Project Presentation: CS676
(Computer Vision)

3D RECONSTRUCTION FROM SEVERAL IMAGES(STRUCTURE FROM MOTION)

Anshu Avinash
10327122
anshuavi@iitk.ac.in

Pranjal Singh
10327511
spranjal@iitk.ac.in

Sunil Kumar
10327742
msunil@iitk.ac.in

GROUP-6

Advisor: Prof. Vinay P. Namboodiri

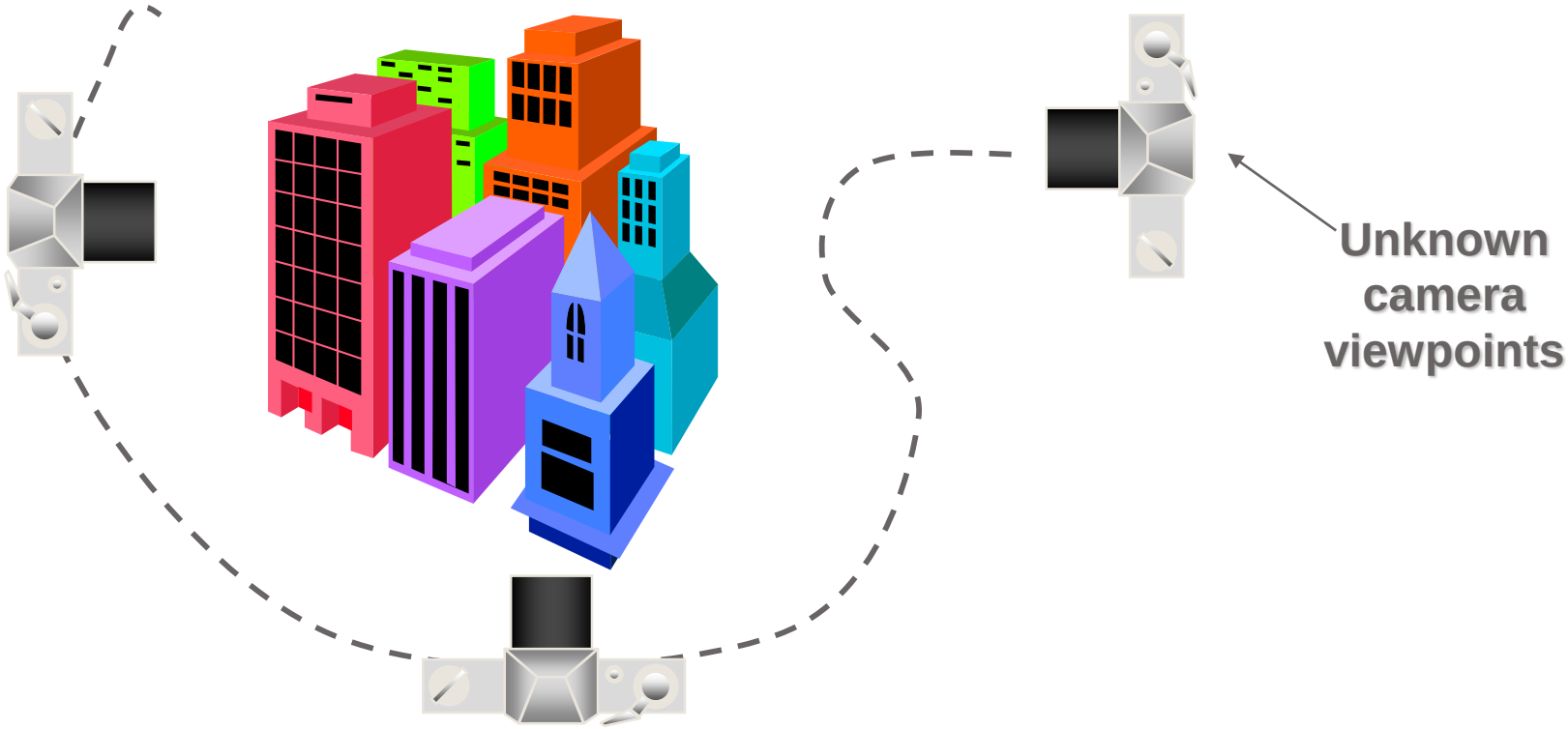
CONTENT

MOTIVATION

- We live in a 3D world
- We see 2D images but perceive the world in 3D
- Intelligent robot should have this 3D reconstruction capability

WHY SFM?

