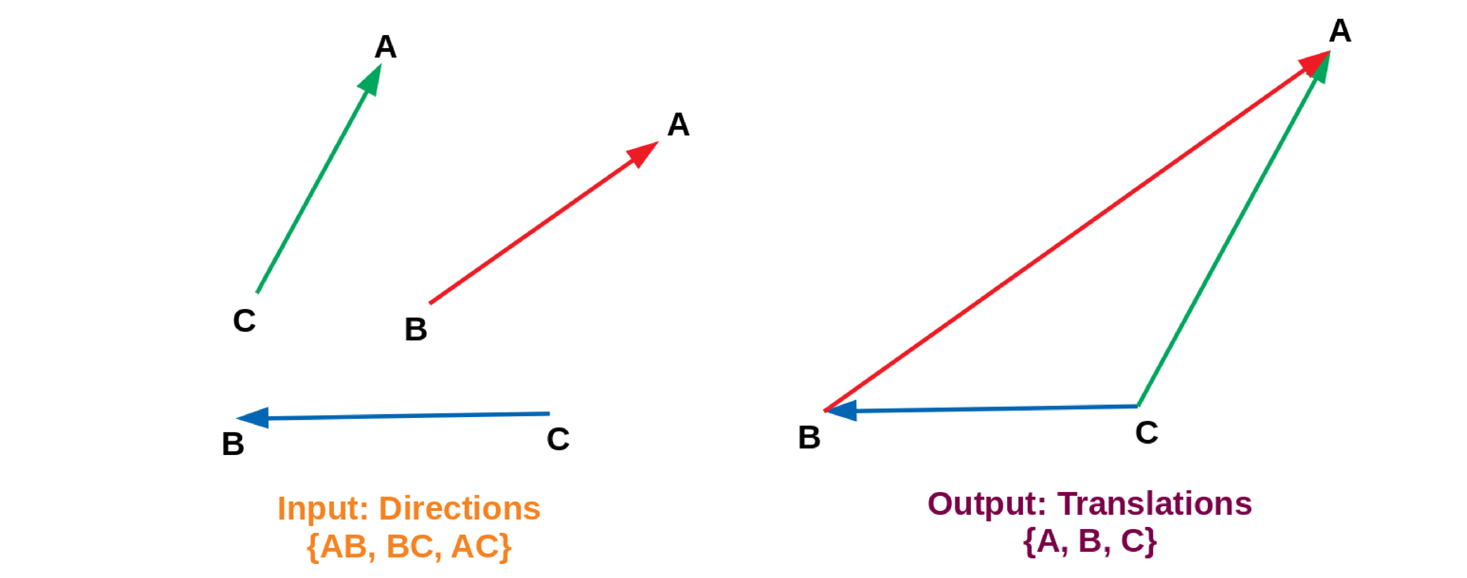
Fusing Directions and Displacements in Translation Averaging

Lalit Manam and Venu Madhav Govindu

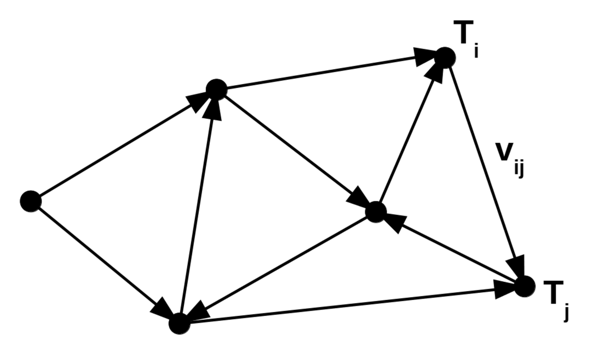
Department of Electrical Engineering, Indian Institute of Science, Bengaluru, INDIA

Translation Averaging



Solving for translations given directions

Problem Setup



- Graph representation: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$
- Input: Directions $\mathbf{v}_{ij}$ on edges
- Output: Translations $\mathbf{T}_i$ on nodes
- Consistency equation:

