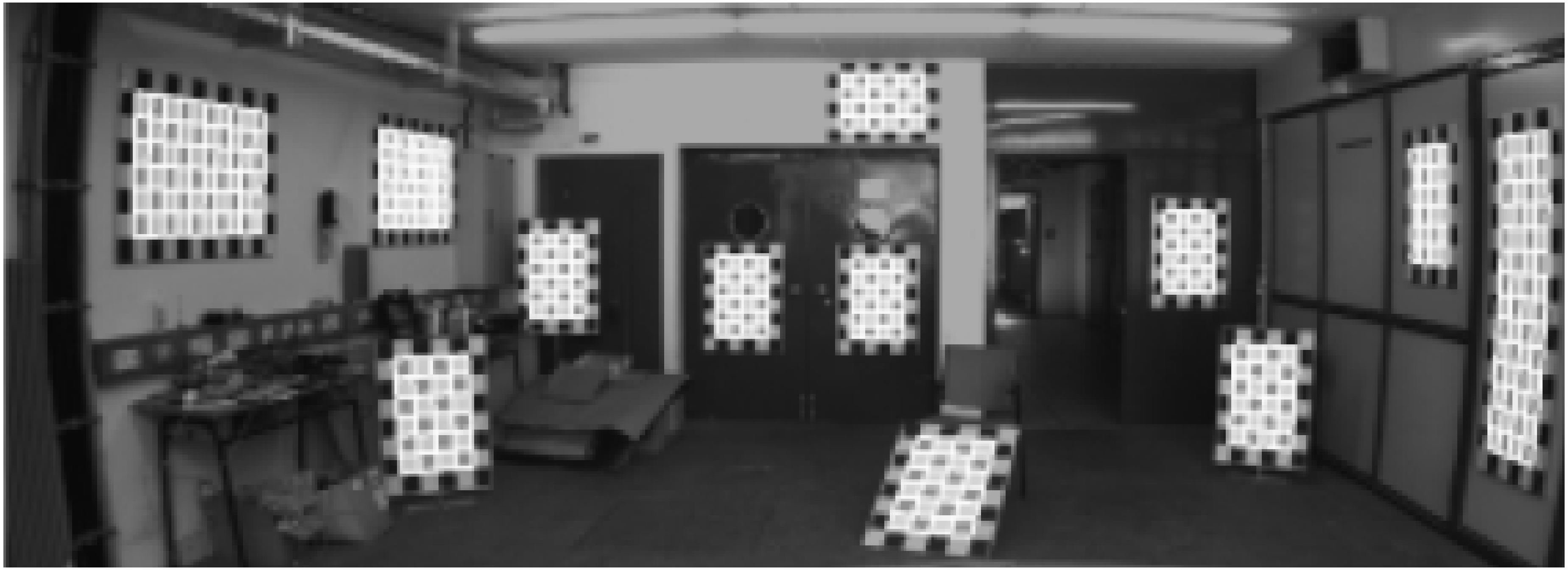


D. Herrera *et al*. *PAMI*'12

C. Zhang and Z. Zhang. *ICME*'11

Related Work

- Calibration with checkerboard pattern (single shot)



A. Geiger *et al*. *ICRA'12*

Related Work

- Hand-eye calibration
  - R. Y. Tsai and R. K. Lenz. *T-RA*’89
  - R. Horaud and F. Dornaika. *IJRR*’95
  - N Andreff, R Horaud, B Espiau. *IJRR*’01
  - Y. Dai *et al*. *ACCV*’09
  - …

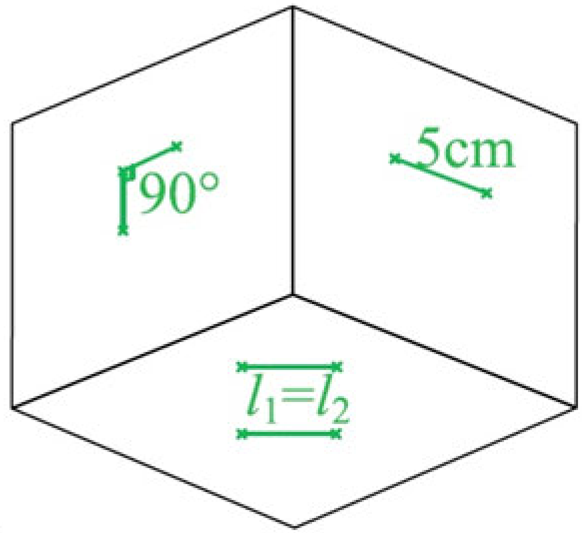
Single-shot Extrinsic Calibration

- Our target:
  - Single-shot verus multi-shot
  - Scene constraints verus calibration patterns
- The principles:
  - Color image and depth map measure the scene in different modalities from different positions
  - Scene constraints are invariant to view and modalities
  - Single-shot with scene constraints provide enough information for extrinsic calibration

