Structure-from-Motion Revisited

Johannes L. Sch¨onberger , Jan-Michael Frahm

Group 1: Jaehoon Choi, Mingyang Xie, Akashkumar Parmar, Rishikesh Jadhav, Yishan Zhao, Haoming Cai, Aditi Bhoir

Scribe: Jaehoon Choi

Problem Settings

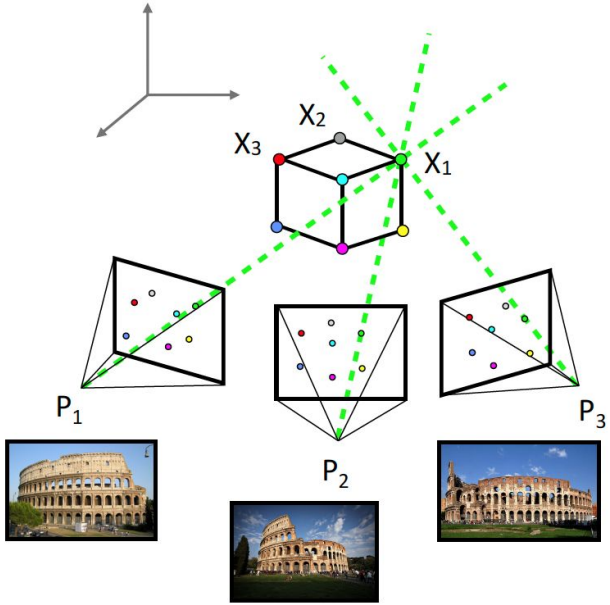
(Jaehoon)

Problem of original SfM settings

- Joint estimation of Structure Xi and Cameras Pj
- from motion, i.e. images at different viewpoints

Goal of this paper: propose a general-purpose SfM in terms of robustness and accuracy

Figure 1. Result of Rome with 21K registered out of 75K images.



Method

(Jaehoon)

- Introduce general SfM pipeline and Open-source software
- Contributions
  - Scene graph augmentation
  - Next best view selection maximizing the robustness and accuracy
  - Robust and efficient triangulation method
  - Iterative BA, re-triangulation, and outlier filtering strategy

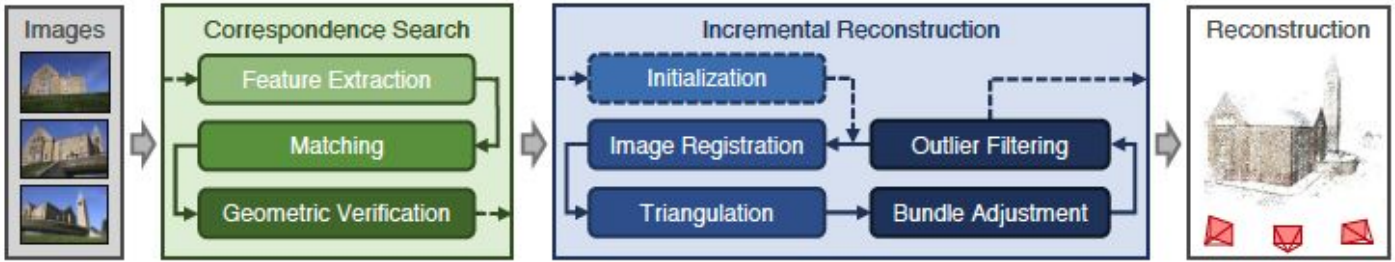


Figure 2. Incremental Structure-from-Motion pipeline.

Method - Scene graph Augmentation

(Jaehoon) propose a multi-model geometric verification strategy to augment the scene graph

- Augmented geometric verification
  - The number of inliers for the fundamental matrix N_F
  - The number of inliers for the essential matrix N_E
  - The number of inliers for the homography N_H
- Check the ratio N_E/N_F , N_H/N_E , N_H/N_F and classify the type of the two view geometry
  - If $N_E/N_F < \epsilon_{EF}$ and $N_H/N_E < \epsilon_{HE}$, then “planar or panoramic (pure rotation)”
  - If $N_E/N_F < \epsilon_{EF}$ and $N_H/N_E > \epsilon_{HE}$, then “calibrated”
  - If $N_E/N_F > \epsilon_{EF}$ and $N_H/N_F < \epsilon_{HF}$, then “planar or panoramic (pure rotation)”
  - If $N_E/N_F > \epsilon_{EF}$ and $N_H/N_F > \epsilon_{HE}$, then “uncalibrated”
- Seed for reconstruction
  - Non-panoramic
  - Calibrated image pairs
- Do not triangulate
  - Panoramic image pairs to avoid degenerate points
- Also, handles images including WTF (watermarks, timestamps, and frames)
  - Estimate a translation transformation with N_T inliers at the image borders
  - Any image pair with $N_T/N_F < \epsilon_{TF}$ is considered a WTF and not inserted to the scene graph

Method - Robust and Efficient Triangulation

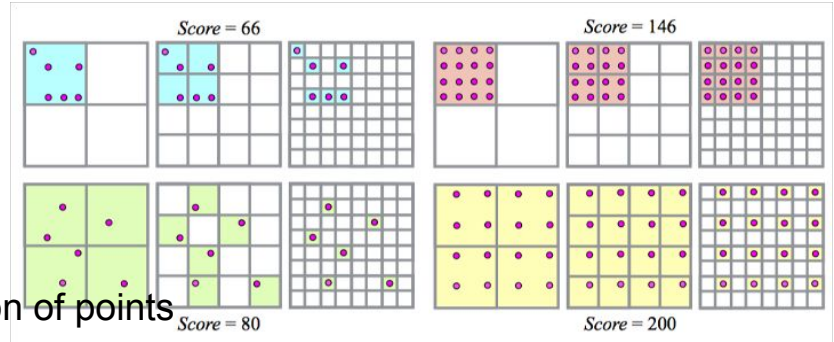
(Jaehoon)

Motivation

- Choosing the next best view is critical, as every decision impacts the remaining reconstruction.
- A single bad decision may lead to a cascade of camera mis-registrations.

Diverse strategies

- MAX_VISIBLE_POINTS_NUM
  - To choose the image that sees most triangulated points
- MAX_VISIBLE_POINTS_RATIO
  - Higher ratio of visible points on observations
- MIN_UNCERTAINTY
  - More visible points and a more uniform distribution of points



More visible points and a more uniform distribution of points: **higher score**

Robust and Efficient Triangulation

(Jaehoon)
Propose an efficient, sampling-based triangulation method that can robustly estimate all points within an outlier-contaminated feature track

- Refinement using multiple view triangulation
- Cheirality constraint
  - Check positive depth w.r.t the camera views
- Sufficient triangulation angle
  - Angle between two rays should be bigger enough

Bundle Adjustment $\min_{P,X} \|x - \pi(P,X)\|$

(Jaehoon) perform local BA on the set of most-connected images after each image registration. perform global BA only after growing the model by a certain percentage.

- Before BA: Re-Triangulation
  - Continuing the tracks of points that previously failed to triangulate to improve the completeness of the reconstruction
  - Continue tracks with observations whose errors are below the filtering thresholds
- After BA: Filtering
  - Filter observations with large reprojection errors
- Iterative refinement
  - Perform Re-triangulation, BA, and filtering in an iterative optimization until the number of filtered observations and post-BA RT points diminishes.

Weakness

(Jaehoon)

- Inherent scale ambiguity of SfM
- Dynamic Objects
- Repetitive Structures
- Illumination Change (e.g. Day-night matching difficult)