- Structure and depth are inherently ambiguous from single views

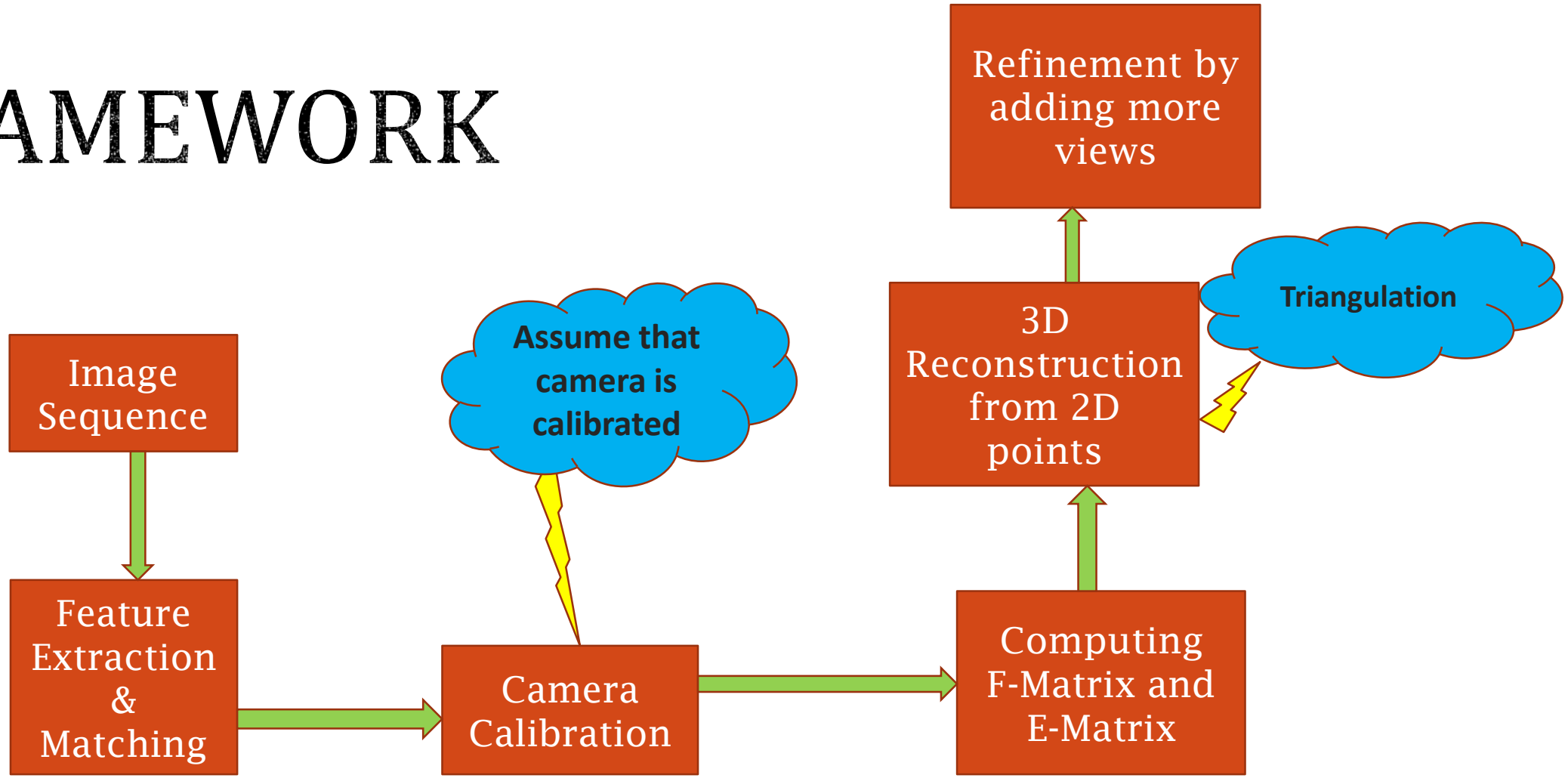


SFM

- Given many points in *correspondence* across several images, $\{(u_{ij},v_{ij})\}$, simultaneously compute the 3D location $\mathbf{x}_i$ and camera (or *motion*) parameters ($\mathbf{K}$, $\mathbf{R}_j$, $\mathbf{t}_j$)

$$\begin{aligned} \hat{u}_{ij} &= f(\mathbf{K}, \mathbf{R}_j, \mathbf{t}_j, \mathbf{x}_i) \\ \hat{v}_{ij} &= g(\mathbf{K}, \mathbf{R}_j, \mathbf{t}_j, \mathbf{x}_i) \end{aligned}$$