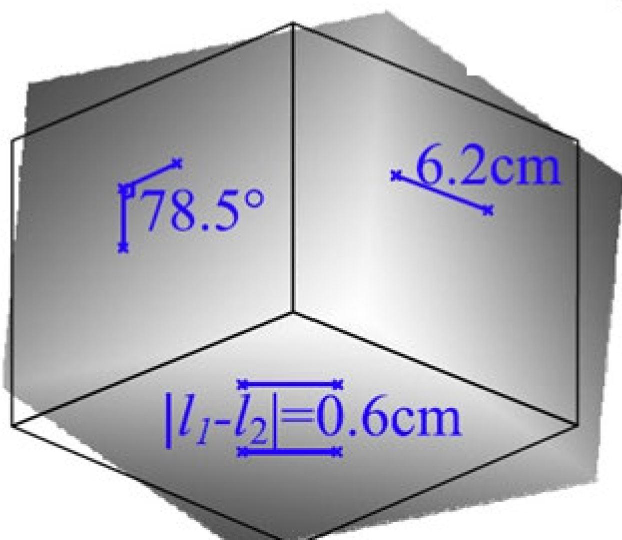
Single-shot Extrinsic Calibration

- Scene constraints:
  - Distance: known distance, distance equivalency
  - Angle: parallel, orthogonal, known angle
- Evaluation:
  - Inverse projection
  - Scene knowledge measurement
  - Scene knowledge discrepancy
- Benefits:
  - Hand held camera application
  - Post processing/ permit varying extrinsic parameters

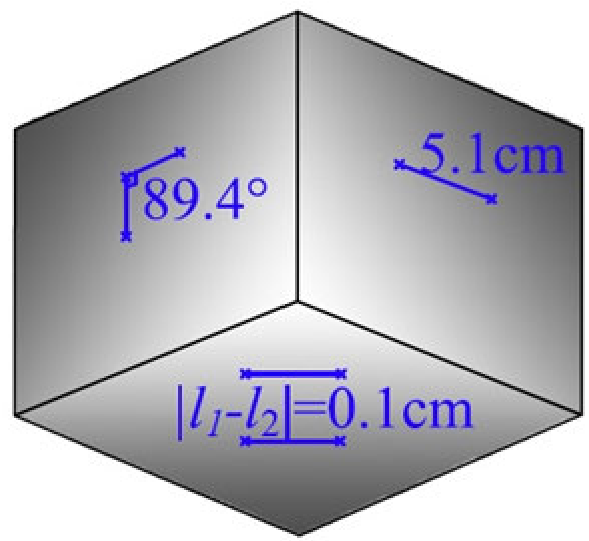
Scene Constraints Evaluation



Scene Constraints
(Ground Truth)



