3D from Photographs: Introduction

Francesco Banterle, Ph.D.
francesco.banterle@isti.cnr.it

3D from Photographs

- 3D from photographs is a technology that allows us to do a 3D reconstruction of a real-world scene starting from a set of photographs as input.
- We can see it as an alternative to 3D scanning but it has some important issues:
    - It is not MEASURING tool!
    - We cannot know the result of a 3D reconstruction beforehand.

3D from Photographs

- Advantages:
  - Fully automatic process.
  - Faster for creating models than modeling (e.g. AutoCAD, Rhinoceros, etc.).
  - Good scalability: from tiny (e.g., a toy) to large models (e.g., an entire city).
  - Unskilled users can create 3D models.
  - Economically cheap; i.e., a digital camera.

3D from Photographs

- Disadvantages:
  - Accuracy may be low; it can be improved with expensive set-ups.
  - Some real-world objects cannot be captured.
  - A generated 3D model may not match ground truth due to skew.

3D from Photographs

- The 3D model is generated using automatic Computer Vision techniques.

- The process has three main steps.

3D from Photographs

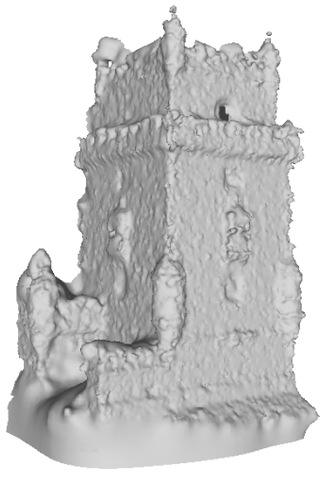
Photographs

Automatic Matching of Images

Camera Calibration

Dense Matching

Surface Reconstruction



3D model

3D from Photographs

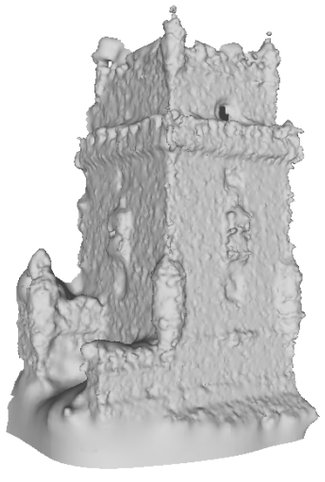
Photographs

Automatic Matching of Images

Camera Calibration

Dense Matching

Surface Reconstruction



3D model

Automatic Matching of Images

- The entire process is based on finding matches between images.

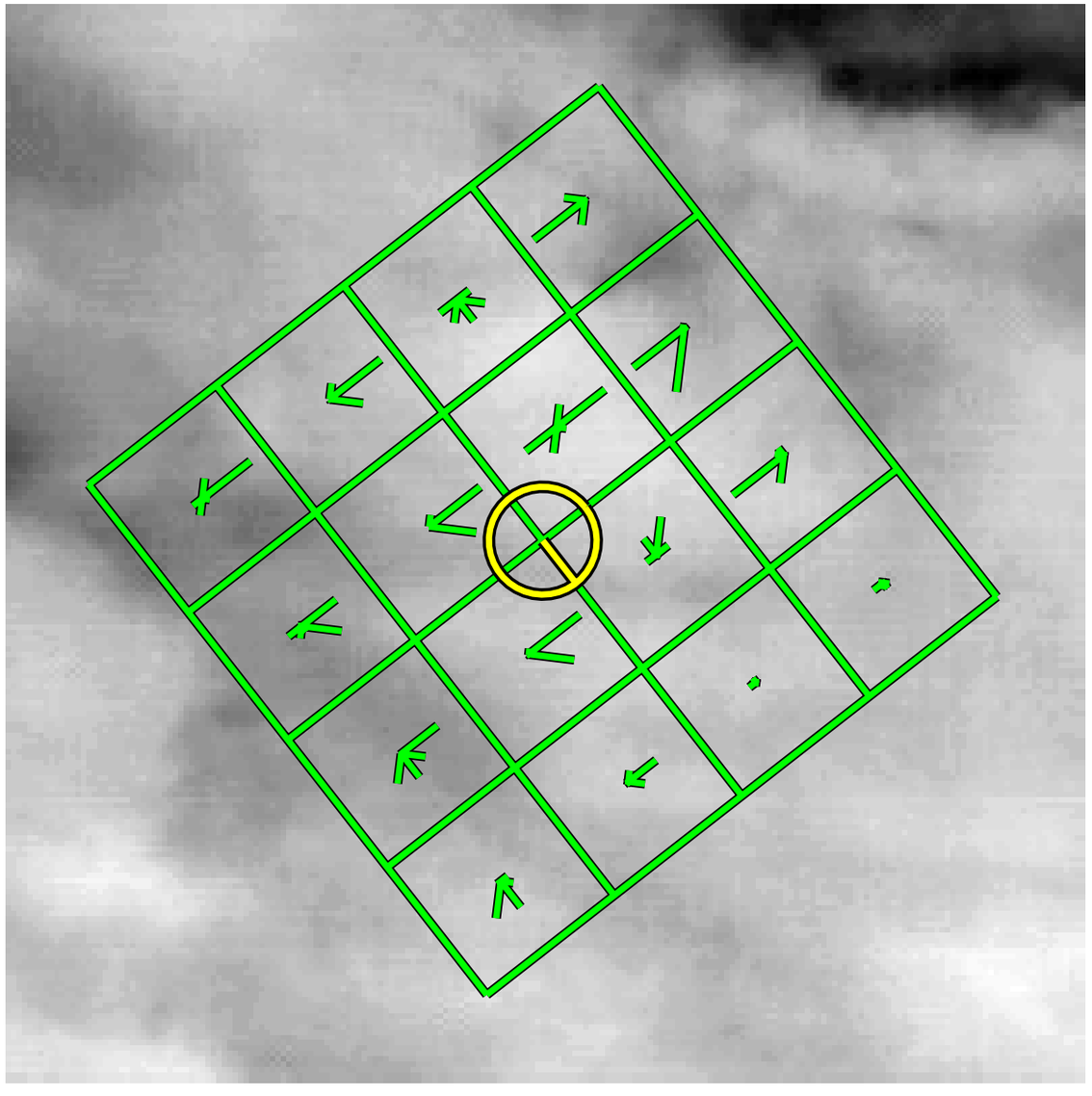
- This means that you have to shoot pictures not too far apart, so that the algorithm can match them easily.