FRAMEWORK


Refinement by adding more views
Triangulation
3D Reconstruction from 2D points
Assume that camera is calibrated
Image Sequence
Feature Extraction & Matching
Camera Calibration
Computing F-Matrix and E-Matrix

DATASET

- **Fountain P-11 Sequence**

http://cvlabwww.epfl.ch/~strecha/multiview/denseMVS.html

- **Herz-Jesu-P8**

http://cvlabwww.epfl.ch/~strecha/multiview/herzjesu_dense.html

FEATURE CORRESPONDENCE

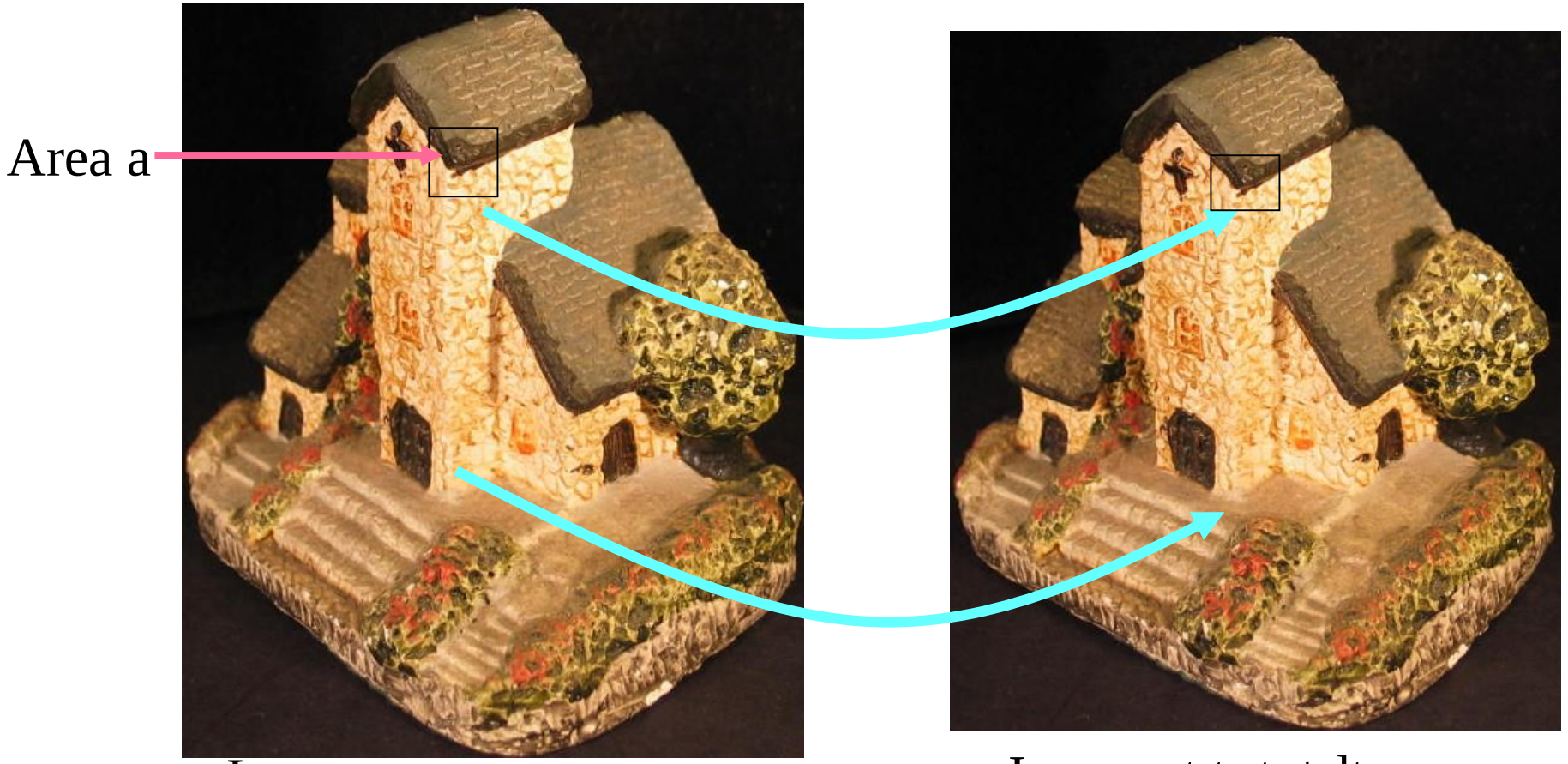


Image at $t=t_0$
(or left image)

