Tensor-based Synchronization and the Low-rankness of the Block Trifocal Tensor

Daniel Miao*, Gilad Lerman *, and Joe Kileel †

* School of Mathematics, University of Minnesota.

† Department of Mathematics and Oden Institute for Computational Engineering and Sciences, University of Texas at Austin.

Introduction

From 2D images, we want to obtain a 3D model of the structure of interest and the camera poses.

Figure: Structure from Motion (from www.cs.cornell.edu/∼snavely/bundler/).

This is a crucial technology for modern advancements, such as autonomous vehicles, large scale scene reconstruction, etc.

Introduction

In this work, we develop a global synchronization algorithm using trifocal tensors. We are able to capture more complex geometric information compared to traditional approaches that are encoded in higher order measurements, and improve reconstruction quality. Specifically,

- We simplify the location synchronization problem. Pairwise measurements can only encode the relative direction of camera locations, yet trifocal tensors can determine the relative scales of directions, including in the collinear case.
- We enable structure from motion algorithms in textureless scenes (insufficient points), as trifocal tensors not only relate points across three views, but also lines, or a mixture of lines and points.

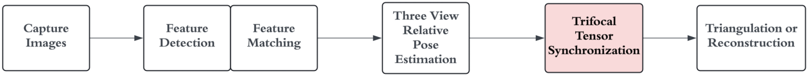


Figure: Trifocal Tensor based Pipeline

The Block Trifocal Tensor

- Let $\{P_i\}_{i=1}^n$ with $n \geq 3$ be camera matrices.
- Define the *block trifocal tensor* T^n to be the $3n \times 3n \times 3n$ tensor, where the $3 \times 3 \times 3$ sized *ijk* block is the trifocal tensor calculated with the canonical equation corresponding to the triplet of cameras P_i, P_j, P_k.

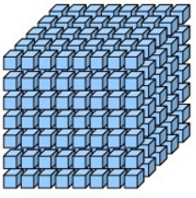
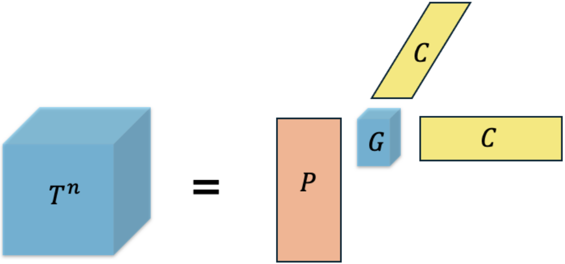


Figure: Block Trifocal Tensor Configuration. Each block is a trifocal tensor. (modified from wikidocs.net/52460)

Properties T^n

We mainly establish the following theoretical property of the block trifocal tensor T^n. Additional properties can be found in the associated paper.

- **Tucker Factorization:** The block trifocal tensor T^n admits a tucker factorization, $T^n = \mathcal{G} \times_1 \mathcal{P} \times_2 \mathcal{C} \times_3 \mathcal{C}$, where $\mathcal{G} \in \mathbb{R}^{6\times4\times4}$ is a sparse tensor, $\mathcal{P} \in \mathbb{R}^{3n\times6}$ are n line projection matrices, and $\mathcal{C} \in \mathbb{R}^{3n\times4}$ are n camera matrices.
  Low Multilinear Rank: If the n cameras that produce T^n are not all collinear, then $mlrank(T^n) = (6, 4, 4)$.



Theoretical Guarantee of Constraint

Retrieve Poses: The cameras can be retrieved from T^n up to a global projective ambiguity using the higher-order SVD. The cameras will be the leading 4 singular vectors of $T^n_{(2)}$ or $T^n_{(3)}$.

In practice, trifocal tensors can only be measured up to an unknown scale. We establish the theoretical tightness of our low multilinear rank constraint, which becomes the guarantee for our algorithm.

- **Theoretical Guarantee:** $\lambda \in \mathbb{R}^{n\times n\times n}$. Assume that $\lambda \odot_b T^n \in \mathbb{R}^{3n\times 3n\times 3n}$ has multilinear rank $(6,4,4)$ where $\odot_b$ denotes blockwise scalar multiplication. Then there exist $\alpha, \beta, \gamma \in \mathbb{R}^n$ such that $\lambda_{ijk} = \alpha_i \beta_j \gamma_k$ whenever i, j, k are not all the same.

In other words, as long as we can find a set of scales such that the block tensor has the multilinear rank of $(6,4,4)$, we can obtain the camera matrices.

