Robust Global Translations with 1DSfM

Kyle Wilson
Noah Snavely
{wilsonkl, snavely}@cs.cornell.edu

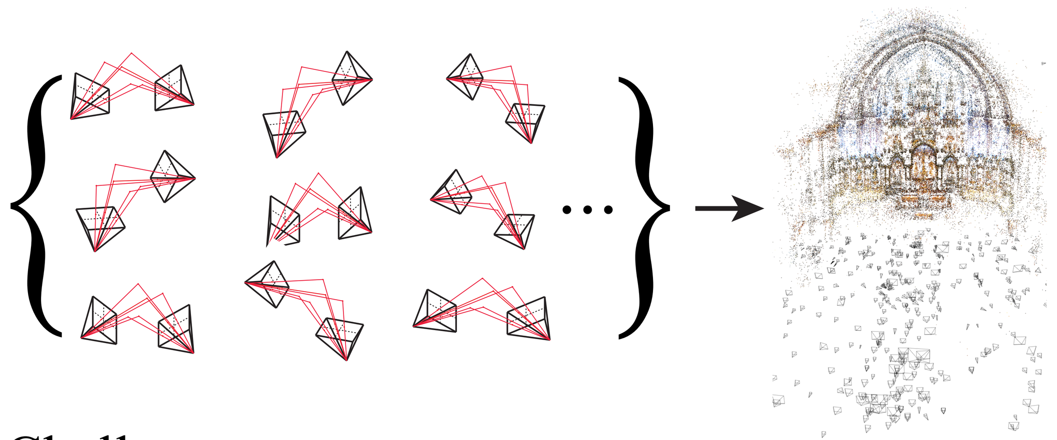
Code and Datasets: www.cs.cornell.edu/projects/1DSfM

Problem Statement

Incremental SfM is expensive and error-prone. We explore global methods to solve the problem in one shot.

Goal:

Build a 3D model in one shot given many two-view models. We use Chatterjee and Govindu [1] to solve for rotations, and focus only on translations.



Challenges:

- Many formulations of the translations problem are non-convex. A solver must find a good solution **reliably**.
- Translations problems generally contain **outliers**. These bad measurements can reduce solution quality and make it harder for solvers to converge.

Contributions:

1DSfM: a simple way to detect outlier translation measurements using 1D subproblems

Solver: a new approach to solving translations problems using nonlinear optimization

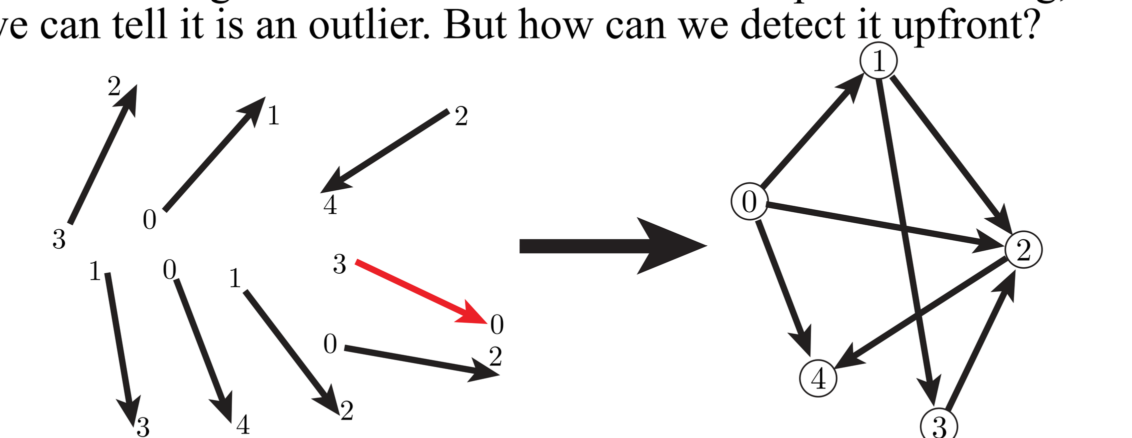
Takeaway:

We pose a translations problem as a standard nonlinear optimization, which, coupled with outlier removal, yields good results even when initialized randomly.

Contribution 1: Outlier Removal with 1DSfM

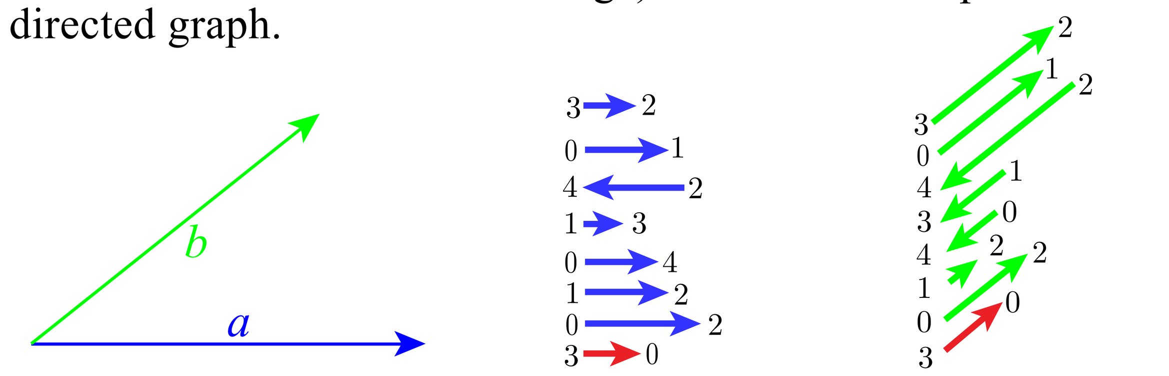
Left: an example translations problem
Right: the correct solution

An outlier edge is shown in red. Given the output embedding, we can tell it is an outlier. But how can we detect it upfront?