Image at $t=t_0+dt$
(or right image)

FEATURE EXTRACTION

- SIFT Features Detection

- Feature Matching

FEATURE EXTRACTION

- Optical flow is another alternative to get better matching

“Optical flow is the process of matching selected points from one image to another, assuming both images are part of a sequence and relatively close to one another”

- **Advantage:** Process is usually faster and can accommodate matching many more points, making the reconstruction denser

CAMERA CALIBRATION

- Determine camera parameters from *known* 3D points or calibration object(s)

1. *internal* or *intrinsic* parameters such as focal length, optical center, aspect ratio:
   what kind of camera?
2. *external* or *extrinsic* (pose) parameters:
   where is the camera?

- Camera intrinsic and extrinsic parameters can be determined for a particular camera and lens combination by photographing a controlled scene.

FUNDAMENTAL MATRIX(F)

- Records motion between cameras
- Used RANSAC to calculate F
- Can be used to calculate Essential matrix which assumes usage of calibrated cameras
- $E = K^T * F * K$

ESSENTIAL MATRIX(E)

- 3 x 3 sized matrix
- Can be used to recover projection matrix for each camera
- Imposes a constraint between a point in one image and a point in the other image with **x‘ E x=0**

where x is a point in image one and x' is the corresponding point in image two.

CAMERA MATRIX

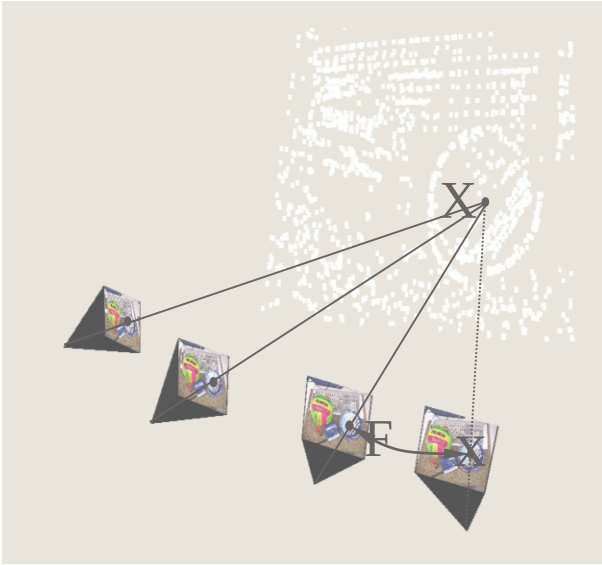
$$P = [R|t] = \begin{bmatrix} r_1 & r_2 & r_3 & t_1 \\ r_4 & r_5 & r_6 & t_2 \\ r_7 & r_8 & r_9 & t_3 \end{bmatrix}$$

Derived using SVD of Essential Matrix

RECONSTRUCTION

- Recovering the 3D structure of the scene from the information we have i.e. 2D points and camera matrices

- Triangulation Method



TRIANGULATION

- Problem: Given some points in *correspondence* across two or more images (taken from calibrated cameras), $\{(u_j,v_j)\}$, compute the 3D location **X**

- **Problem:** Rays usually do not intersect

TRIANGULATION

- **Approach**: **Linear Method**
- Two key equations arising from the 2D point matching and P matrices:

$$x=PX \text{ and } x'=P'X$$

where x and x' are matching 2D points and X is a real world 3D point

- If we rewrite the equations, we can formulate a system of linear equations that can be solved for the value of X, which is what we desire to find.

