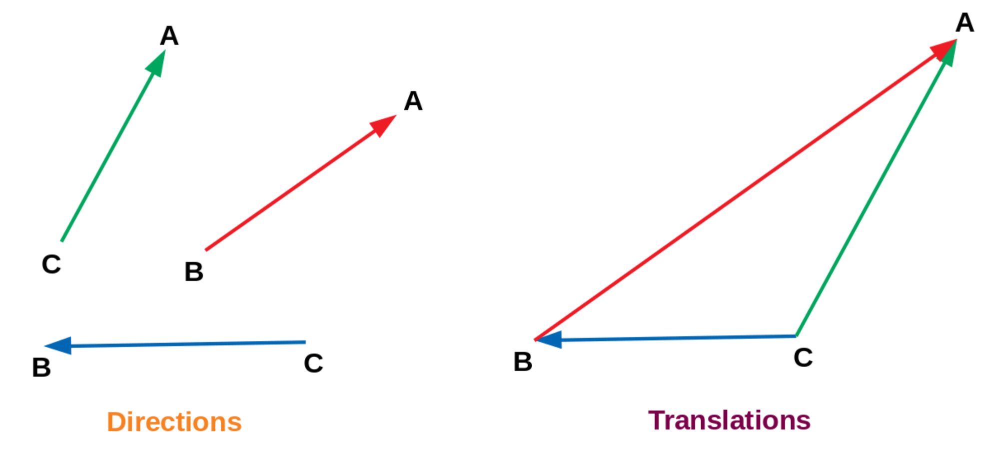
Sensitivity in Translation Averaging

Lalit Manam and Venu Madhav Govindu

Computer Vision Lab, Department of Electrical Engineering, Indian Institute of Science, Bengaluru, INDIA

NEURAL INFORMATION PROCESSING SYSTEMS

Translation Averaging in 3D Computer Vision



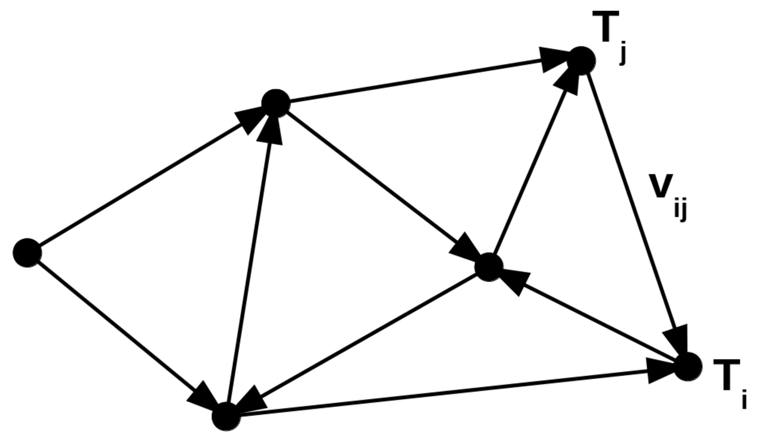
Solving for **translations** given **directions**.

- Input: Directions.
- Output: Translations.
- Solution upto scale and origin.
- Assumes known 3D rotations.
- Application: Structure-from-Motion.