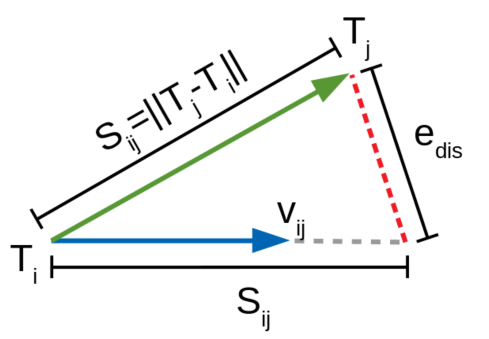
$$\frac{\mathbf{T}_j - \mathbf{T}_i}{\|\mathbf{T}_j - \mathbf{T}_i\|} = \mathbf{v}_{ij}$$

Cost Functions

Comparing Displacements

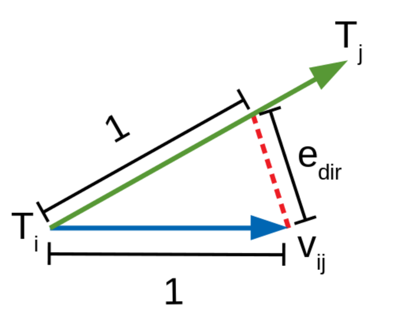
Comparing Directions

Original Costs



$$\sum_{(i,j)\in\mathcal{E}} \rho_{dis}\left(\|\mathbf{T}_j - \mathbf{T}_i - \|\mathbf{T}_j - \mathbf{T}_i\|\mathbf{v}_{ij}\|\right)$$

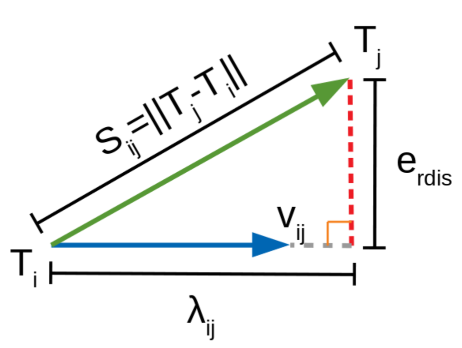
$e_{dis}(\mathbb{T})$



$$\sum_{(i,j)\in\mathcal{E}} \rho_{dir}\left(\left\|\frac{\mathbf{T}_j - \mathbf{T}_i}{\|\mathbf{T}_j - \mathbf{T}_i\|} - \mathbf{v}_{ij}\right\|\right)$$

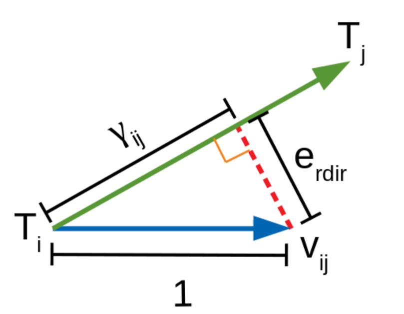
$e_{dir}(\mathbb{T})$

Relaxed Costs



$$\sum_{(i,j)\in\mathcal{E}} \rho_{rdis}\left(\|\mathbf{T}_j - \mathbf{T}_i - \lambda_{ij}\mathbf{v}_{ij}\|\right)$$