Automatic Matching of Images

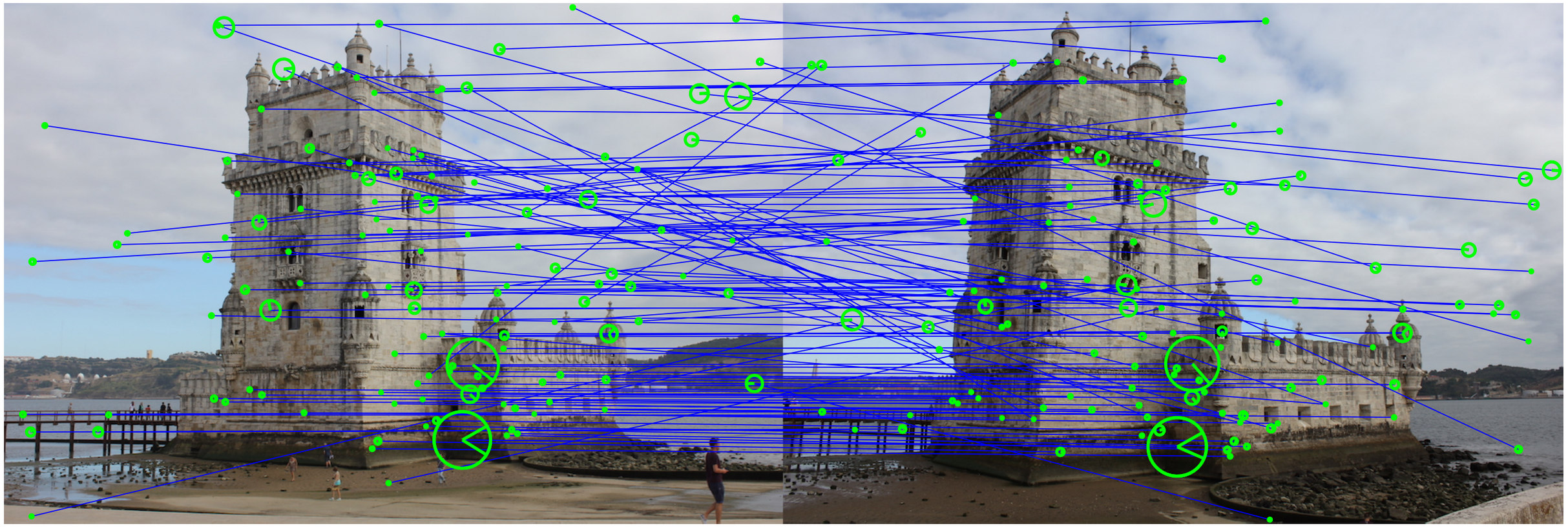
- For any object in an image, "*interesting points*" (or **corners**) on the object can be extracted to provide a "*feature description*" (or **descriptor**) of the corner.
- A descriptor of an object corner (extracted from an image) can be employed to locate the object in another image containing many other objects.

Automatic Matching of Images

Automatic Matching of Images



Automatic Matching of Images



3D from Photographs

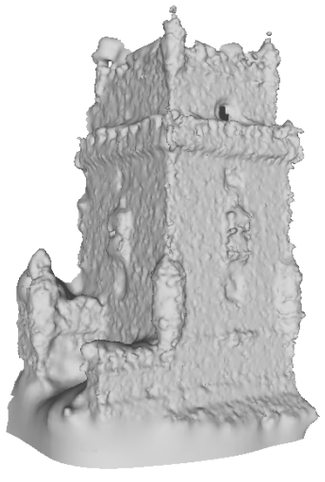
Photographs

Automatic Matching of Images

Camera Calibration

Dense Matching

Surface Reconstruction



3D model

Camera Calibration

- No prior knowledge about camera calibration is available.
  - All information must be recovered from photographs.

- It is crucial that we have enough information in photographs.

- Important factors:
  - Motion of the camera
  - General structure of the scene
  - Enough overlap: only points that are visible in at least three images are useful.
  - **Note that** what you want reconstruct and how you get the photographs have great influence on the final reconstruction!

3D from Photographs

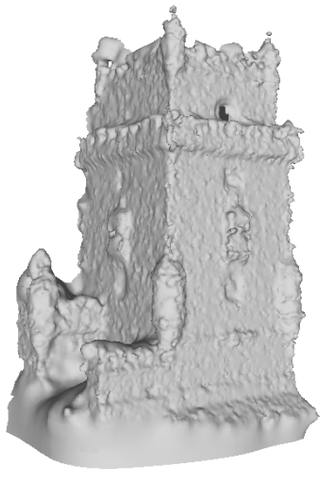
Photographs

Automatic Matching of Images

Camera Calibration

Dense Matching

Surface Reconstruction

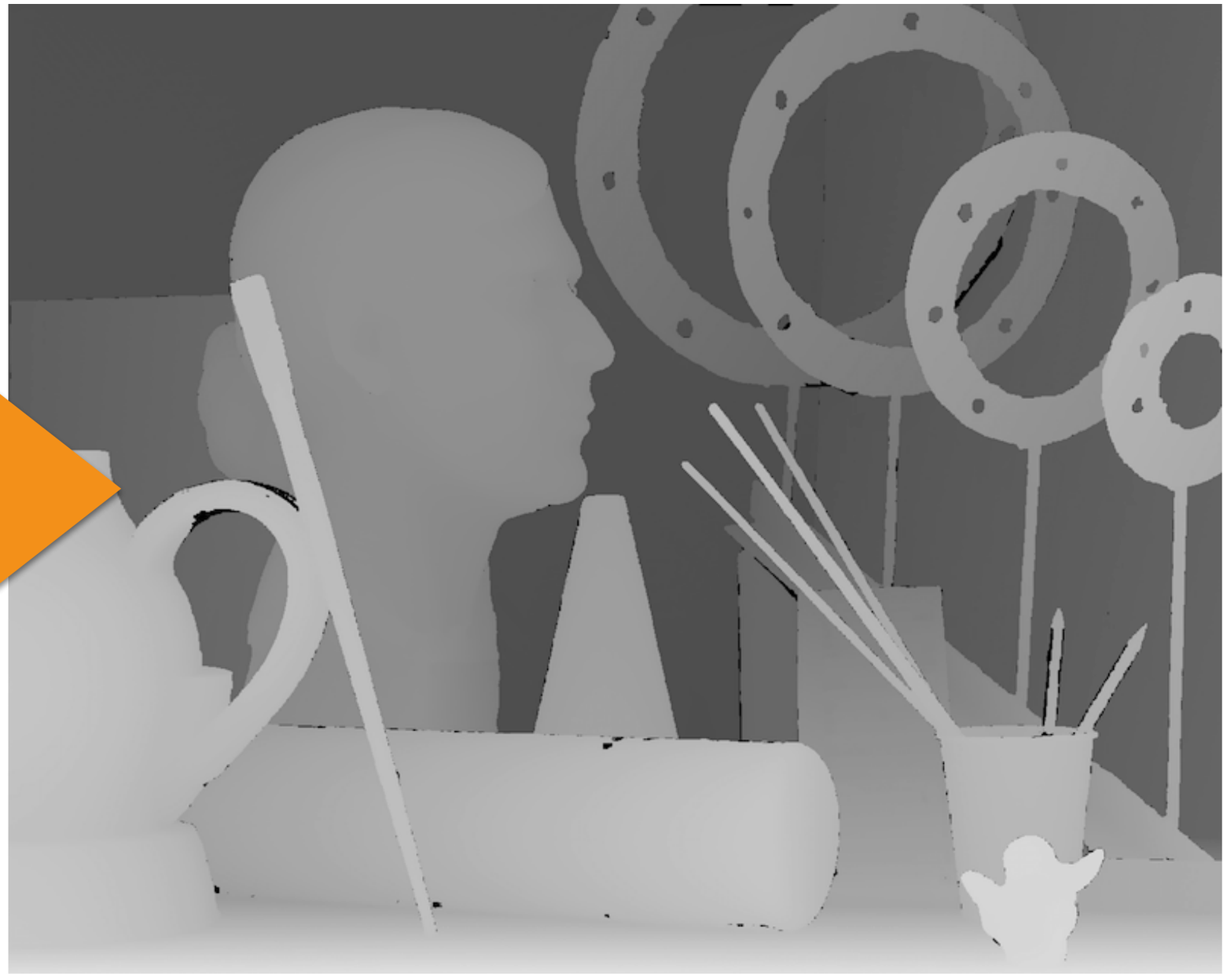


3D model

Dense Matching

- After recovery of the camera calibration, we can compute dense depth maps:
  - We need a pair of images for each depth map.
- These contain the depth of every pixel and a quality measure (how confident we are of each particular pixel).

