A New Rank Constraint on Multi-view Fundamental Matrices, and its Application to Camera Location Recovery

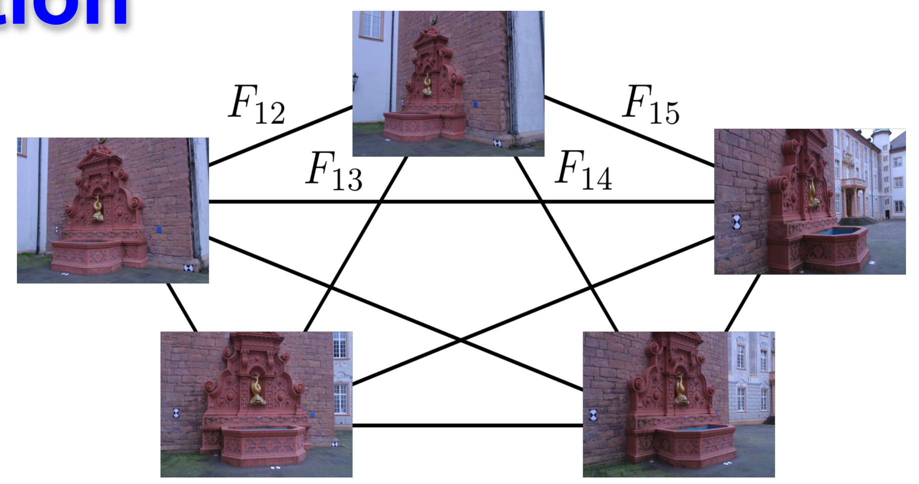
S. Sengupta, T. Amir, M. Galun, A. Singer, T. Goldstein, D. Jacobs, R. Basri

University of Maryland · Weizmann Institute of Science

IEEE 2017 Conference on Computer Vision and Pattern Recognition

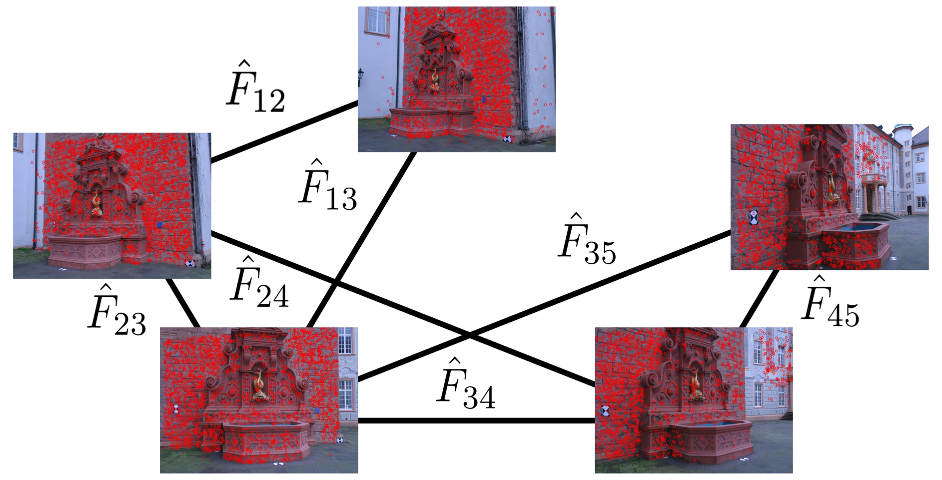
CVPR 2017, July 21-26, Honolulu

Our Contribution



- A **novel rank constraint** on Fundamental matrices in multi-view settings
- **Constrained low rank optimization** to **de-noise and recover missing** fundamental matrices
- **Improvement over state-of-the-art** in structure from motion

Our Result



$$\underbrace{\begin{bmatrix} 0 & \hat{F}_{12} & \hat{F}_{13} & - & - \\ \hat{F}_{21} & 0 & \hat{F}_{23} & \hat{F}_{24} & - \\ \hat{F}_{31} & \hat{F}_{32} & 0 & \hat{F}_{34} & \hat{F}_{35} \\ - & \hat{F}_{42} & \hat{F}_{43} & 0 & \hat{F}_{45} \\ - & - & \hat{F}_{53} & \hat{F}_{54} & 0 \end{bmatrix}}_{\hat{F}} \approx \underbrace{\begin{bmatrix} 0 & \lambda_{12} & \lambda_{13} & \lambda_{14} & \lambda_{15} \\ \lambda_{21} & 0 & \lambda_{23} & \lambda_{24} & \lambda_{25} \\ \lambda_{31} & \lambda_{32} & 0 & \lambda_{34} & \lambda_{35} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & 0 & \lambda_{45} \\ \lambda_{51} & \lambda_{52} & \lambda_{53} & \lambda_{54} & 0 \end{bmatrix}}_{\Lambda} \odot \underbrace{\begin{bmatrix} 0 & F_{12} & F_{13} & F_{14} & F_{15} \\ F_{21} & 0 & F_{23} & F_{24} & F_{25} \\ F_{31} & F_{32} & 0 & F_{34} & F_{35} \\ F_{41} & F_{42} & F_{43} & 0 & F_{45} \\ F_{51} & F_{52} & F_{53} & F_{54} & 0 \end{bmatrix}}_{F}$$

with $F = A + A^T$ and $rank(A) = 3$.