Ref: Triangulation, R.I. Hartley and P.Strum, 1997

RECONSTRUCTION FROM MANY VIEWS

- Resection or camera pose estimation, also known as **Perspective N-Point(PNP)**

- Try to solve for the position of a new camera using the scene points we have already found

- We have used **SolvePnPRansac** function available in **OpenCV**

REFINEMENT OF RECONSTRUCTION

- Known as **Bundle Adjustment**

- Both the position of the 3D points and the positions of cameras are optimized, so re-projection errors are minimized

- We have used SSBA(Simple Sparse Bundle Adjustment) Library

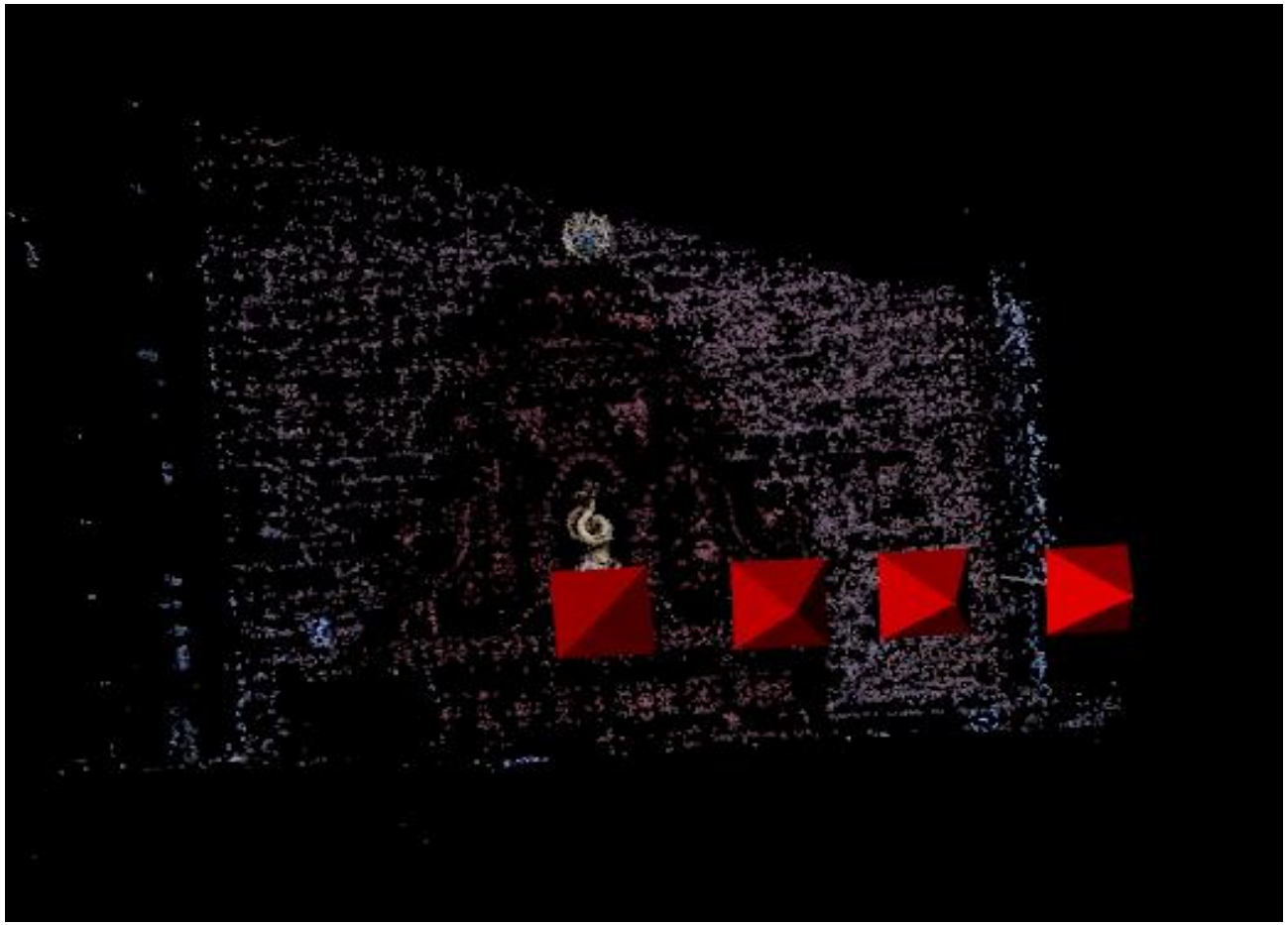
3D VISUALIZATION