Dense Matching



Input

Depth Map

3D from Photographs

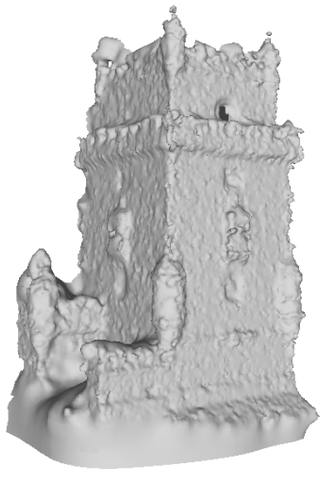
Photographs

Automatic Matching of Images

Camera Calibration

Dense Matching

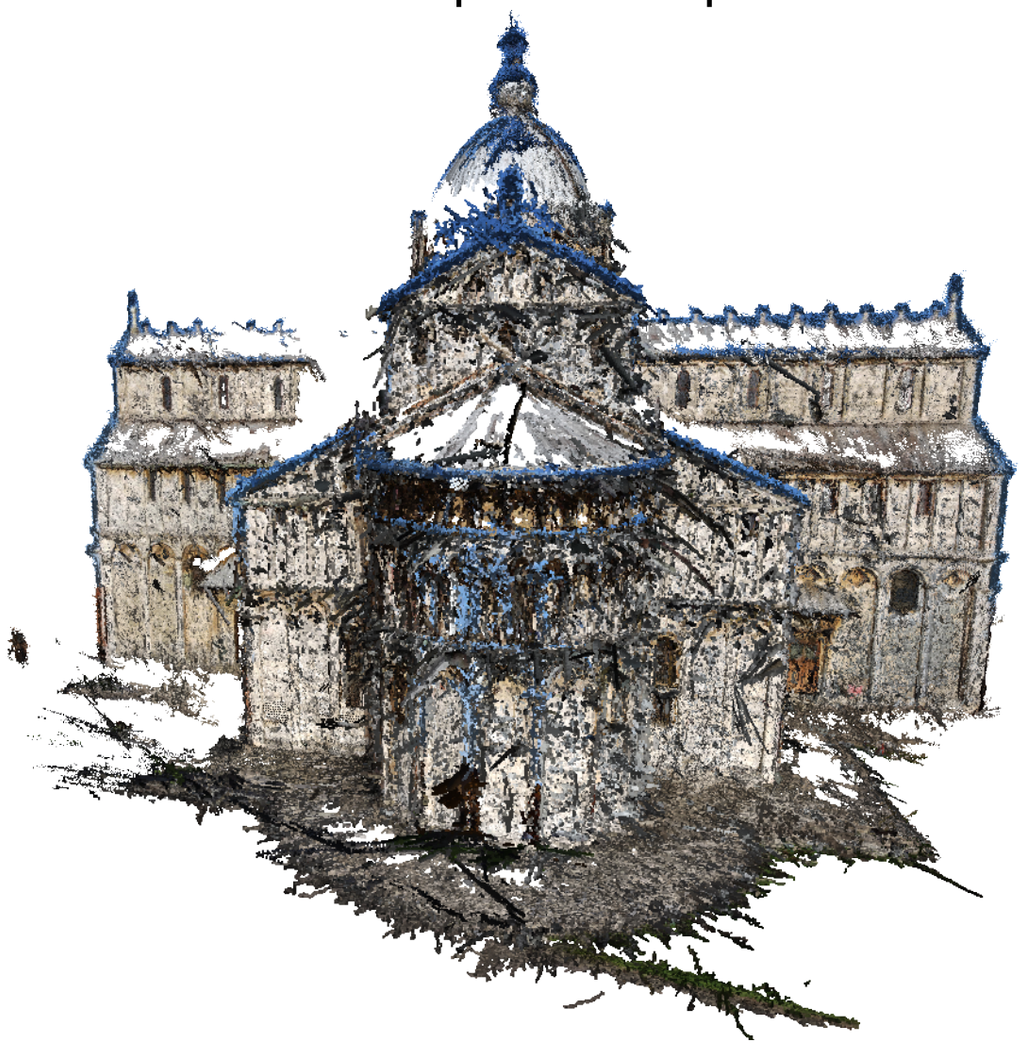
Surface Reconstruction



3D model

Surface Reconstruction

- To compute an unique 3D surface by integration of the all the depth maps of each image:



Dense Point Cloud

Final 3D Model

Photographs

Best Practice

Best Practice

- How do we shoot pictures?

- Practical suggestions and limitations to avoid failures during reconstruction.