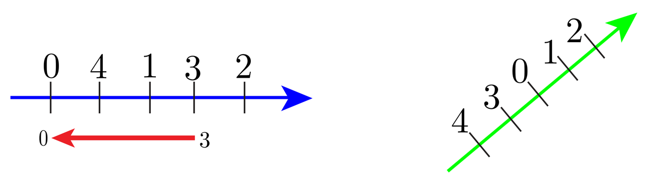
Input unit translation directions   Output: absolute camera positions

1D subproblems are easier: we project the problem onto a single unit vector, so each edge becomes a simple plus/minus sign (due to the unknown scale of each edge) which we can represent as a directed graph.



Two projection directions   1D subproblem (a)   1D subproblem (b)

These 1D problems are instances of Minimum Feedback Arc Set [2]. Solving them means choosing a best ordering. Outlier edges may not be consistent with the others.



solution to (a)   solution to (b)

Outliers won't be detected in some projections. We project in many random directions and reject edges that are frequently inconsistent.

Contribution 2: New Translations Solver

We want to solve problems of this general form:

Given:	a directed graph $G = (V, E)$ 3D translation directions $t : E \to S^2$
Compute:	an embedding $X : V \to \mathbb{R}^3$ (up to scale and translation)
Such that:	the translation directions induced by X are close to t

We compare poses in the **measurement space** of unit vectors with the squared chordal distance.

$$\hat{X} = \underset{X}{\operatorname{argmin}} \sum_{(i,j)\in E} d_{ch}\left(t_{ij}, \frac{X_j - X_i}{\|X_j - X_i\|}\right)^2$$

$$d_{ch}(u, v) = \|u - v\|$$

Properties:

- Nonlinear Least Squares problem (**NLLS**)—we use Ceres [3]
- Well-behaved error surface, especially after 1DSfM
- Can additionally use a Huber loss for even greater robustness
- Geometrically meaningful: **MLE** of the error model below

$$f(t_{ij}|X) \propto exp\left[\frac{-d_{ch}^2}{\sigma^2}\right]$$

Convergence:

- NLLS is a local optimizer—global convergence not guaranteed
- Surprisingly, **we find good solutions, even from random initializations**
- Plausibility: for a noise free problem, the error surface is decreasing towards the global optimum. It deviates from this behavior slowly as noise increases:

$$d_{ch}^2(t, X_\lambda) \leq d_{ch}^2(t, X_1) + d_{ch}^2(t, X_{opt})$$
$$\text{where } X_\lambda = \lambda X_1 + (1-\lambda)X_{opt}, \quad 0 \leq \lambda \leq 1$$

Results

- 13 large datasets—all new (except Notre Dame, from [5])
- state of the art results
- datasets and code available

We evaluate our results by robustly rigidly aligning solutions to models produced by Bundler, in incremental SfM solver [5].

The numbers below are errors in meters after a final bundle adjustment.

			no 1DSfM		with 1DSfM		[4]
Name	Size	N_c	$\tilde{x}$	$\bar{x}$	$\tilde{x}$	$\bar{x}$	$\tilde{x}$
Piccadilly	80	2152	**0.3**	9e3	0.7	**7e2**	10
Union Square	300	789	**3.2**	2e2	3.4	**9e1**	10
Roman Forum	200	1084	2.7	9e5	**0.2**	**3e0**	37
Vienna Cathedral	120	836	0.7	7e4	**0.4**	2e4	12
Piazza del Popolo	60	328	**1.6**	**9e1**	2.2	2e2	16
NYC Library	130	332	**0.2**	8e1	0.4	**1e0**	1.4
Alamo	70	577	**0.2**	**7e5**	0.3	2e7	2.4
Metropolis	200	341	0.6	**3e1**	**0.5**	7e1	18
Yorkminster	150	437	0.4	9e3	**0.1**	**5e2**	6.7
Montreal N.D.	30	450	**0.1**	**4e-1**	0.4	1e0	9.8
Tower of London	300	572	**0.2**	3e4	1.0	**4e1**	44
Ellis Island	180	227	0.3	3e0	0.3	3e0	8.0
Notre Dame	300	553	**0.8**	7e4	1.9	**7e0**	2.1

Dataset sizes are given in both meters and number of cameras. The table shows median and mean camera error.

We significantly outperform an existing method [4]. 1DSfM often results in a similar median error, but a greatly improved average. Runtimes are 3-12x faster than [5].

References

[1] Chatterjee, A., Govindu, V.M. Efficient and robust large-scale rotation averaging. ICCV 2013.
[2] Eades, P., Lin, X., Smyth,W.F. A fast and effective heuristic for the feedback arc set problem. Information Processing Letters (1993).
[3] Agarwal, S., Mierle, K., Others. Ceres solver. `https://code.google.com/p/ceres-solver/`
[4] Govindu, V.M. Combining two-view constraints for motion estimation. CVPR 2001.
[5] Snavely, N., Seitz, S., Szeliski, R.: Photo tourism: Exploring photo collections in 3D. SIGGRAPH 2006.

All Results

Vienna Cathedral, Montreal Notre Dame, Alamo, Piazza del Popolo, NYC Library, Metropolis, Ellis Island, Notre Dame, Tower of London, Yorkminster, Roman Forum, Union Square, Piccadilly (2152 images), Trafalgar (4597 images)