- Used **Point Cloud Library(PCL)** for 3D Visualization

DATASET: FOUNTAIN P-11 SEQUENCE

RESULT

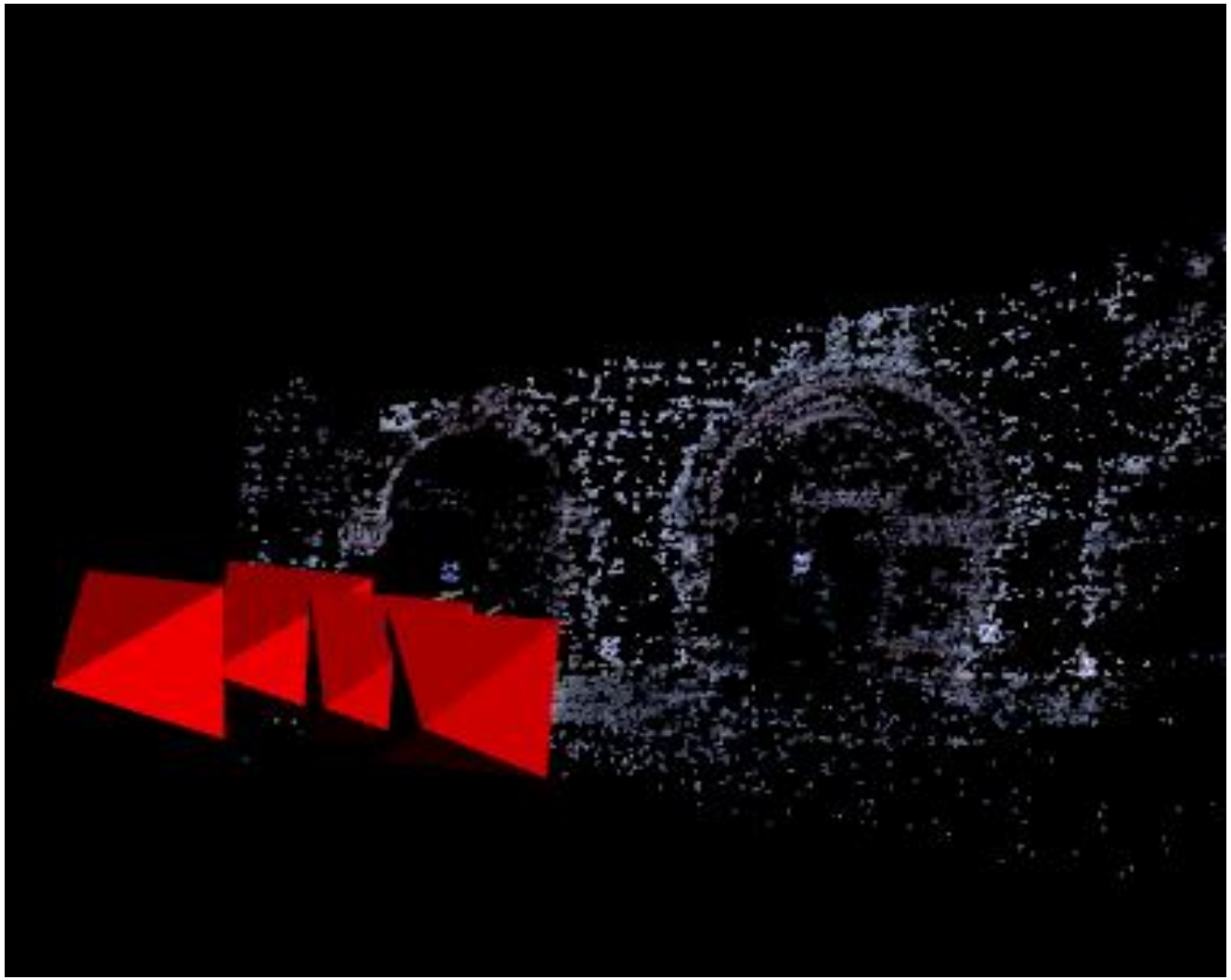
3D points recovered from the set of 4 images



DATASET: HERZ-JESU-P8

RESULT

3D points recovered from the set of 4 images



MORE WORK

- Can have better Triangulation Method that minimizes the re-projection error

- There are many state-of-the art tools available
- **E.g. Bundler, VisualSFM, LibMV**

REFERENCES

- R. I. Hartley. and A. Zisserman. Multiple View Geometry. Cambridge University Press, 2000.

- Triangulation, R.I. Hartley and P.Strum, 1997

THANK YOU!!

QUESTIONS??