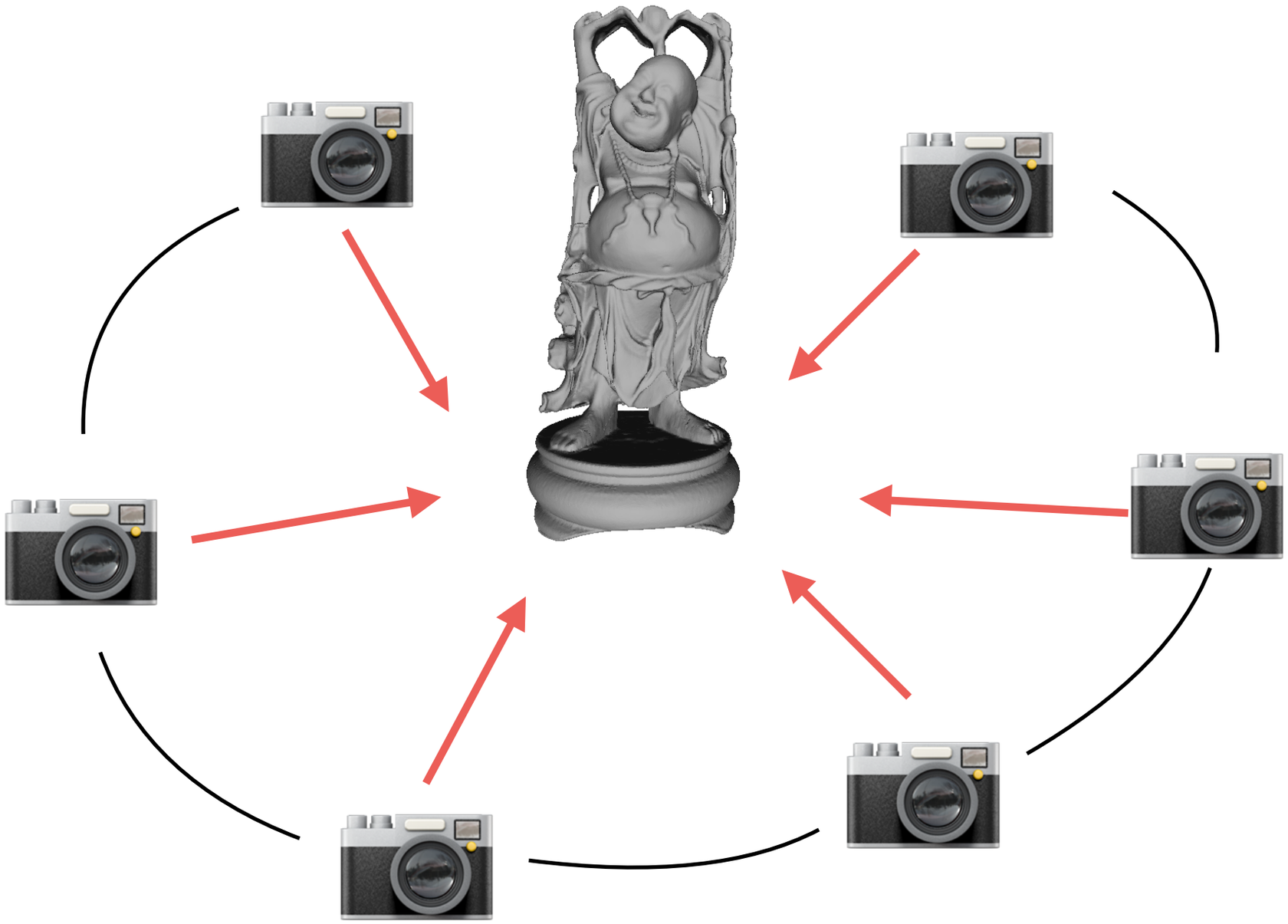
Best Practice:
A Good Sequence

- We have to shoot a picture of the same location for every step made in the shooting sequence.
    - Each picture needs to be of the same scene, but captured from a slightly different point of view.

- We have to walk with the camera in an arc around the scene and keeping the entire scene all times.

- We have to keep the same focal length; i.e., zoom!

Best Practice:
A Good Sequence



Best Practice:
A Good Sequence

- We have to capture as many photographs as we can:
  - The more the better.
  - We need at least 5-6 photographs for a very basic reconstruction!
  - A reconstruction algorithm can fail if only four photographs or less are given as input!

Best Practice:
Bad Sequences

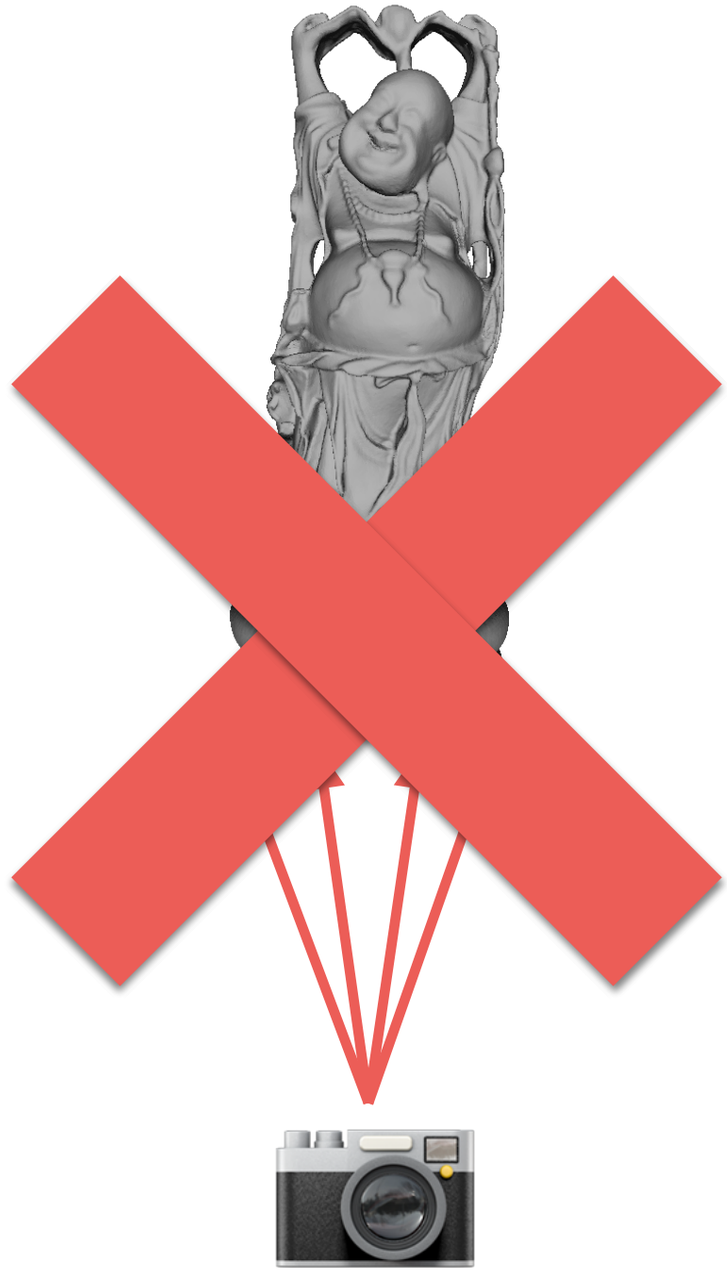
- We have to avoid “pan” sequences (panoramas sequences); i.e., capturing photographs on a plane.

- These sequences do not have 3D information.