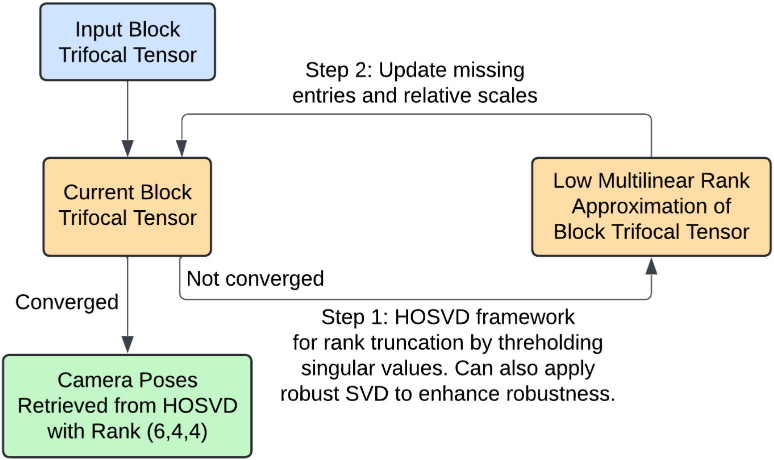
A Heuristic Approach using HOSVD

We develop a synchronization framework to optimize the block trifocal tensor and retrieve camera orientation and location using HOSVD. There are 2 practical challenges.

- **Missing Measurements**: many trifocal tensor measurements may be missing.
- **Scales Retrieval**: trifocal tensors are measured up to an unknown scale, yet our theorem relies on having a correct set of scales.

Algorithm Solution: Heuristically, we use the information of the relative scales and entries in a multilinear rank truncated tensor to synchronize the estimated block tensor.

TrifocalSync



Numerical Results for EPFL

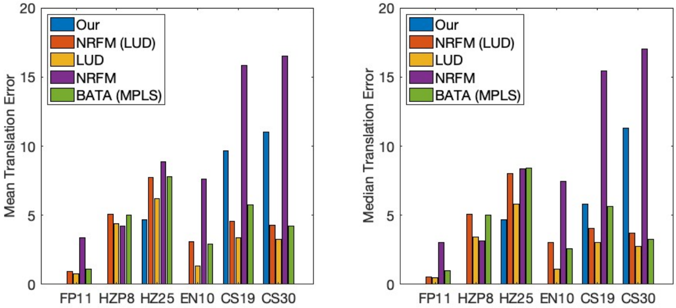


Figure: EPFL translation error comparison between our method, NRFM initialized by LUD, LUD, and NRFM initialized randomly. BATA(MPLS) stands for BATA initialized by MPLS.

Median Translation Errors for PhotoTourism

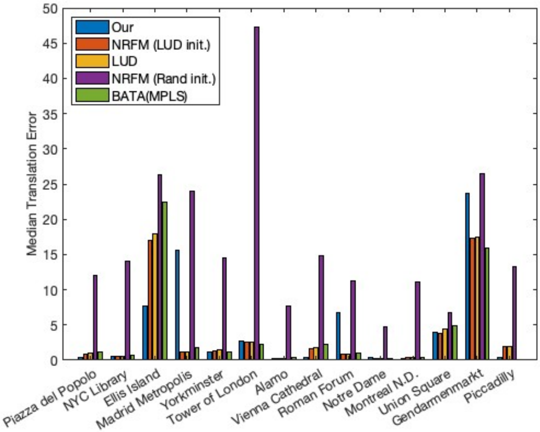


Figure: Median Translation Errors

Mean Translation Errors for PhotoTourism

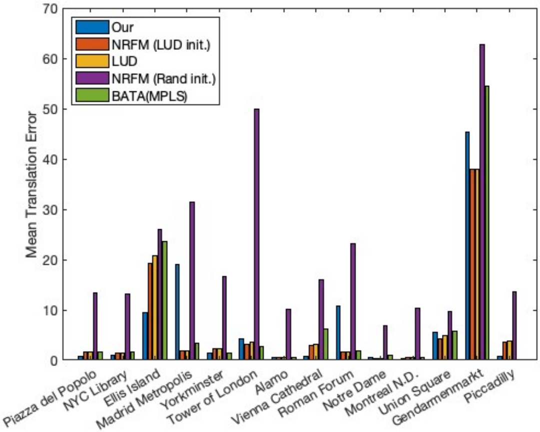


Figure: Mean Translation Error

Conclusion

Paper Summary:

- We introduce the Block Trifocal Tensor and establish the Low Multilinear Rank of $(6, 4, 4)$.
- We develop an algorithmic framework to synchronize the Block Trifocal Tensor.

Advantages of TrifocalSync:

- To the best of our knowledge, this is the first global synchronization algorithm operating directly on trifocal tensors.
- Theoretical guarantee for the tightness of the constraint.
- Strong potential of improving location synchronization quality.