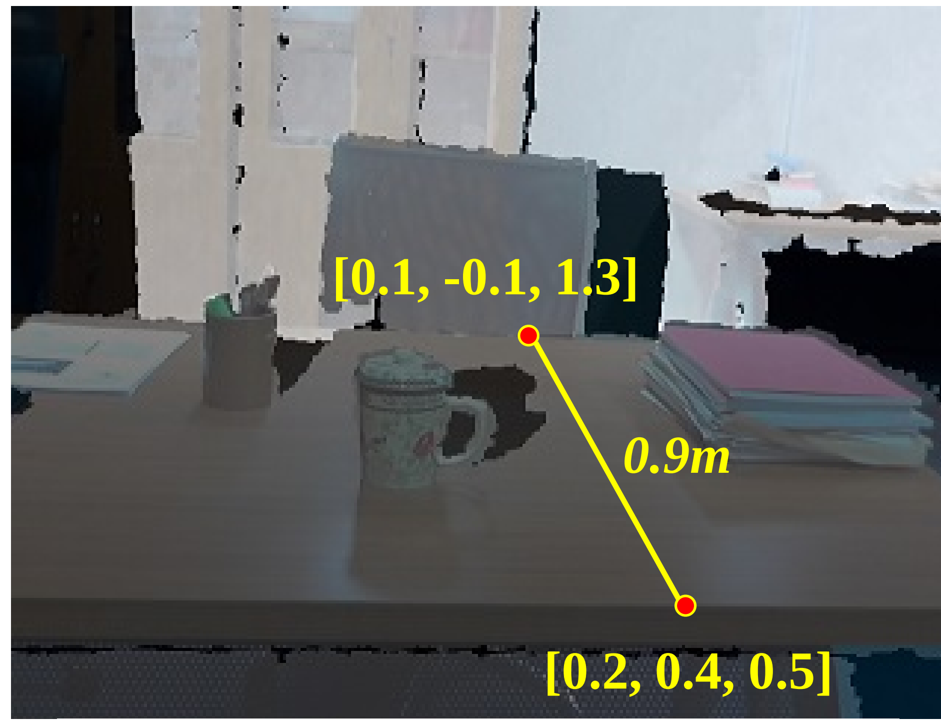
Measurements with
Wrong Transformation



Measurements with
Optimal Transformation

Single-shot Extrinsic Calibration

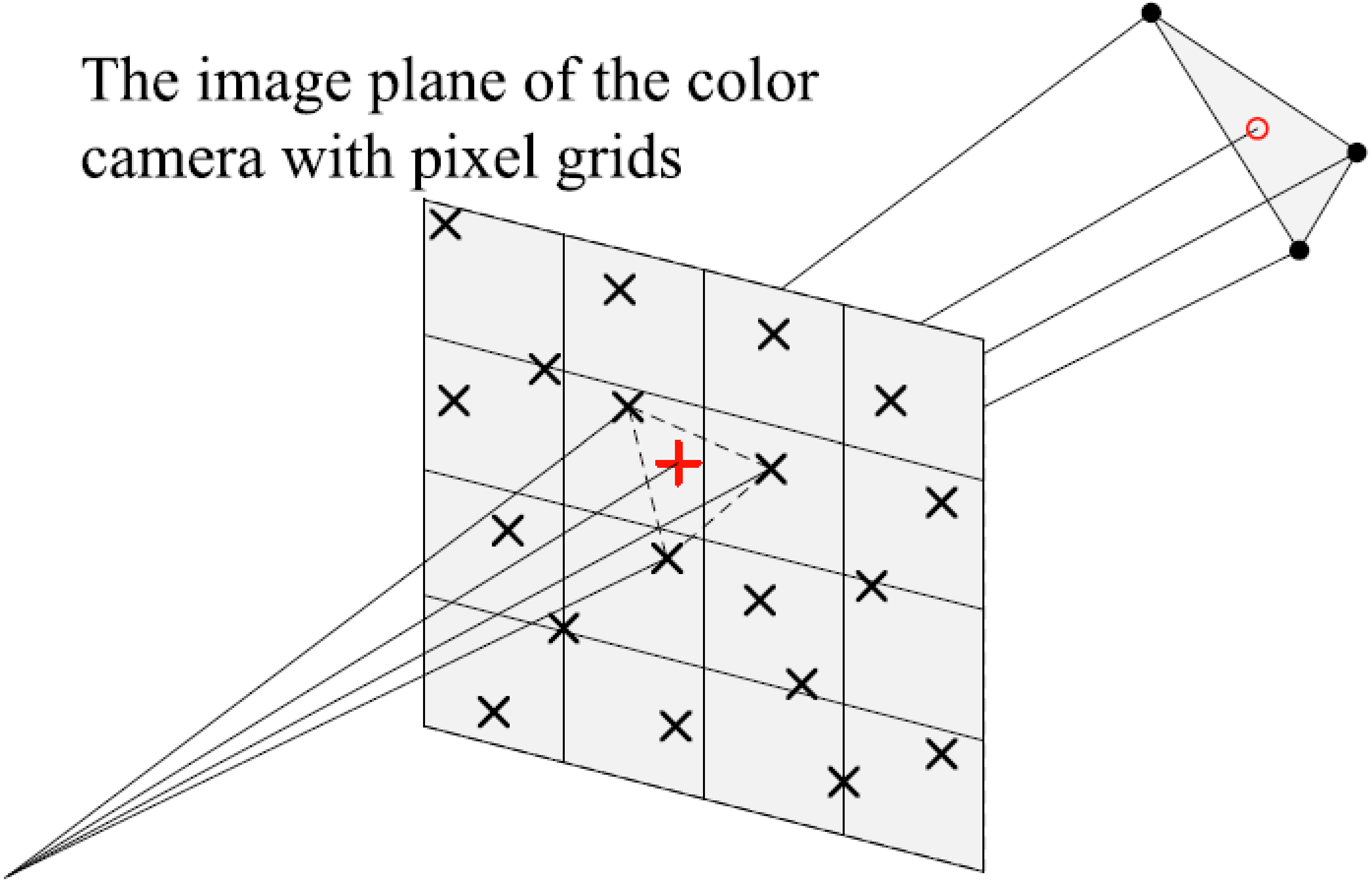
$$(u_i^{cd}, v_i^{cd}) = g(\Theta) \circ (x_i^d, y_i^d, z_i^d)$$



$$(x_i^d, y_i^d, z_i^d) = g^{-1}(\Theta) \circ (u_i^c, v_i^c)$$

Single-shot Extrinsic Calibration

- Inverse projection estimation by triangulation



Single-shot Extrinsic Calibration

Given a Θ we can compute the ***discrepancy/error*** between a scene constraint and its measurement.

- Known distance constraint:

$$e_k(\Theta) = |||g^{-1}(\Theta) \circ (u_i^c, v_i^c) - g^{-1}(\Theta) \circ (u_j^c, v_j^c)|| - l_{ij}|$$

- Distance equivalency constraint:

$$e_d(\Theta) = |||g^{-1}(\Theta) \circ (u_i^c, v_i^c) - g^{-1}(\Theta) \circ (u_j^c, v_j^c)|| - ||g^{-1}(\Theta) \circ (u_k^c, v_k^c) - g^{-1}(\Theta) \circ (u_l^c, v_l^c)|||$$

- Angular constraints:

....