$e_{rdis}(\mathbb{T}, \Lambda)$



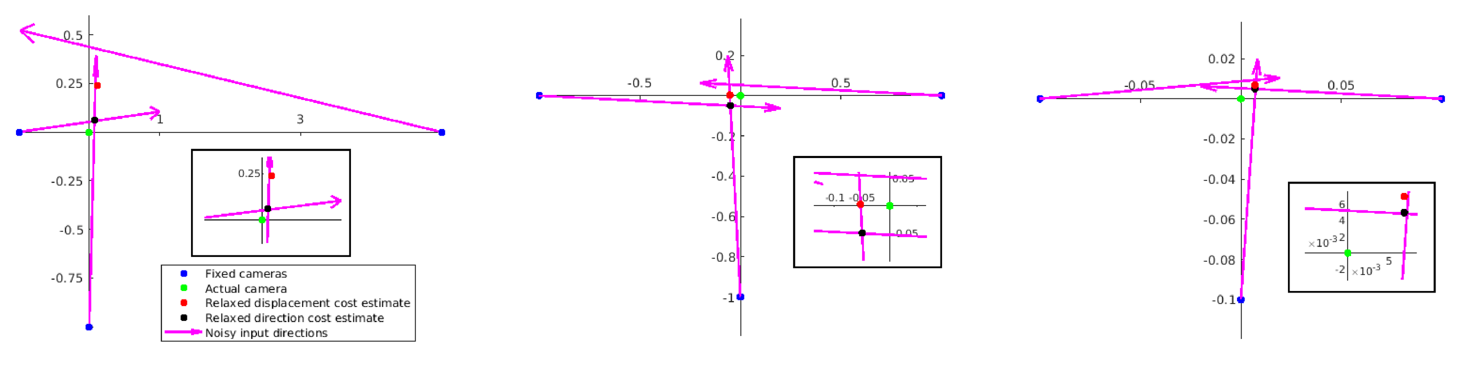
$$\sum_{(i,j)\in\mathcal{E}} \rho_{rdir}\left(\|(\mathbf{T}_j - \mathbf{T}_i)\gamma_{ij} - \mathbf{v}_{ij}\|\right)$$

$e_{rdir}(\mathbb{T}, \Gamma)$

Constraint Set:

$$\mathcal{C} = \left\{\mathbb{T} \,\middle|\, \frac{1}{|\mathcal{E}|}\sum_{(i,j)\in\mathcal{E}} \langle \mathbf{T}_j - \mathbf{T}_i, \mathbf{v}_{ij}\rangle = 1, \sum_{i\in\mathcal{V}} \mathbf{T}_i = \mathbf{0}\right\}$$

Behaviour of the Relaxed Costs



Disparate baselines $\eta = \{3°, 4°, 5°\}$ Baselines $= \{1, 1, 5\}$ units

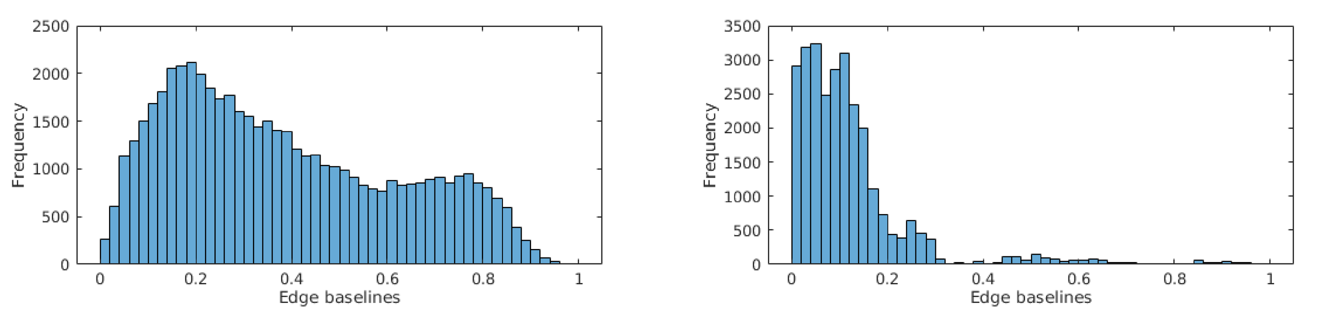
Similar baselines $\eta = \{3°, 3°, 3°\}$ Baselines $= \{1, 1, 1\}$ units