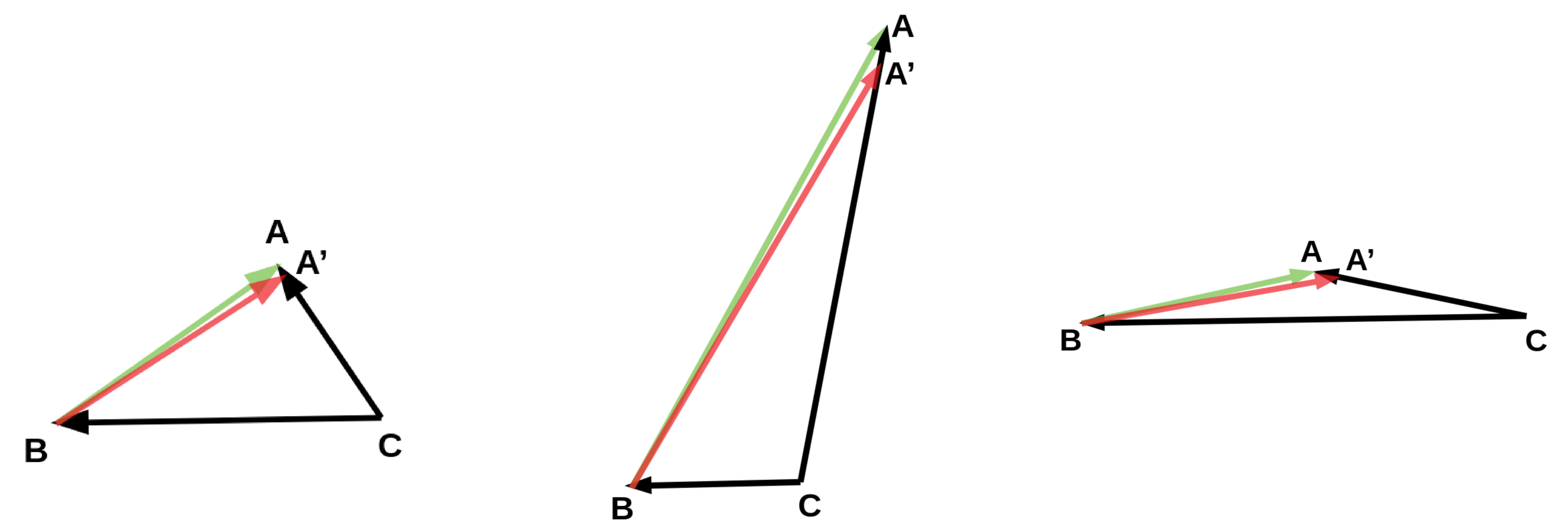
Problem Setup

- Represented as a network: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$.
- Nodes: absolute translations $\mathbf{T}_i \in \mathbb{R}^3, i \in \mathcal{V}$.
- Edges: directions $\mathbf{v}_{ij} \in \mathbb{S}^2, (i,j) \in \mathcal{E}$.
- Called bearing-based network.
- Consistency equation with scale (s) and origin ($\mathbf{T}_O$) ambiguity:

$$\mathbf{v}_{ij} = \frac{\mathbf{T}_j - \mathbf{T}_i}{\|\mathbf{T}_j - \mathbf{T}_i\|} = \frac{s(\mathbf{T}_j - \mathbf{T}_O) - s(\mathbf{T}_i - \mathbf{T}_O)}{\|s(\mathbf{T}_j - \mathbf{T}_O) - s(\mathbf{T}_i - \mathbf{T}_O)\|}.$$



Main Considerations



Well-conditioned triangle | Ill-conditioned triangle (Type-I) | Ill-conditioned triangle (Type-II)

Conditioning with respect to perturbations in directions (green $\rightarrow$ red $\implies$ A $\rightarrow$ A').

- Parallel rigidity [1]: Deals with the uniqueness of the solution.
- Outlier detection/suppression [2, 5]: Deals with gross errors in input directions.
- **Sensitivity [Ours]: Deals with small perturbations in input directions.**

Theoretical Results: Smallest Solvable Network

Formulation: Consider the smallest solvable bearing-based network i.e. a network $\mathcal{G}_\Delta$ of 3 nodes, $\mathcal{V}_\Delta = \{1,2,3\}$, with all possible edges, $\mathcal{E}_\Delta = \{(1,2),(2,3),(3,1)\}$. Least squares problem to estimate scales: $\min_{\mathbf{s}} \|\mathbf{V}\mathbf{s}\|^2$ s.t. $\|\mathbf{s}\|^2 = 1$.
Solution: Eigen vector corresponding to smallest eigenvalue of $\mathbf{V}^\mathbf{T}\mathbf{V}$.