Single-shot Extrinsic Calibration

Our goal is to find optimal transformation minimizing the total geometric error:

$$\Theta^* = \underset{\Theta \in SE(3)}{\operatorname{argmin}} \sum_i e_i(\Theta)^2$$

Minimization:

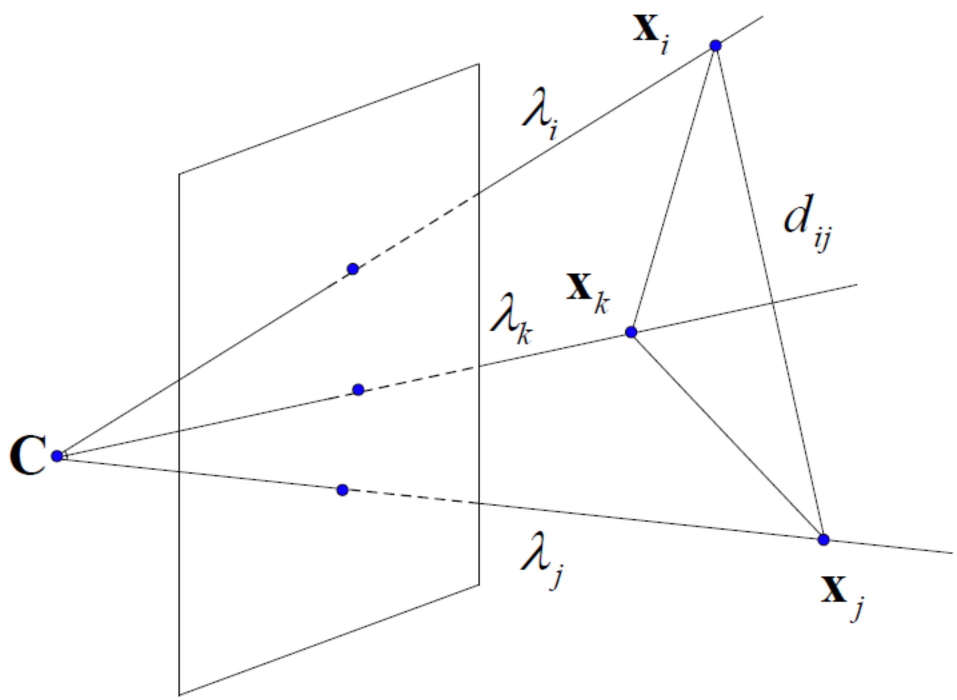
- Levenberg-Marquardt algorithm (numerical gradients)
- Nelder-Mead simplex downhill on *SE*(3) manifold (gradient free)

Initialization

- Build pointset in the color camera by **Single View Reconstruction (SVR) with Scene Constraints**
- Find transformation between color and depth point clouds by **Pointset Registration**

Initialization

- **Single View Reconstruction (SVR)**



The corresponding 3D point of (u_i^c, v_i^c) lies on the ray with direction $\mathtt{K}^{-1}[u_i^c \;\; v_i^c \;\; 1]^T$ with unknown projective depth λ_i to determine.

Initialization

- **Single View Reconstruction (SVR)**

  The known distance (*see paper for more constraints and minimal configuration*) between two 3D points gives constraint on the projective depth

$$\min \operatorname{trace}(\mathtt{Y}) \\ \text{such that } \operatorname{tr}(\mathtt{A}_{ij}\mathtt{Y}) = d_{ij}^2, \forall (i,j) \in \mathcal{N}.$$

Semi-Definite Programming (SDP) problem

Initialization

- **Pointset registration**

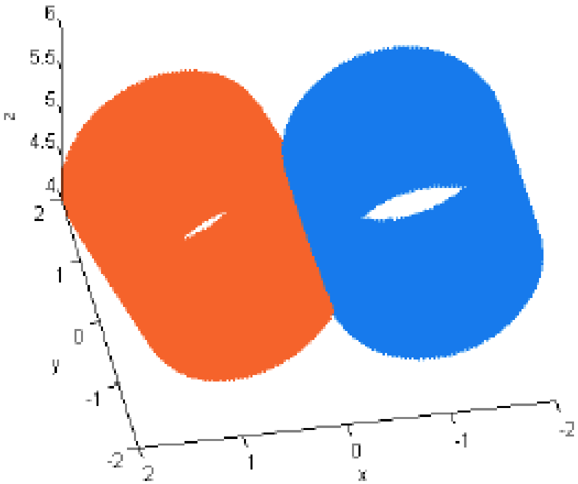
  Two point clouds: $\{\mathrm{X}_i^c\}$ $\{\mathrm{X}_j^d\}$

  – ICP: $\Theta_0 = \underset{\Theta \in \mathrm{SE}(3)}{\operatorname{argmin}} \sum_i \min_j \|(\mathrm{R}\mathrm{X}_i^c + \mathbf{t}) - \mathrm{X}_j^d\|^2.$

