Fast Multiview Registration of 3D Scans using Planar Structures

Uttaran Bhattacharya, Sumit Veerawal, Venu Madhav Govindu

{uttaran.bhattacharya127@gmail.com | sumit.veeru@gmail.com | venu@ee.iisc.ernet.in}

Indian Institute of Science, Bengaluru, INDIA

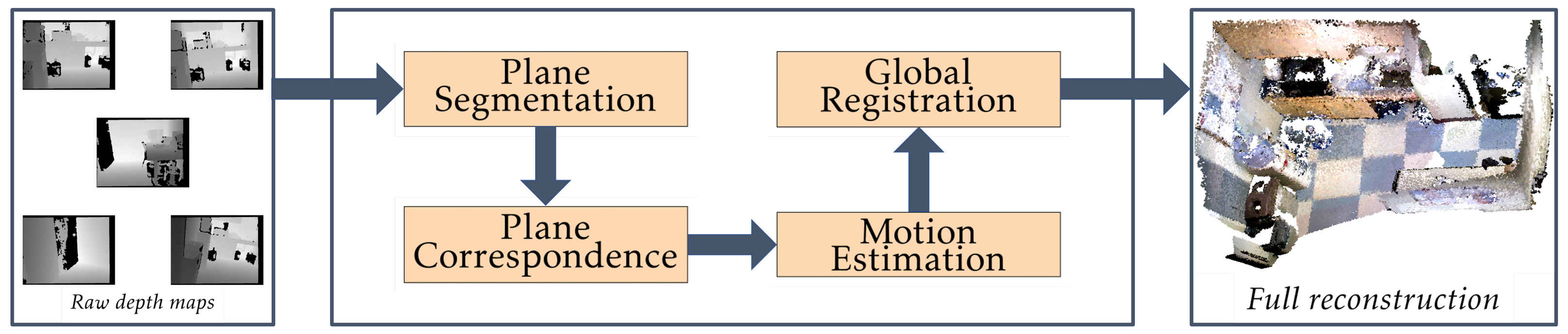
Abstract

We present a fast and lightweight method for 3D registration of scenes by exploiting the presence of planar regions.

Introduction

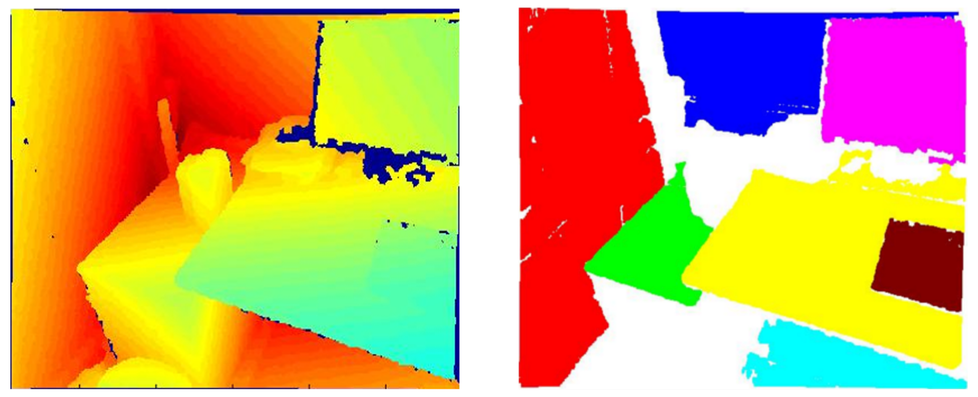
- Generation of 3D maps of real world scenes for robotic navigation, area surveillance and other such applications is of growing interest.
- Most of the available 3D registration techniques based on point correspondences between scans do not scale well for faster implementations.
- Our method exploits the presence of planes in common real world indoor environments to produce a fast and lightweight registration pipeline.