Similar baselines $\eta = \{5°, 4°, 3°\}$ Baselines $= \{0.1, 0.1, 0.1\}$ units

Behaviour of relaxed costs under varied baselines and noise η

Neither cost performs the best in all the scenarios

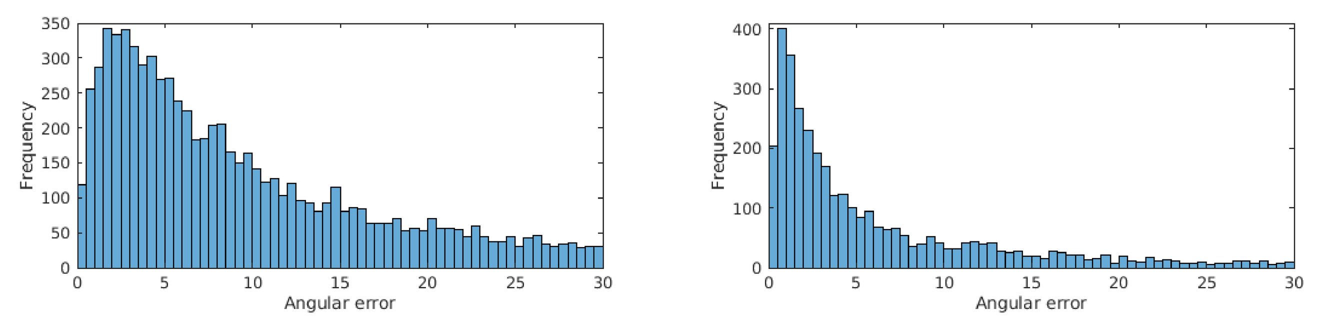
Analysis of Real Data



Montreal Notre Dame — Yorkminster

Histograms of ground truth normalized baselines

Distribution of baselines impacts performance

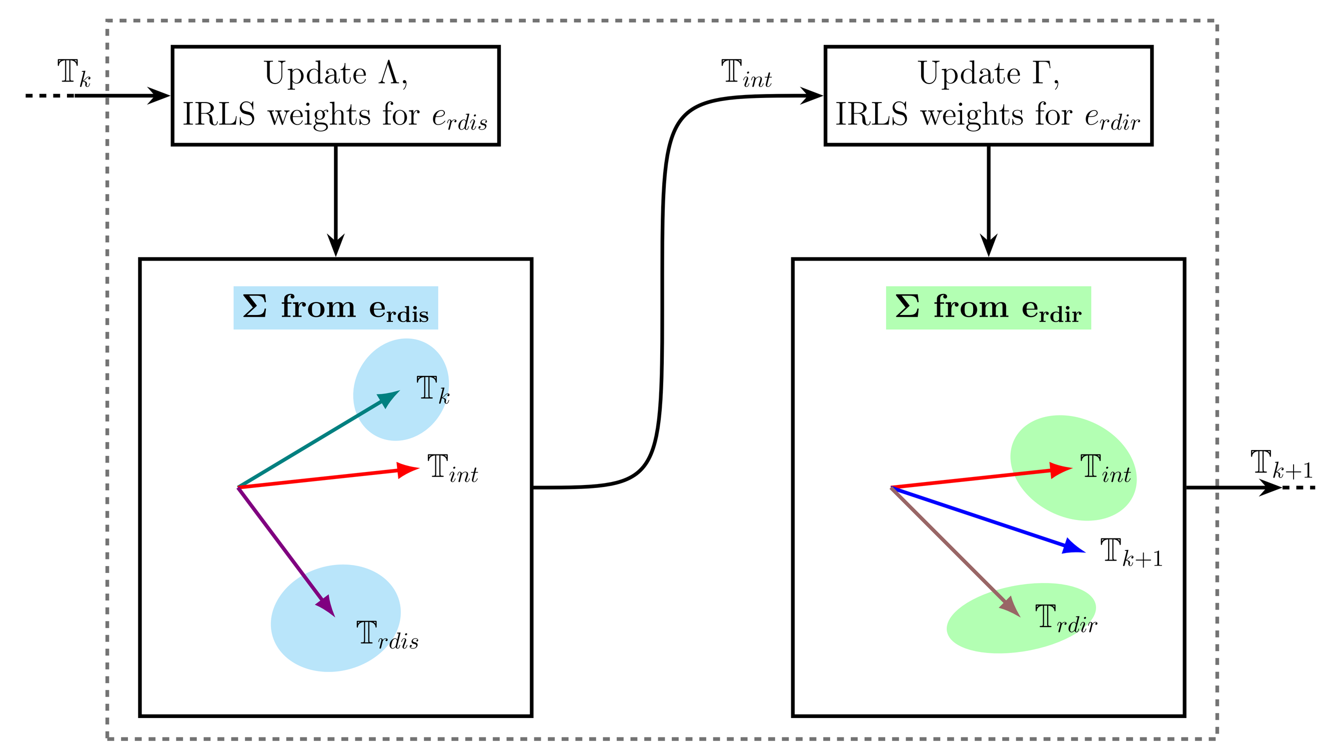


Ellis Island — Madrid Metropolis

Histograms of input direction errors (in degrees)