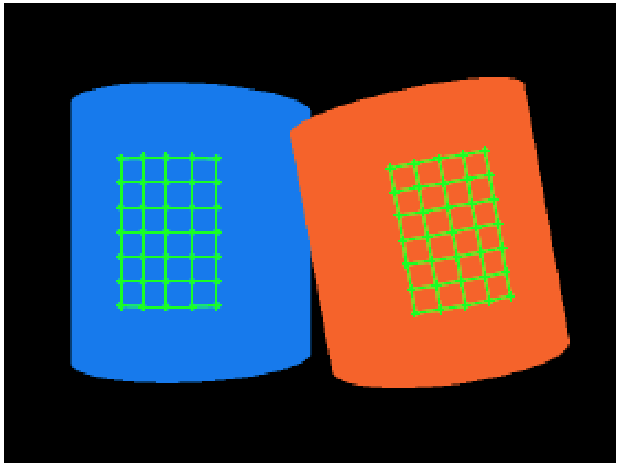
  – Go-ICP(Globally-optimal ICP) (*J. Yang, H. Li, Y. Jia, to appear in ICCV'13*)

Experiments

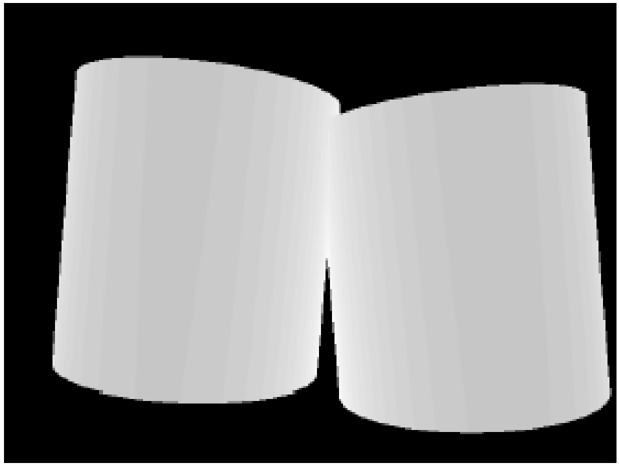
- Synthetic scene: two cylinders



(a) A synthesized scene.



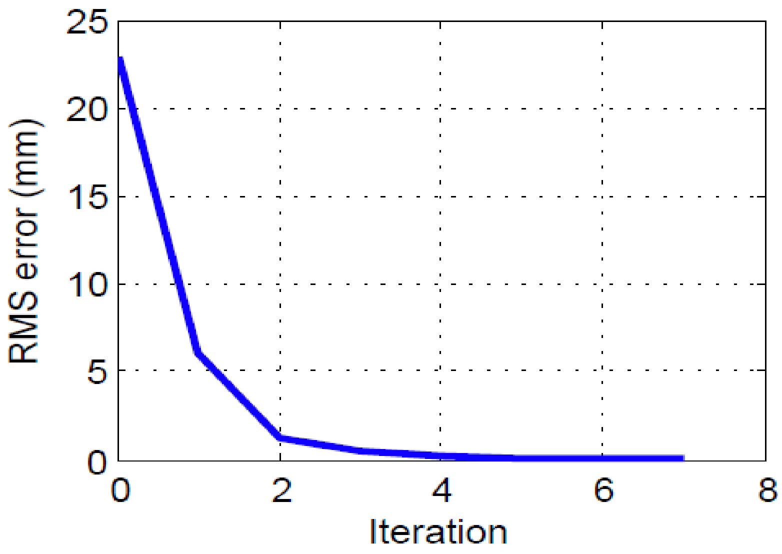
(b) Synthesized color image



(c) Synthesized depth image

Experiments

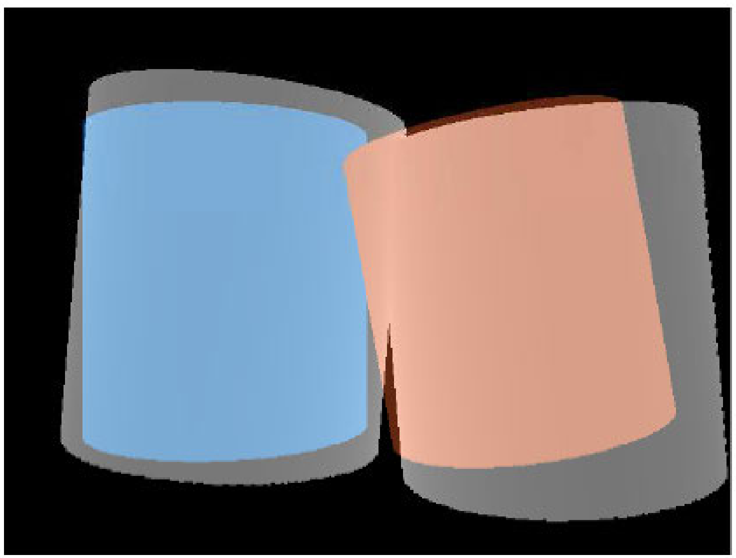
- Synthetic data: two cylinders
  - Quantitative results