Overall Pipeline



Plane Segmentation

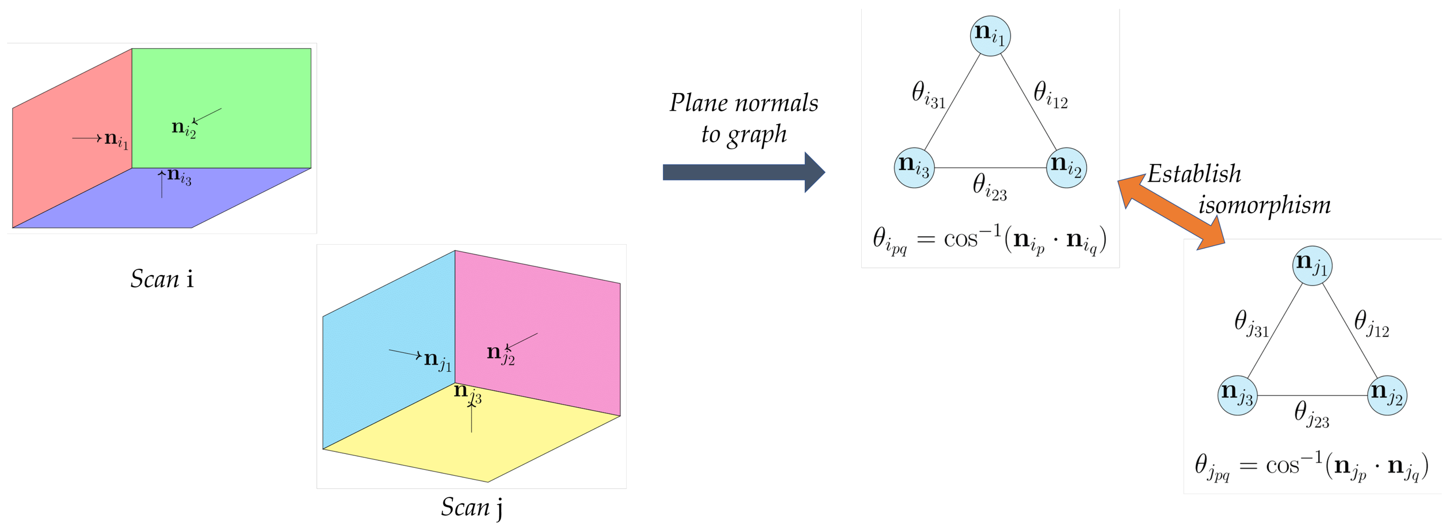
- Exploit **affine relationship** between image points corresponding to 3D planar regions and their disparity map.
- Pass the disparity maps through **LoG filters**: zero response in the output up to noise correspond to planar regions.



Raw depth map *Segmented planes*

Plane Correspondence

- Angles between plane normals in scans are **invariant to rigid motions**.
- **Graph isomorphism** used to establish correspondence between the plane normals.

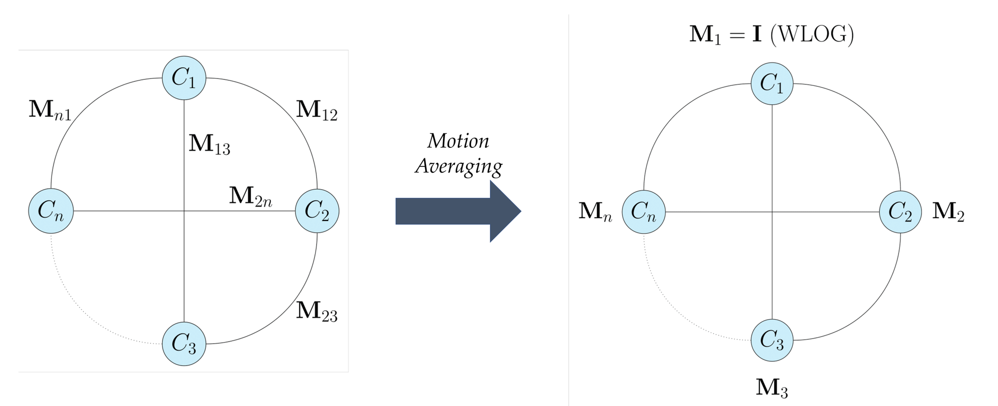


Motion Estimation

	≥ 3 corresponded planes	< 3 corresponded planes
Rotation Estimation $[\mathbf{Y} = \mathbf{N}_j\mathbf{N}_i^\top]$	**Umeyama's method:** $\mathbf{R}_{ij} = \mathbf{USV}^\top$; $\mathbf{UDV}^\top = svd(\mathbf{Y})$, $\mathbf{D} = diag(d_i), d_i \geq d_j \ \forall\ i > j$, $\mathbf{S} = \begin{cases} \mathbf{I}, & det(\mathbf{Y}) \geq 0 \\ diag(1,1,-1), & det(\mathbf{Y}) < 0 \end{cases}$	**Modified Horn's method:** $\mathbf{R}_{ij} = \mathbf{YQ} \pm \frac{\mathbf{Z}}{\sqrt{\lvert trace(\mathbf{Z})\rvert}}$; $\mathbf{Q} = \left(\frac{\mathbf{u}_1\mathbf{u}_1^\top}{\sqrt{\lambda_1}} + \frac{\mathbf{u}_2\mathbf{u}_2^\top}{\sqrt{\lambda_2}}\right)$, $\mathbf{Z} = \left[(\mathbf{YS})(\mathbf{YS})^\top \quad -\mathbf{I}\right]\mathbf{u}_3\mathbf{u}_3^\top$, $\{\lambda_j\}, \{\mathbf{u}_j\} \leftarrow eig(\mathbf{Y}^\top\mathbf{Y})$
Translation Estimation	Solve for $\mathbf{t}_{ij}$: $\begin{bmatrix} \mathbf{n}_{i_1}^\top \mathbf{R}_{ij} \\ \vdots \\ \mathbf{n}_{i_X}^\top \mathbf{R}_{ij} \end{bmatrix} \mathbf{t}_{ij} = \begin{bmatrix} d_{i_1} - d_{j_1} \\ \vdots \\ d_{i_X} - d_{j_X} \end{bmatrix}$; $X = \#$ correspondences	**ICP** for translation *only*

Global Registration

- $C_k, k = 1 \ldots n$: Cameras.
- $\mathbf{M}_{ij} = \begin{bmatrix} \mathbf{R}_{ij} & \mathbf{t}_{ij} \\ \mathbf{0}^\top & 1 \end{bmatrix}$: Relative motion between Cameras i and j.
- $\mathbf{M}_k$: Absolute motion of Camera k w.r.t. global frame of reference, $\mathbf{M}_{ij} = \mathbf{M}_j\mathbf{M}_i^{-1}$.