Distribution of noise impacts performance

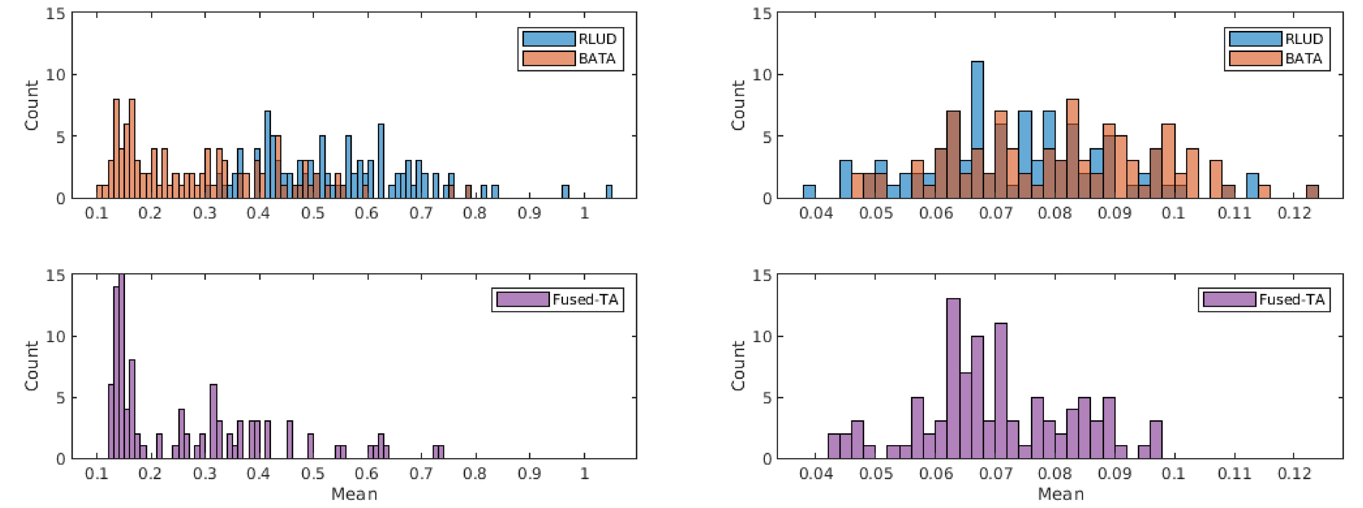
Our IRLS Method: Fused Translation Averaging



Gaussian assumption:

- Σ^{-1} = Hessian of the corresponding cost
- Bayesian update: $\mathbb{T} = \left(\Sigma_1^{-1} + \Sigma_2^{-1}\right)^{-1}\left(\Sigma_1^{-1}\mathbb{T}_1 + \Sigma_2^{-1}\mathbb{T}_2\right)$

Results



Disparate baselines with input noise $\sigma = 5$ — Similar baselines with input noise $\sigma = 5$

Histogram of mean translation errors for two synthetic datasets.