Theorem 1: For a set of consistent directions, $\mathbf{v}_{ij}, (i,j) \in \mathcal{E}_\Delta$, in $\mathcal{G}_\Delta$, the absolute change in any eigenvalue of $\mathbf{V}^\mathbf{T}\mathbf{V}$, denoted as $|\delta\lambda|$, when the directions $\mathbf{v}_{ij}$ are perturbed by small rotations $\delta\mathbf{R}_{ij}$, with $\mathbf{n}_{ij}$ and $\delta\theta_{ij} > 0$ being the rotation axis and angle, is bounded by

$$|\delta\lambda| \leq \sum_{(k,i,j) \in TI(\Delta)} \delta\theta_{ij} \cdot \|\mathbf{v}_{ki}^{\mathbf{n}_{ij}\perp}\| \cdot \|\mathbf{v}_{kj}^{\mathbf{n}_{ij}\perp}\| \cdot \frac{\left|\sin\phi_{(k,i),(k,j)}^{\mathbf{n}_{ij}\perp}\right|}{\sin^2\phi_{(k,i),(k,j)}} \cdot$$
$$\left[\left(1 + (\mathbf{v}_{ik}^T\mathbf{v}_{jk})^2\right)\left((\mathbf{v}_{ij}^T\mathbf{v}_{ik})^2 + (\mathbf{v}_{ij}^T\mathbf{v}_{jk})^2\right) - 4 \cdot \mathbf{v}_{ik}^T\mathbf{v}_{jk} \cdot \mathbf{v}_{ij}^T\mathbf{v}_{ik} \cdot \mathbf{v}_{ij}^T\mathbf{v}_{jk}\right]^{\frac{1}{2}},$$

where $TI(\Delta) = \{(1,2,3),(2,3,1),(3,1,2)\}$, $\mathbf{v}^{\mathbf{n}_{ij}\perp}$ is the component of $\mathbf{v}$ orthogonal to $\mathbf{n}_{ij}$, $\phi_{(k,i),(k,j)}^{\mathbf{n}_{ij}\perp}$ is the angle between $\mathbf{v}_{ki}^{\mathbf{n}_{ij}\perp}$ and $\mathbf{v}_{kj}^{\mathbf{n}_{ij}\perp}$, and $\phi_{(k,i),(k,j)}$ is the angle between $\mathbf{v}_{ki}$ and $\mathbf{v}_{kj}$.

Angle Matrix for $\mathcal{G}_\Delta$: $\mathbf{A}_{\mathcal{G}_\Delta} \in \mathbb{R}^{3\times3}$ ($\phi_{(k,i),(k,j)}$: angle between $\mathbf{v}_{ki}$ and $\mathbf{v}_{kj}$)

$$\mathbf{A}_{\mathcal{G}_\Delta} = \begin{bmatrix} \phi_{(1,2),(1,2)} & \phi_{(2,1),(2,3)} & \phi_{(1,2),(1,3)} \\ \phi_{(2,3),(2,1)} & \phi_{(2,3),(2,3)} & \phi_{(3,2),(3,1)} \\ \phi_{(1,3),(1,2)} & \phi_{(3,1),(3,2)} & \phi_{(3,1),(3,1)} \end{bmatrix} = \begin{bmatrix} 0 & \phi_{(2,1),(2,3)} & \phi_{(1,2),(1,3)} \\ \phi_{(2,1),(2,3)} & 0 & \phi_{(3,2),(3,1)} \\ \phi_{(1,2),(1,3)} & \phi_{(3,2),(3,1)} & 0 \end{bmatrix}.$$

Theorem 2: Consider the bearing based-network of 3 nodes and 3 edges, $\mathcal{G}_\Delta = (\mathcal{V}_\Delta, \mathcal{E}_\Delta)$, and the corresponding angle matrix $\mathbf{A}_{\mathcal{G}_\Delta}$. The conditioning of the matrix $\mathbf{A}_{\mathcal{G}_\Delta}$ signifies the skewness of the triangle formed using the directions in $\mathcal{E}_\Delta$.

