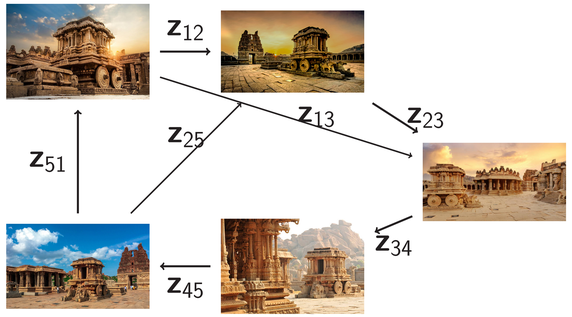
Adaptive Annealing for Robust Averaging

Chitturi Sidhartha and Venu Madhav Govindu

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

EUROPEAN CONFERENCE ON COMPUTER VISION
MILANO 2024

Introduction



- Global pipelines for large-scale 3D reconstruction entail solving the global poses of cameras given the relative poses between pairs of cameras.
- **Averaging**: Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and a set of outlier-prone pairwise observations $\{z_{ij} \mid (i,j) \in \mathcal{E}\}$, where z_{ij} is a function of x_i and x_j, estimate the node values $\{x_i \mid i \in \mathcal{V}\}$.
- Minimization of a robust cost to mitigate outliers.

$$f_\sigma(x) = \sum_{ij \in E} \rho_\sigma(\|r_{ij}(x_i, x_j; z_{ij})\|) \tag{1}$$

Graduated Non-Convexity (GNC)

- Solve Eq. 1 by minimizing a sequence of increasingly hard-to-solve robust costs $\rho_\sigma(\cdot)$, $\sigma = \sigma_0 > \sigma_1 > \sigma_2 > \cdots > \sigma_{min}$.
- Adaptive $\{\sigma_k\}$ schedule [4] $\Rightarrow$ Good accuracy-speed tradeoff.
- **Core Idea**: Track the positive-definiteness of the Hessian (H) to determine the next σ.
- We exploit the **graph Laplacian structure** of the Hessian matrices in averaging problems.

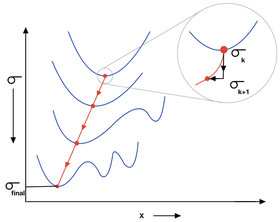


Figure: Annealing σ

$$H(\sigma) = \begin{bmatrix} D_1(\sigma) & -W_{12}(\sigma) & -W_{13}(\sigma) & \cdots & -W_{1N}(\sigma) \\ -W_{21}(\sigma) & D_2(\sigma) & -W_{23}(\sigma) & \cdots & -W_{2N}(\sigma) \\ -W_{31}(\sigma) & -W_{32}(\sigma) & D_3(\sigma) & \cdots & -W_{3N}(\sigma) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -W_{N1}(\sigma) & -W_{N2}(\sigma) & -W_{N3}(\sigma) & \cdots & D_N(\sigma) \end{bmatrix}$$

where $D_i(\sigma) = \sum_{j=1, j\neq i}^{N} W_{ij}(\sigma)$, and $W_{ij}(\sigma)$ is the weight matrix corresponding to z_{ij}.

Proposed Annealing Scheme

Laplacian Property-1: *The Laplacian matrix $H(\sigma)$ is positive semidefinite if $W_{ij}(\sigma) \succcurlyeq \mathbf{0}$, $\forall (i,j) \in \mathcal{E}$.*

$\Downarrow$

Anneal σ to ensure $W_{ij}(\sigma) \succcurlyeq \mathbf{0}\ \forall (i,j) \in \mathcal{E}$. Once this is not possible, reduce σ so that at most p% of edges have $W_{ij} \not\succcurlyeq \mathbf{0}$.

Percentile Approach:

- Let $\sigma_{ij} = \text{argmin}_\sigma\{\lambda_{min}(W_{ij}(\sigma)) = 0\}$. Set σ as the $(100-p)^{th}$ percentile of $\{\sigma_{ij}\}$.
- Increase p in fixed steps until $p = 50$ or $\sigma \geq \sigma_{min}$.

Advantages

- σ_{ij} is a closed-form function of r_{ij}, enabling efficient σ computation for the next GNC stage.
- σ_{min} can be adaptively chosen as the k^{th} percentile of $\{\sigma_{ij}\}$, effective for up to k% outliers.
- Efficiently detect $\lambda_{min}(H) > 0$ via power iteration by assessing the eigenvector's proximity to the all-ones vector.

$\Uparrow$

Laplacian Property-2: *A Laplacian with no negative eigenvalue has its smallest eigenvector as the all-ones vector.*

Illustrative Problems

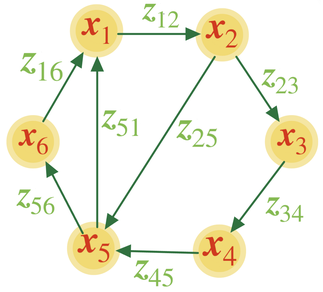


Figure: Vector Averaging

Input: Relative **Displacements** (z_{ij})

$$\min_{\{x_i\}} \sum_{ij \in E} \rho_\sigma(\|x_i - x_j - z_{ij}\|)$$

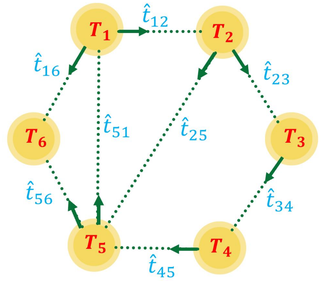


Figure: Translation Averaging

Input: Relative **Directions** ($\hat{t}_{ij}$)

$$\min_{\{T_i\},\{d_{ij}\}} \sum_{ij \in E} \rho_\sigma(\|T_i - T_j - d_{ij}\hat{t}_{ij}\|)$$

Output: **Camera/Sensor Locations** (x_i, T_i)

Results

Superior performance on synthetic datasets for varying outlier fractions and graph densities.