For n not all collinear cameras, the multi-view fundamental matrix F of size $3n \times 3n$ satisfies :

$$F_{ii} = \mathbf{0}, \quad F = A + A^T, \quad rank(A) = 3, \quad rank(F) = 6.$$

Construction

For simplicity, we show our construction for Essential matrices

$$E_{ij} = R_i^T (T_i - T_j) R_j$$

$T_i = [t_i]_\times$: Camera location, R_i: Camera orientation

Let $A_{ij} = R_i^T T_i \, R_j$ Then $E_{ij} = A_{ij} + A_{ji}^T$

$$A = \begin{bmatrix} R_1^T T_1 \\ \vdots \\ R_n^T T_n \end{bmatrix} [R_1 \dots R_n] \text{ and } E = A + A^T$$

$$\implies rank(A) \leq 3 \text{ and } rank(E) \leq 6$$

Equality is proved in the paper

Result applies also to fundamental matrices since $F_{ij} = K^{-T} E_{ij} K^{-1}$

Rank constrained Optimization

$$\min_{F, \{\lambda_{ij}\}} \frac{1}{2} \sum_{(i,j) \in \Omega} \| \hat{F}_{ij} - \lambda_{ij} F_{ij} \|_F$$

Robust L1 cost function

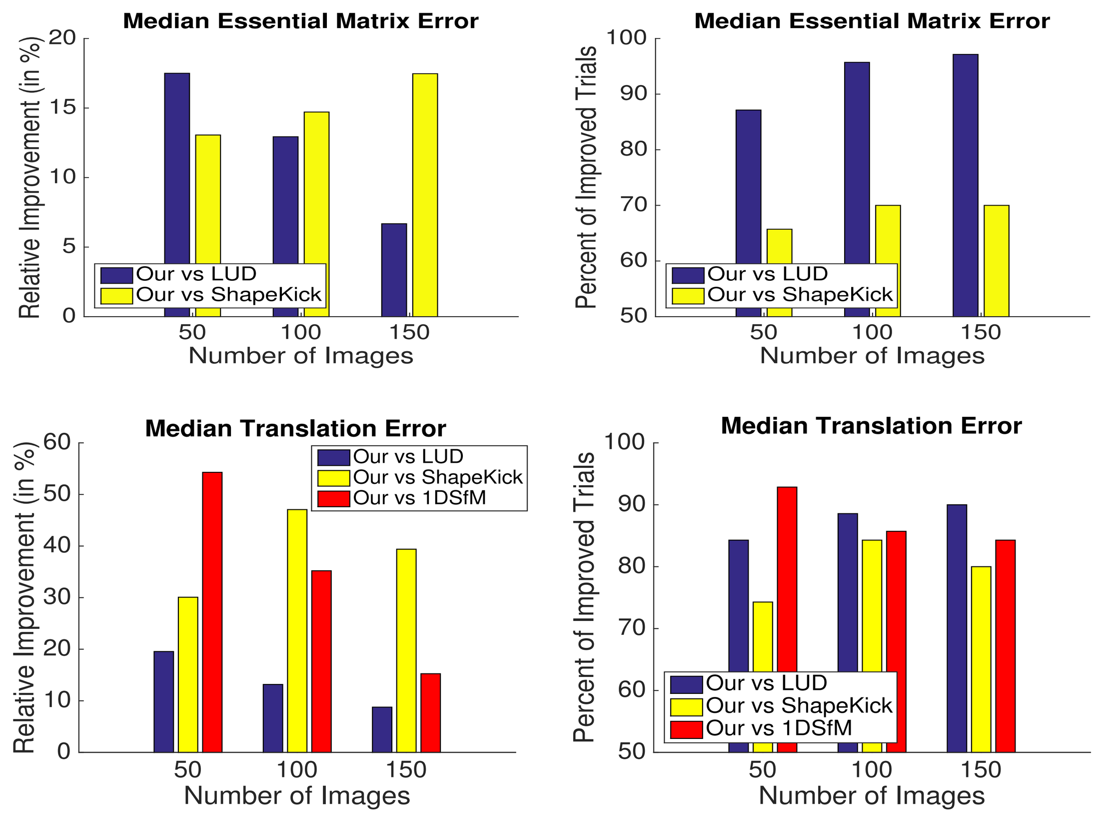
$$s.t. \; F = A + A^T, \; rank(A) = 3, \; F_{ii} = \mathbf{0}.$$

Our Rank Constraint

Solve using IRLS - ADMM

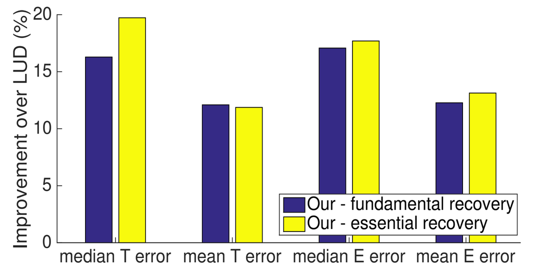
Experimental Results

- 14 scenes from Photo-tourism dataset
- For each scene, 5 different sub-samples of 50, 100 and 150 images
- Measure error on Essential matrix and camera location recovery
- Performance Measures:
  - **Relative Improvement** (in %).
  - **Number of Improved Trials.**



Essential Matrix : Improves over LUD :
- by 18% on average
- In 86% of all trials

Camera Location : Improves over LUD :
- by 20% on average
- In 84% of all trials



Our method can also optimize directly over Fundamental matrices by jointly optimizing camera rotation, location and calibration.

References

1.“Robust Global Translation with 1DSfM” K. Wilson *et.al.* ECCV 2014.(1DSfM)
2.“ Robust camera location estimation by convex programming” O. Ozyesil *et.al.* CVPR 2015. (LUD)
3. “ Shapefit and shapekick for robust, scalable structure from motion” T. Goldstein *et.al.* ECCV 2016. (ShapeKick)

Code

www.umiacs.umd.edu/~sengupta/SfM_CVPR2017_code.zip