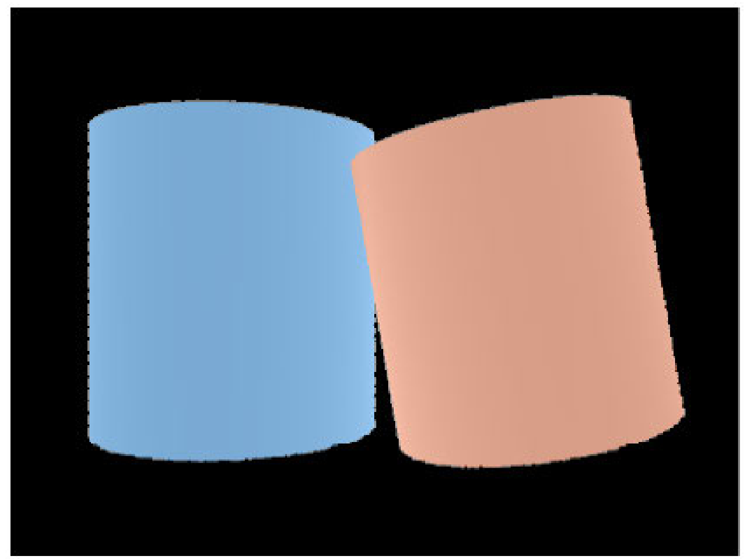
	Angle (°)	Axis	Translation (m)
Ground truth	5.067	-0.100,-0.128,-0.987	-0.113,-0.086,0.500
Our method	5.106	-0.096,-0.128,-0.986	-0.112,-0.078,0.503

