Automatic Calibration of a Multi-Camera System

with Limited Overlapping Fields of View for 3D Surgical Scene Reconstruction

Tim Flückiger[1,2], Jonas Hein[1,2], Valery Fischer[1], Philipp Fürnstahl[1], Lilian Calvet[1]

[1]Research in Orthopedic Computer Science, University Hospital Balgrist, University of Zurich, Switzerland
[2]Computer Vision and Geometry, ETH Zurich, Switzerland

1 Introduction

- Known camera extrinsics are required for many state-of-the-art methods, but their calibration remains a manual process [1]
- Calibration quality is **operator-dependent** (coverage, motion blur, reflections) and **time-consuming**.
- A single fixed-size marker may not be detected in all cameras due to **large scale changes**.


2 Method Overview

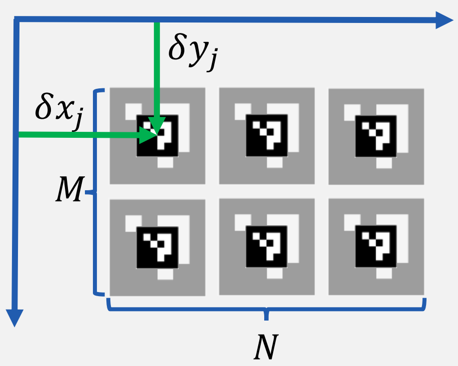
We propose a **fully automated calibration** solution for indoors based on multi-scale markers and a ceiling-mounted projector.

Multi-Scale Markers

- A set of markers with a common center point that is invariant under homography. In practice, most common markers are suitable, e.g. ArUco [2], CCTags [3].

Automatic Projection of Marker Arrays

- We generate a sequence of frames with an MxN grid of markers at $s_i \in S$ scales and spatial offsets $(\delta x_j, \delta y_j) \in \Delta$.
- The frames are projected onto the OR floor using an entry-grade projector.



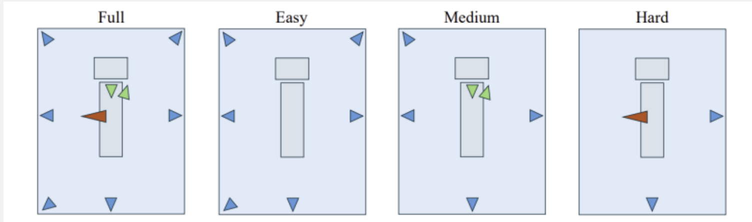
Optimization of Extrinsic Camera Parameters

- Camera extrinsics are recovered via incremental Bundle Adjustment, similar to standard SfM pipelines [4]
- During initialization, a degenerate essential matrix estimation from co-planar 3D points must be circumvented.

→ Instead, estimate the pose $\mathbf{P}_{c_2}$ of camera c_2 relative to camera c_1 by decomposing the inter-image homography $H_{c_2}^{c_1}$

3 Results

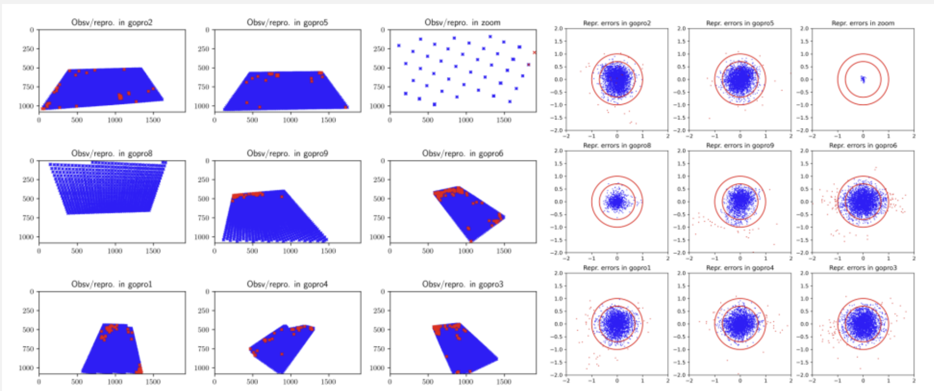
We test our method in a mock-up OR with four camera configurations. We use a calibration sequence of **70 seconds**, totalling 22 400 markers.



Our fully automated calibration approach achieves **mean reprojection errors** of less than 0.35px in our experiments, on par with a manual calibration using a ChArUco board.

	Full		Easy		Medium		Hard	
	ChAr.	Ours	ChAr.	Ours	ChAr.	Ours	ChAr.	Ours
Calibration	0.33	0.35	0.30	0.34	0.35	0.32	0.30	0.23
Evaluation	0.18	0.28	0.17	0.27	0.21	0.27	0.16	0.16

Spatial distribution of markers and reprojection errors:



Our approach yields higher **success rates** (for 0.5 / 2 / 5 pixel) compared to marker-free and manual calibration methods:

	COLMAP		GLOMAP		ChArUco	Proposed
	Standard	Projected Texture	Standard	Projected Texture		
Full	0/0/0	0/33/89	0/0/22	0/22/89	89/89/89	100/100/100
Easy	0/0/0	0/50/100	0/0/50	0/0/83	100/100/100	100/100/100
Medium	0/0/0	0/33/100	0/0/0	0/33/83	100/100/100	100/100/100
Hard	0/0/0	0/0/0	0/0/0	0/0/0	67/67/67	100/100/100

4 Conclusion

We propose a **fully automated** method for the calibration of camera extrinsic parameters in indoor environments, achieving **state-of-the-art accuracy in 70 seconds without user input**.

References

[1] Zhang, Z. (2002). A flexible new technique for camera calibration. IEEE Transactions on pattern analysis and machine intelligence, 22(11), 1330-1334.

[2] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recognition, 47(6), 2280-2292.

[3] Calvet, L., Gurdjos, P., & Charvillat, V. (2012, September). Camera tracking using concentric circle markers: Paradigms and algorithms. In 2012 19th IEEE International Conference on Image Processing (pp. 1361-1364). IEEE.

[4] Schonberger, J. L., & Frahm, J. M. (2016). Structure-from-motion revisited. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4104-4113).