Best Practice:
Bad Sequences

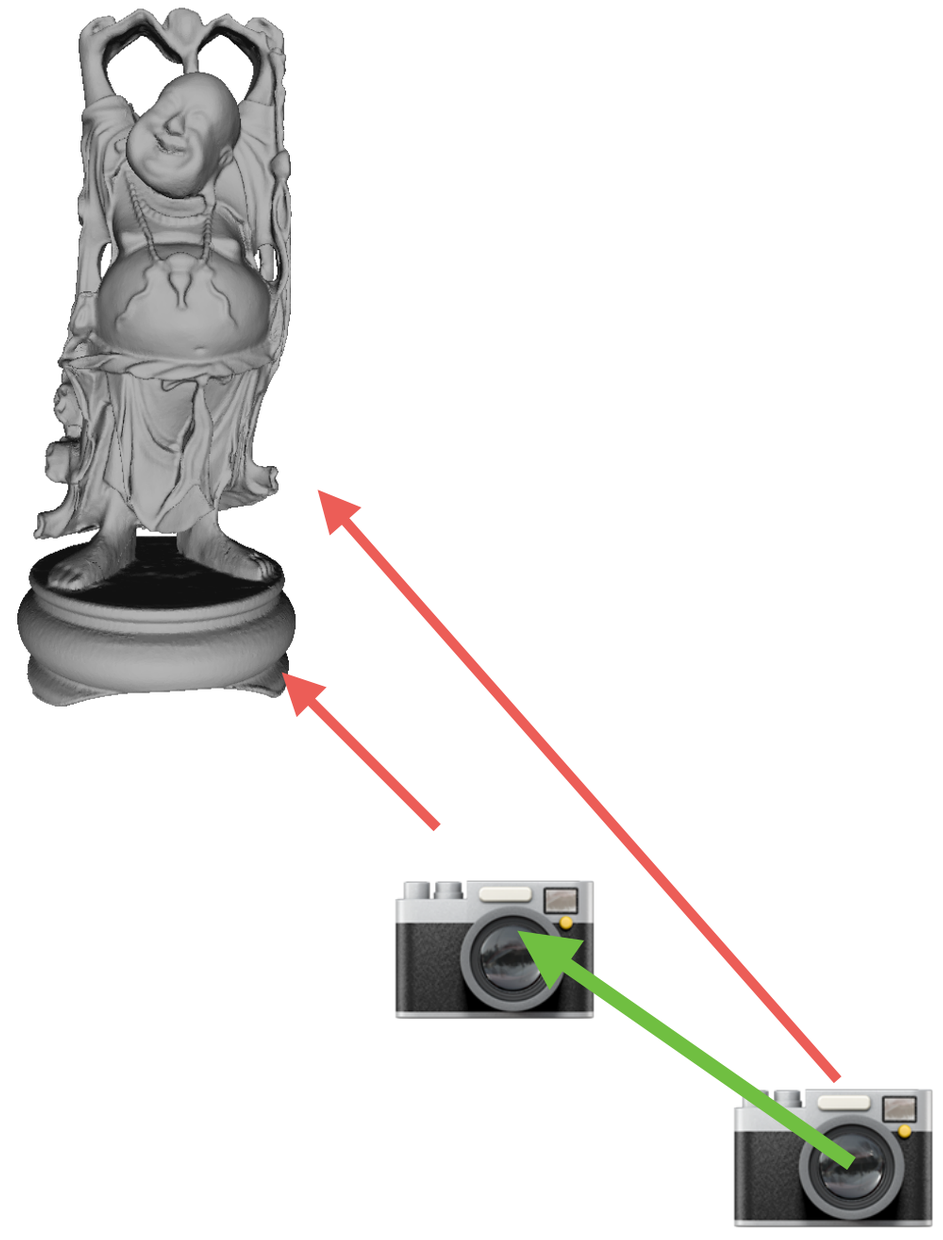
- We have to avoid “pan” sequences (panoramas sequences); i.e., capturing photographs on a plane.

- These sequences do not have 3D information.



Best Practice:
Bad Sequences

- We have to avoid photo sequences in which we shoot toward/outward the scene to capture!

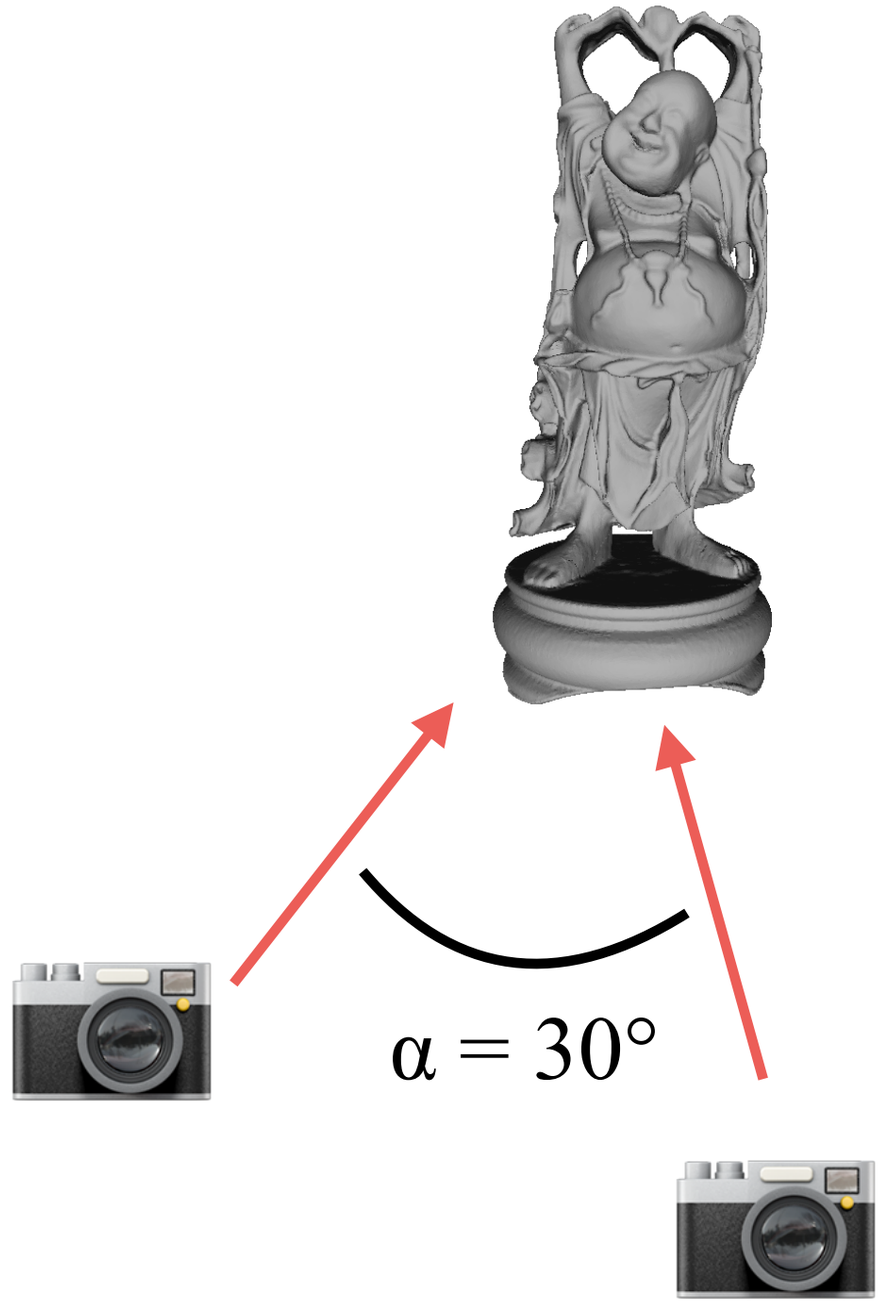


Best Practice:
Bad Sequences

- We have to avoid photo sequences in which we shoot toward/outward the scene to capture!