Experiments

- Synthetic data: two cylinders
  - Qualitative results



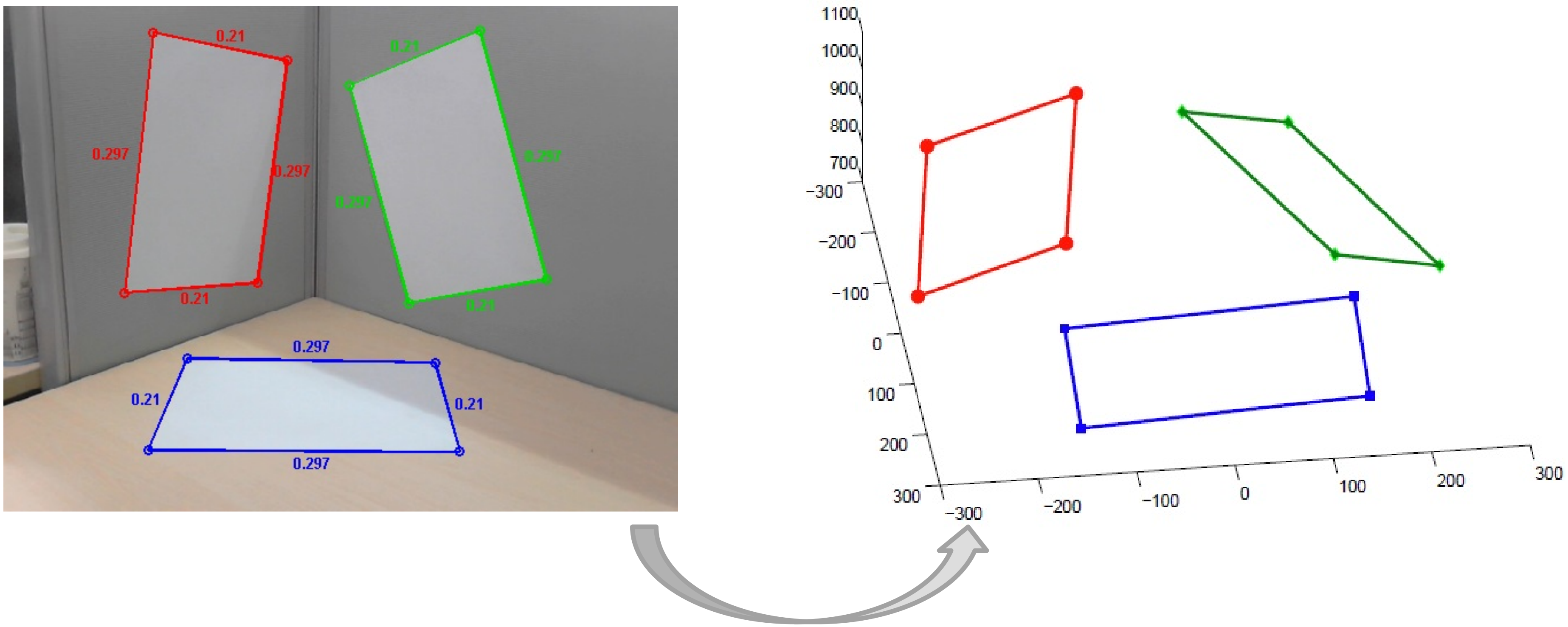
(d) Initial Alignment



(e) Alignment result

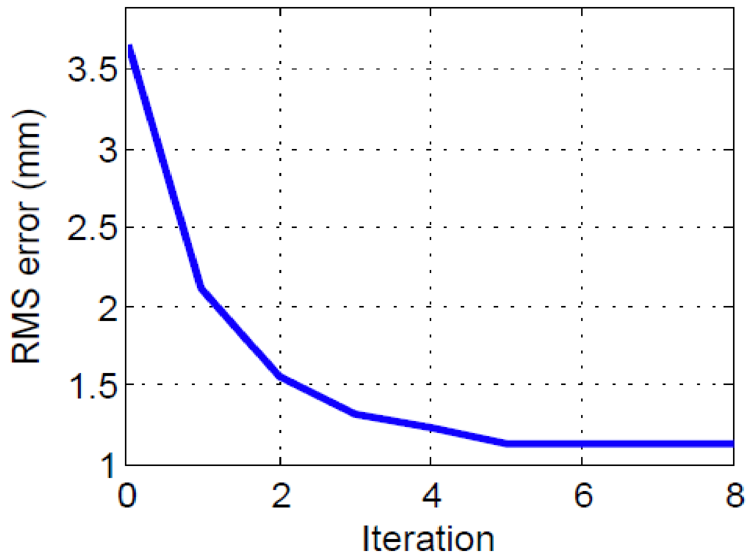
Experiments

- Real-world scene: three A4 paper



Experiments

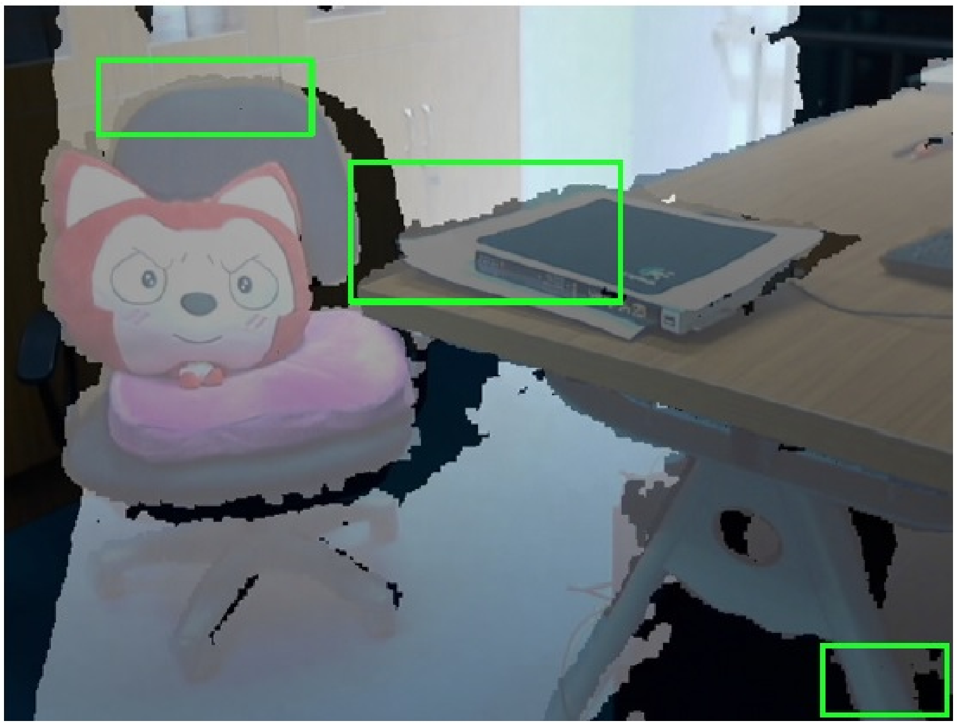
- Real-world scene: three A4 paper
  - Quantitative results

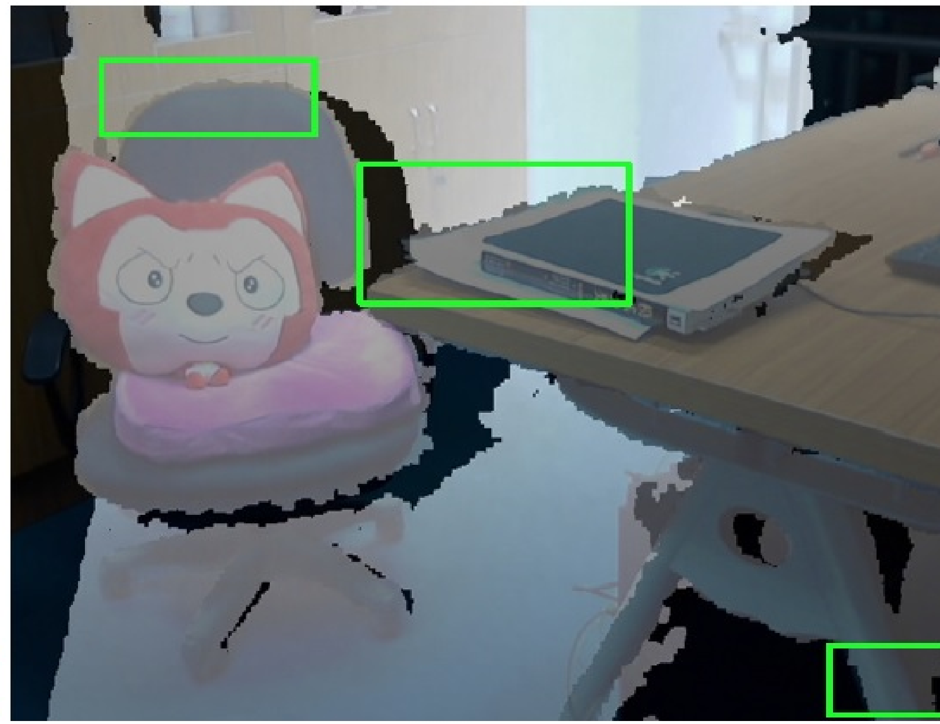


	Angle (°)	Axis	Translation (m)
Herrera *et al.*[9]	17.225	0.102 -0.986 0.131	0.280 0.046 0.083
Our method	17.619	0.104 -0.983 0.153	0.273 0.043 0.091