Theoretical Results: General Network Consisting of Triplets

Extending angle matrix:
- **Angle Matrix for $\mathcal{G}$: $\mathbf{A}_\mathcal{G}$** $\in \mathbb{R}^{M\times M}$ (M is the number of edges in $\mathcal{G}$).
- **Assumption:** All edges in $\mathcal{G}$ are a part of at least one triplet.

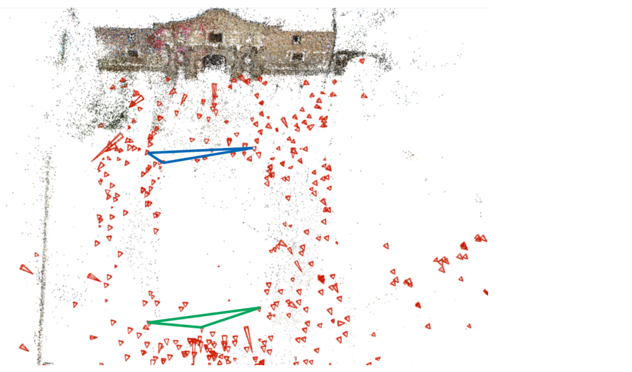
Conditioning of Translation Averaging: Defined as condition number of the angle matrix $\mathbf{A}_\mathcal{G}$.

Theorem 3: Consider a bearing-based network $\mathcal{G}$, with all edges contributing to triplets. The angle matrix $\mathbf{A}_\mathcal{G}$, corresponding to $\mathcal{G}$, is well conditioned if the minimum angle (or equivalently all the angles) in all the triangles formed by the triplets are sufficiently large.

Dealing with uniqueness of solution:
- Theorem 3 does not assume parallel rigidity of the network $\mathcal{G}$.
- Construct triplet network: $\mathcal{G}_T = (\mathcal{V}_T, \mathcal{E}_T)$, Nodes $\mathcal{V}_T$ denote a triplet in $\mathcal{G}$, Edges $\mathcal{E}_T$ connect the nodes if an edge is common between the triplets in $\mathcal{G}$.

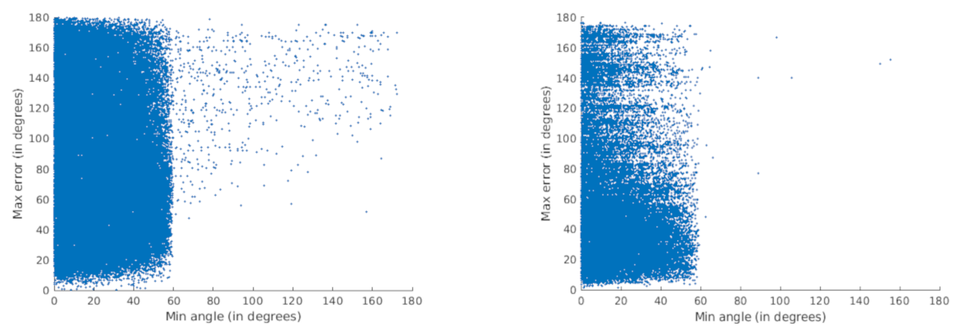
Theorem 4: Given a bearing-based network $\mathcal{G}$, with all edges contributing to triplets forming triangles, and its corresponding triplet network $\mathcal{G}_T$, the maximal parallel rigid component of $\mathcal{G}$ can be determined by the edges in $\mathcal{G}$ contributing to the largest connected component of $\mathcal{G}_T$.

Analysis of Real Data



Alamo reconstruction [4] using COLMAP [3].

Blue triangle: Ill-conditioned (Type-I).
Green triangle: Ill-conditioned (Type-II).



Piccadilly | Tower of London

Maximum error of directions vs minimum angle between edges.

Abundance of skewed triangles.

Outlier presence $\neq$ Skewed triangle.