The leftward shift indicates the superior performance of our method

Dataset	$\|\mathcal{V}\|$	$\|\mathcal{E}\|$	Mean Errors					Median Errors				
			1DSfM [3]	LUD [2]	ShapeFit [1]	BATA [4]	Fused-TA (Ours)	1DSfM [3]	LUD [2]	ShapeFit [1]	BATA [4]	Fused-TA (Ours)
Alamo (ALM)	1016	16831	8e3	7.6	35.3	7.5	**7.0**	**2.4**	3.0	2.6	**2.4**	**2.4**
Ellis Island (ELS)	908	10383	107.3	45.1	39.7	42.6	**27.3**	32.3	26.0	**15.6**	25.3	17.5
Gendarmenkmart (GMM)	1026	14175	3e4	**42.3**	51.3	43.6	44.4	27.5	**25.9**	32.3	26.3	27.7
Madrid Metropolis (MDR)	627	4941	2e4	17.5	170.8	**15.3**	15.7	5.4	8.7	5.5	**4.8**	6.1
Montreal Notre Dame (MND)	599	18390	1e3	**4.0**	4.3	4.5	4.3	1.9	2.4	2.3	1.9	**1.8**
Notre Dame (ND)	1421	70771	4e4	4.0	3.7	**3.3**	**3.3**	**1.2**	1.8	1.3	**1.2**	**1.2**
NYC Library (NYC)	858	8173	3e4	**6.4**	3e3	8.2	7.6	3.1	3.0	3.6	2.5	**2.3**
Piazza del Popolo (PDP)	1038	15182	6e4	8.0	19.1	8.1	**7.6**	7.2	5.5	15.8	5.0	**4.5**
Piccadilly (PIC)	3124	46754	3e4	**6.5**	25.2	7.0	6.7	2.2	2.6	2.4	2.3	**2.1**
Roman Forum (ROF)	1575	19593	5e5	22.4	33.5	15.5	**15.3**	4.3	8.6	8.6	**4.4**	5.3
Tower of London (TOL)	824	9457	8e4	28.2	58.0	**22.9**	24.5	8.0	10.7	8.6	6.9	**6.5**
Trafalgar (TFG)	7483	115052	4e4	21.4	33.3	22.4	**21.3**	11.7	7.6	8.3	6.9	**6.2**
Union Square (USQ)	1166	13460	2e4	**14.5**	23.0	17.1	31.7	9.2	8.5	14.3	8.1	**7.9**
Vienna Cathedral (VNC)	1647	27386	7e4	14.8	172.0	14.9	**14.3**	12.8	6.5	7.1	6.5	**6.1**
Yorkminster (YKM)	1834	19177	7e4	21.1	72.6	21.6	**20.7**	22.9	12.9	15.8	13.2	**13.0**

Camera translation errors (in meters) on 1DSfM [3] datasets

Alamo — Montreal Notre Dame — Piazza del Popolo — Tower of London

Reconstructions obtained with triangulation using our Fused-TA translation estimate

Conclusion

Solving translation averaging:

- Displacement vs Direction costs
- No clear winner
- Our method: Similar to Bayes recursive filter
- Improves over individual methods

References

[1] Goldstein, T., Hand, P., Lee, C., Voroninski, V., Soatto, S.: Shapefit and shapekick for robust, scalable structure from motion. In: ECCV (2016)

[2] Ozyesil, O., Singer, A.: Robust camera location estimation by convex programming. In: CVPR (2015)

[3] Wilson, K., Snavely, N.: Robust global translations with 1dsfm. In: ECCV (2014)

[4] Zhuang, B., Cheong, L.F., Lee, G.H.: Baseline desensitizing in translation averaging. In: CVPR (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 Indian Institute of Science and Pratiksha Trust Travel Fellowship.

{lalitmanam,venug}@iisc.ac.in

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