Experiments

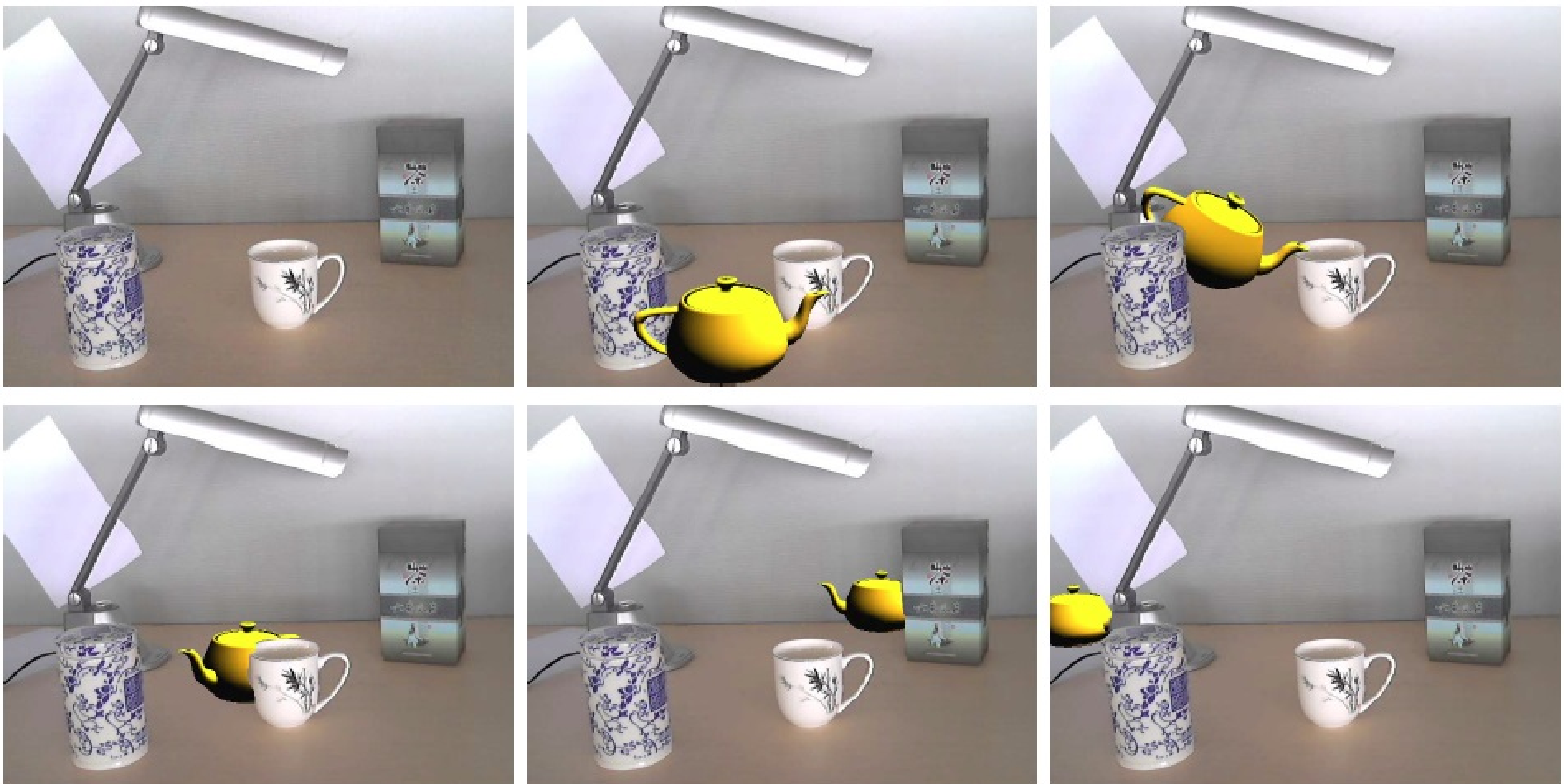
- Real-world scene: three A4 paper
  - Qualitative results





Experiments

- Real-world scene: three A4 paper
  - AR Application



Summary

- Single-shot extrinsic calibration of generally configured RGB-D camera rig from scene constraints
- Single view reconstruction + pointset registration for initialization
- Without specifically designed patterns
- Correspondence-free
- Geometric error minimization
- Less human intervention

Thanks