Experimental Results

Dataset	$\#N_{rem}$	$\#M_{rem}$	Mean-ATE ↓		RMS-ATE ↓		Removed Node Errors		P_{tri} ↑		BA_{iters} ↓	
			w/o filter	w/ filter	w/o filter	w/ filter	Mean	RMS	w/o filter	w/ filter	w/o filter	w/ filter
Alamo	12	197	4.7	**4.5**	11.1	**10.5**	22.9	39.5	142	**144**	73	**54**
Ellis Island	8	69	23.2	**22.1**	**50.7**	51.8	98.6	118.1	57	**58**	40	**33**
Gendarmenmarkt	24	257	50.6	**40.9**	77.8	**61.1**	149.7	190.8	101	**104**	31	**21**
Madrid Metropolis	24	197	13.9	**12.7**	29.4	**26.1**	54.2	61.2	46	**49**	**66**	83
Montreal Notre Dame	4	147	4.4	**4.3**	10.0	**9.9**	31.2	34.1	115	**118**	59	**37**
NYC Library	13	211	6.5	**5.2**	15.8	**12.6**	28.7	38.1	77	**79**	100	**74**
Notre Dame	3	108	**3.3**	**3.3**	**6.4**	**6.4**	6.9	7.0	**376**	375	34	**22**
Piazza del Popolo	18	268	**8.0**	**8.0**	**13.0**	13.1	15.0	17.3	91	**94**	**34**	44
Piccadilly	61	901	5.3	**5.1**	11.0	**10.8**	20.8	29.2	231	**234**	53	**47**
Roman Forum	41	394	12.8	**10.4**	27.0	**19.6**	65.4	100.5	**196**	**196**	31	**21**
Tower of London	17	141	15.6	**14.5**	32.2	**30.3**	63.7	90.3	107	**110**	89	**85**
Union Square	26	232	14.5	**10.6**	24.8	**18.1**	34.5	49.9	57	**56**	122	**52**
Vienna Cathedral	29	282	10.2	**10.2**	**17.8**	18.4	21.4	27.5	219	**222**	52	**25**
Yorkminster	34	363	20.4	**19.3**	29.5	**28.3**	44.0	51.4	**175**	168	58	**33**

Absolute translations errors (ATE) (in m), points triangulated ($P_{tri} \times 10^3$) and bundle adjustment iterations (BA_{iters}) on 1DSfM [4] datasets using BATA [5]. Removed Node Errors: Errors of removed nodes in the unfiltered network.

Conclusion

This work deals with sensitivity in translation averaging under input uncertainty;
- studies sensitivity in estimating edge scales suggesting skewed triangles are unstable,
- defines conditioning of translation averaging problem and provides a sufficient criterion to ensure it is well-conditioned,
- proposes an efficient algorithm to remove skewed triangles from the network while ensuring parallel rigidity,
- demonstrates the effectiveness of the proposed filter with better absolute translations, more 3D points triangulated and faster convergence of bundle adjustment for filtered networks.

References

[1] Arrigoni, F., Fusiello, A.: Bearing-based network localizability: a unifying view. IEEE Transactions on Pattern Analysis and Machine Intelligence (2018)
[2] Ozyesil, O., Singer, A.: Robust camera location estimation by convex programming. In: IEEE Conference on Computer Vision and Pattern Recognition (2015)
[3] Schonberger, J.L., Frahm, J.M.: Structure-from-motion revisited. In: IEEE Conference on Computer Vision and Pattern Recognition (2016)
[4] Wilson, K., Snavely, N.: Robust global translations with 1dsfm. In: European Conference on Computer Vision (2014)
[5] Zhuang, B., Cheong, L.F., Lee, G.H.: Baseline desensitizing in translation averaging. In: IEEE Conference on Computer Vision and Pattern Recognition (2018)

Acknowledgments

Lalit Manam is supported by a Prime Minister's Research Fellowship, Government of India. This research was supported in part by a Core Research Grant from the Science and Engineering Research Board, Department of Science and Technology, Government of India. Travel supported by Google Travel Grant, NeurIPS 2023 Scholar Award and Indian Institute of Science.

{lalitmanam,venug}@iisc.ac.in