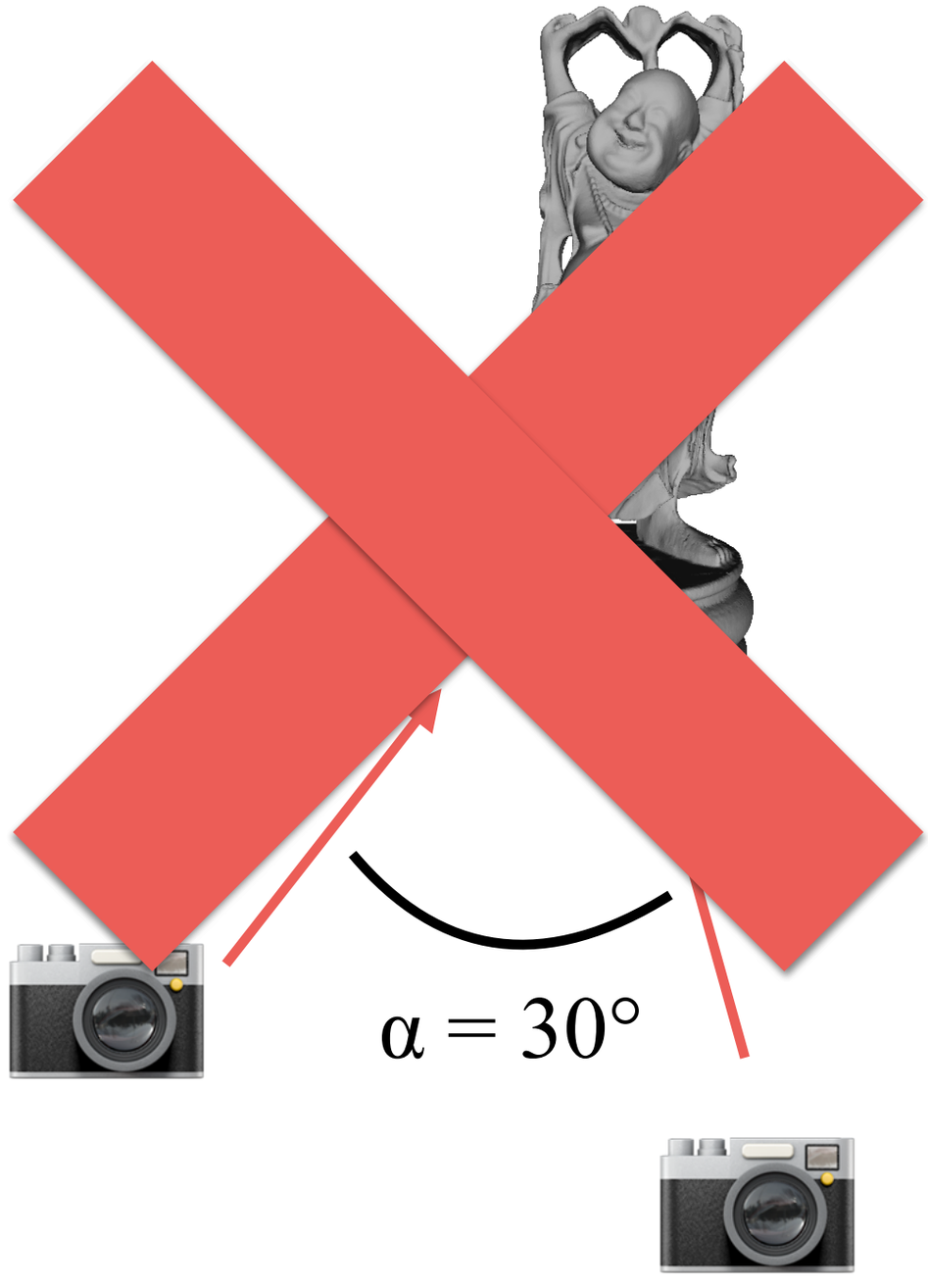
Best Practice:
Bad Sequences

- If the angle between a photograph and another is too small, the reconstruction algorithm may fail or produce low quality reconstruction!



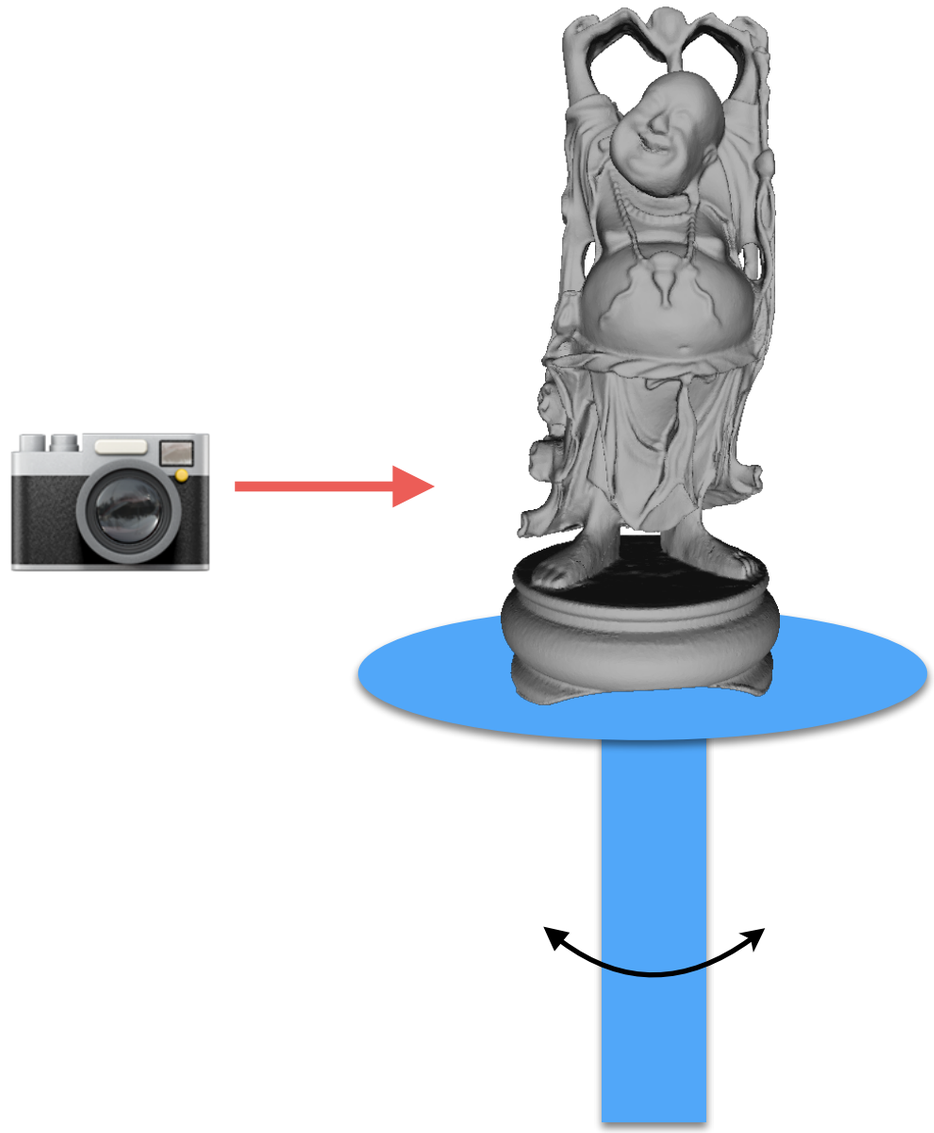
Best Practice:
Bad Sequences

- If the angle between a photograph and another is too small, the reconstruction algorithm may fail or produce low quality reconstruction!