	IRLS			GNC ($\delta \downarrow$)			GNC ($\delta \uparrow$)			SE-Sync [3]			GNCp		
lr1	2.15	2.04	0.05	**1.93**	1.60	0.65	2.00	1.66	0.15	2.15	2.11	7.7	**1.93**	**1.59**	0.22
lr2	1.39	1.32	0.27	**1.25**	**1.31**	3.64	1.31	**1.31**	0.53	1.55	1.53	28	**1.25**	**1.31**	3.11
of1	1.87	1.89	0.06	1.81	1.80	0.52	1.87	1.88	0.10	1.84	1.88	9.5	**1.80**	**1.79**	0.18
of2	2.15	2.20	0.18	**2.12**	1.94	1.68	2.17	2.24	0.34	2.16	2.24	14	**2.12**	**1.93**	0.95

Table: Vector averaging on Augmented ICL-NUIM dataset (*lr*:livingroom, *of*:office): {Mean | median errors | time taken} of the baselines

Dataset	N	M	Mean | Median errors ↓									
			LUD [2]		1DSfM [5]		ShapeFit [1]		BATA [6]		GNCp-LUD	
Alamo	586	81437	2.7	0.5	4.E+03	0.5	20.9	0.6	2	0.6	**1.9**	0.5
Ellis Island	229	14728	6.9	3.6	2.E+04	3.4	35	6.1	6.7	3.2	**6.6**	3.1
Gendarmenkmart	686	27145	31.2	11.3	1.E+04	11.4	36.8	14.8	31.3	11.4	**27.8**	11.3
Madrid Metropolis	325	11995	8.4	1.9	2.E+05	2	24	2.8	6.9	**1.6**	**6.4**	1.6
Montreal Notre Dame	461	45737	0.9	0.5	6.E+03	0.5	2.2	0.9	**0.8**	0.5	**0.8**	0.5
NYC Library	337	14365	2.2	0.8	4.E+04	0.8	2.9	1.2	**2.1**	1.7	2.3	**0.8**
Notre Dame	552	80647	1.2	0.3	1.E+05	0.3	6.8	0.5	**1**	0.2	1.1	**0.2**
Piazza del Popolo	334	20974	3.8	2.8	8.E+03	1.7	20.1	**1.6**	**3.4**	2	3.5	2.3
Piccadilly	2362	201600	**2.8**	1.3	2.E+04	1.2	19.8	1.3	3.2	**1.1**	3.2	1.2
Roman Forum	1069	54207	11.9	3.3	1.E+05	1.8	18.5	6.3	8.3	2	**7.3**	1.7
Tower of London	474	19252	14.9	3.2	1.E+05	3	15.3	6.6	12.1	2.9	**10.8**	3
Union Square SF	1369	123081	2.4	0.8	5.E+04	1.1	6.7	1.6	2.6	0.8	**2.3**	**0.7**
Vienna Cathedral	876	52274	6.2	0.5	4.E+05	0.5	10.4	1.1	6.3	0.5	**5.8**	0.5
Yorkminster	1017	45117	10.2	4.7	2.E+04	4.3	14.8	6.3	8.3	3.2	**8**	4.2

Table: Translation Averaging on real data: Mean | Median errors (in metres) of the baselines

Conclusion

- In averaging problems, we exploit the graph Laplacian structure of the Hessian matrices to devise a principled method to adaptively anneal the scale in Graduated Non-Convexity (GNC).
- We demonstrate the superior accuracy and efficiency of our approach compared to the state-of-the-art on two problems: Vector averaging, Translation averaging.

Acknowledgments

Chitturi Sidhartha was supported by the Prime Minister's Research Fellowship, and this research was supported partially by a Core Research Grant, both from the Government of India.

References

[1] Goldstein, T., Hand, P., Lee, C., Voroninski, V., Soatto, S.: Shapefit and shapekick for robust, scalable structure from motion. In: European Conference on Computer Vision. pp. 289–304. Springer (2016)

[2] Ozyesil, O., Singer, A.: Robust camera location estimation by convex programming. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2674–2683 (2015)

[3] Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J.: Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group. The International Journal of Robotics Research **38**(2-3), 95–125 (2019)

[4] Sidhartha, C., Manam, L., Govindu, V.M.: Adaptive annealing for robust geometric estimation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 21929–21939 (June 2023)

[5] Wilson, K., Snavely, N.: Robust global translations with 1dsfm. In: European Conference on Computer Vision. pp. 61–75. Springer (2014)

[6] Zhuang, B., Cheong, L.F., Lee, G.H.: Baseline desensitizing in translation averaging. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4539–4547 (2018)

{chitturis,venug}@iisc.ac.in

https://ee.iisc.ac.in/cvlab/research/adanroavg/