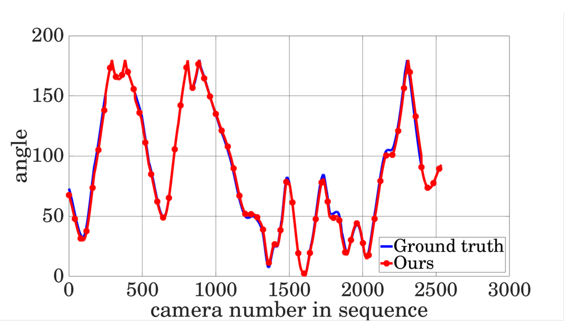


Results

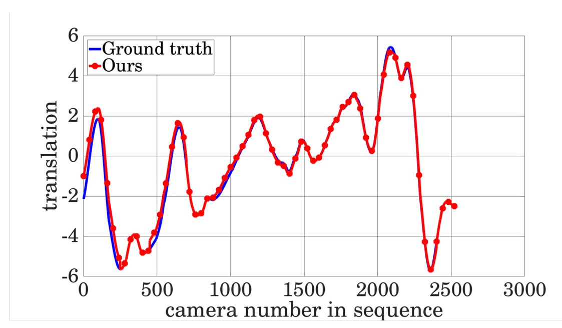
Table: *Mean distance* from ground truth surface for full reconstruction in METERS. *Running time* measured on Intel Core i7-5960X 3 GHz processor with 32 GB RAM in SECONDS

	Dataset	Mean distance			Running time		
		ICP+MA	Zhou *et al.*	Our method	ICP+MA	Zhou *et al.*	Our method
Aug. ICL-NUIM	livingroom1_noisy	0.10	**0.05**	**0.05**	14,260	7,460	**1,380**
	livingroom2_noisy	0.07	**0.06**	**0.06**	11,680	6,110	**970**
	office1_noisy	0.08	**0.03**	**0.03**	13,370	6,990	**1,270**
	office2_noisy	0.08	**0.05**	**0.05**	12,620	6,600	**1,050**
Sun3D	harvard_c8/hv_c8_3	0.07	**0.05**	**0.05**	4,900	2,620	**950**
	mit_32_d507/d507_2	0.07	**0.04**	**0.04**	26,590	14,240	**5,500**

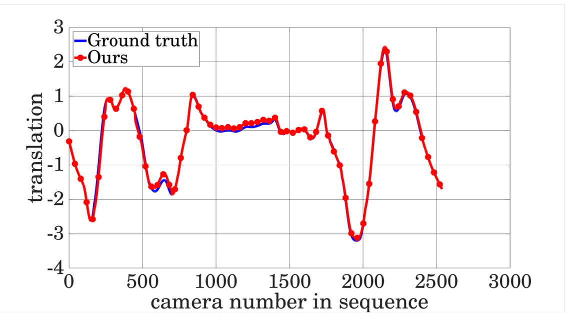
Figure: Registration alignment achieved by the different methods on a sample pair of scans from mit_32_d507/d507_2



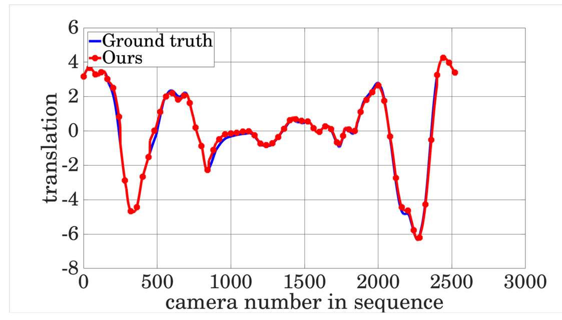
Comparison of recovered rotation angle with ground truth



Comparison of recovered translation along x-coordinate with ground truth



Comparison of recovered translation along y-coordinate with ground truth



Comparison of recovered translation along z-coordinate with ground truth

Figure: Comparison of full camera trajectory recovered by our method with the ground truth for office2_noisy

Conclusion

- In the presence of planes in scans, our registration method provides accuracy at par with state-of-the-art approaches in the literature while being significantly faster than them.
- Since planar representations adequately summarize the 3D information of many points in a scan, our approach results in significantly smaller memory requirements.

References

[1] Qian-Yi Zhou, Jaesik Park, and Vladlen Koltun. Fast Global Registration. In *Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II*, pages 766–78, Cham, 2016. Springer International Publishing.