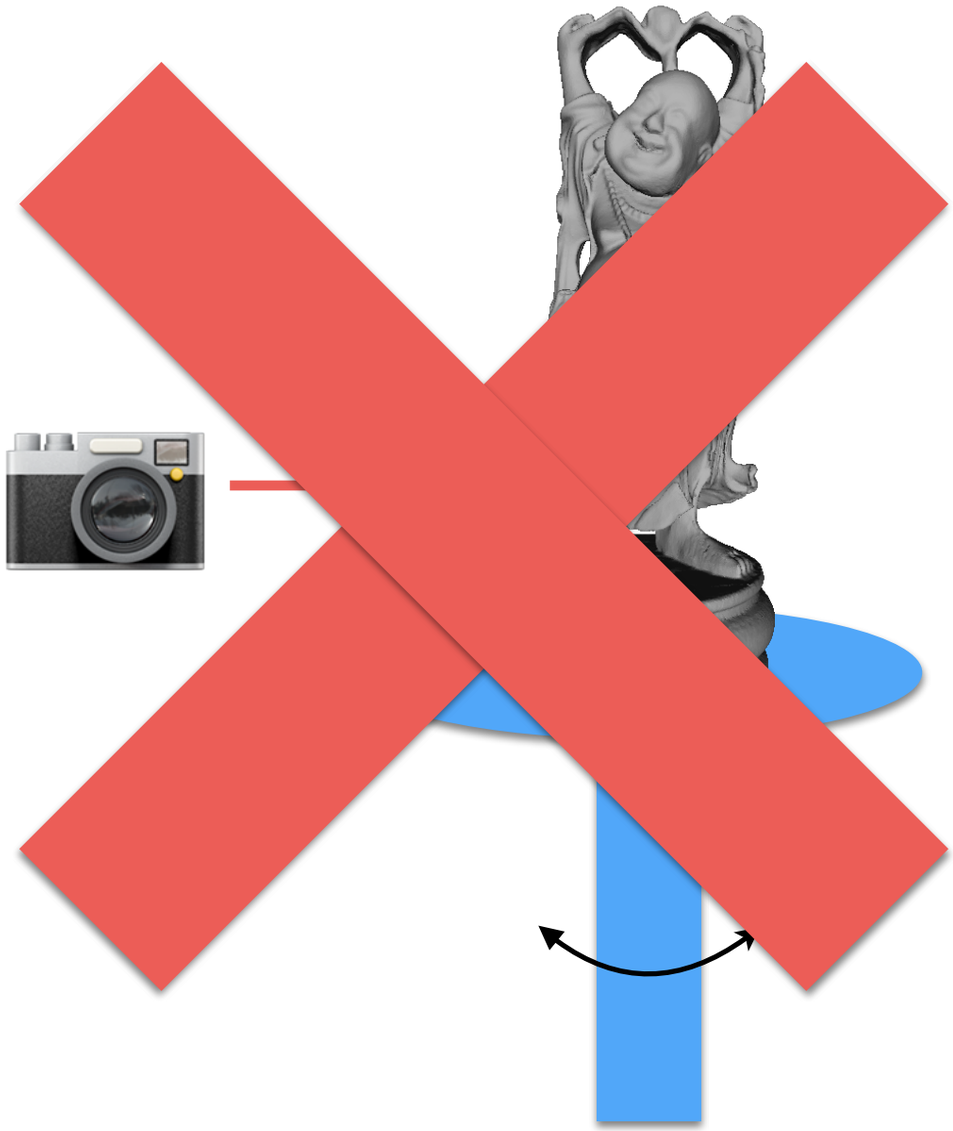
Best Practice:
Bad Sequences

- We cannot take photographs by rotating the person/object using a turnable table!!



Best Practice:
Bad Sequences

- We cannot take photographs by rotating the person/object using a turnable table!!



Best Practice: Planar Objects

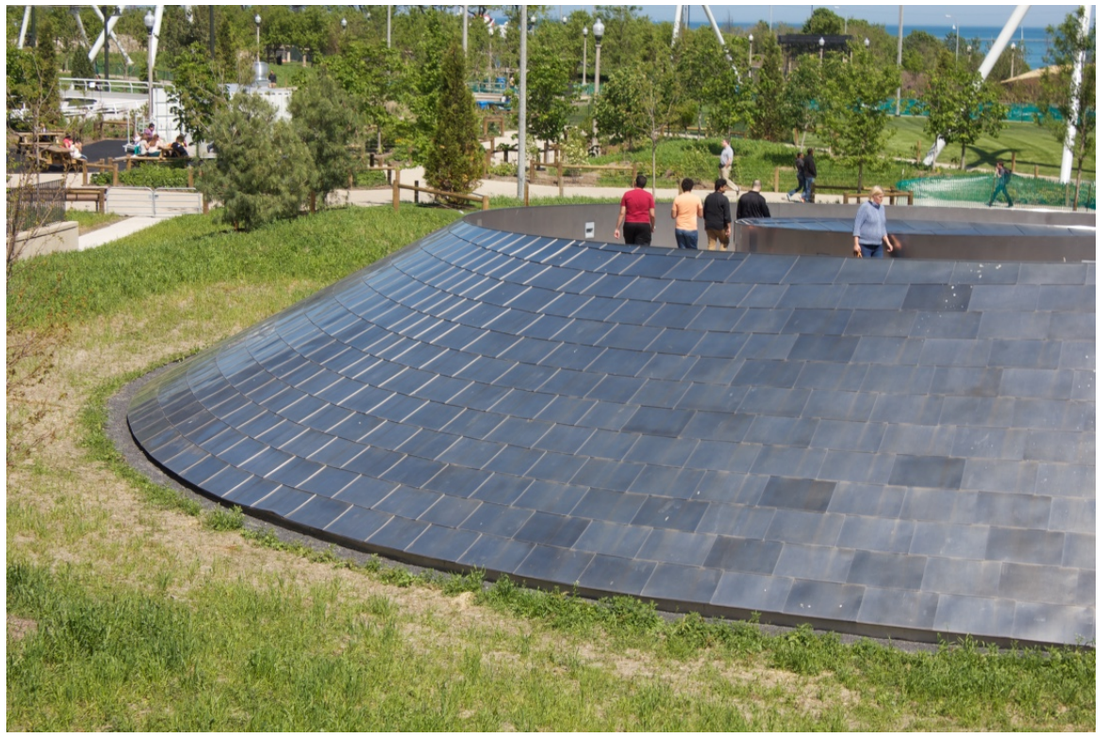
- We cannot take photographs of planar objects!

Best Practices:
Not Enough Textures

Best Practices:
Non Constant Appearance

Best Practices:
Non Constant Appearance

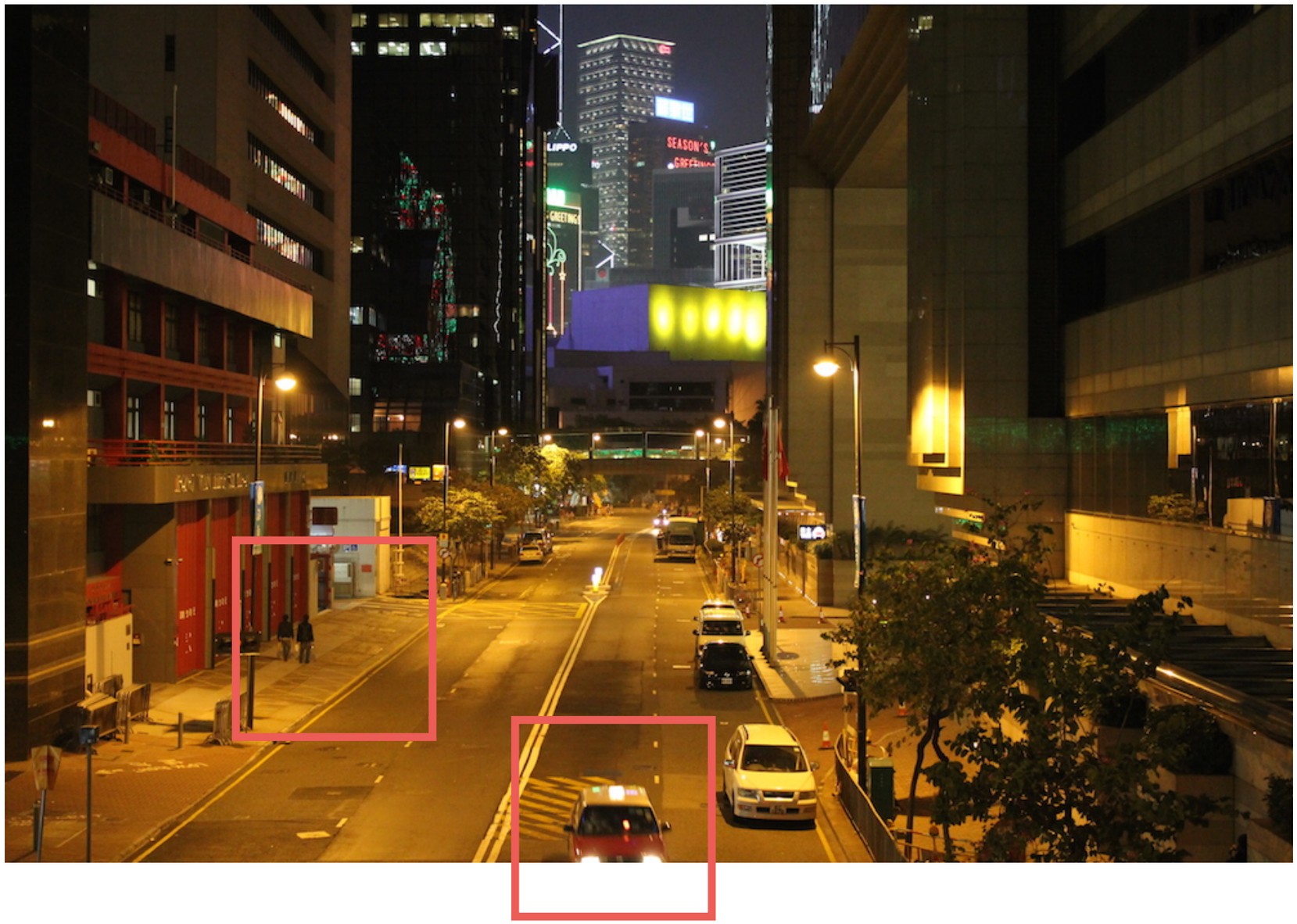
Best Practices:
Non Constant Appearance



Best Practices: Dynamic Scenes

Moving people or objects appear/disappear!

Best Practices: Dynamic Scenes

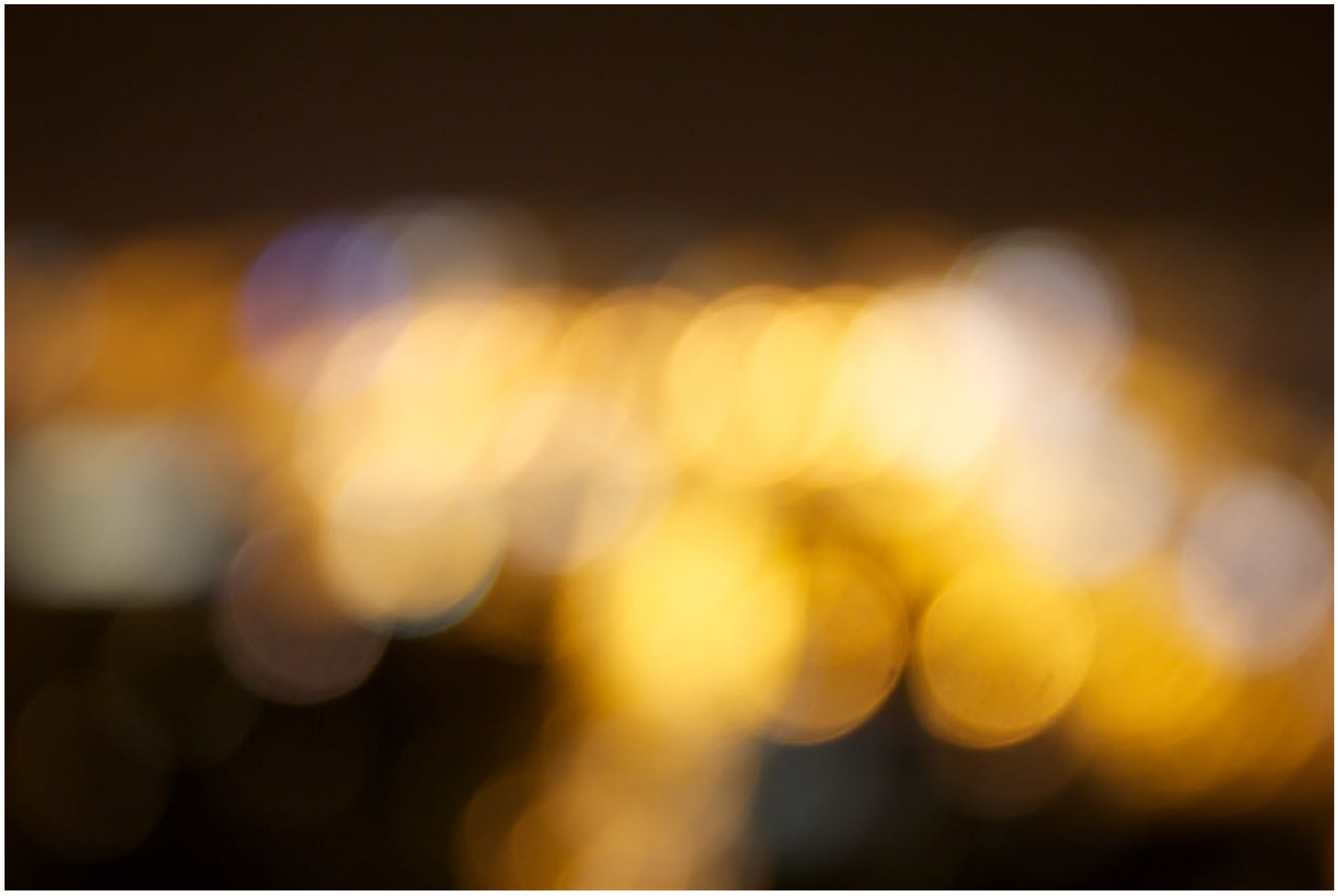


Moving people or objects appear/disappear!

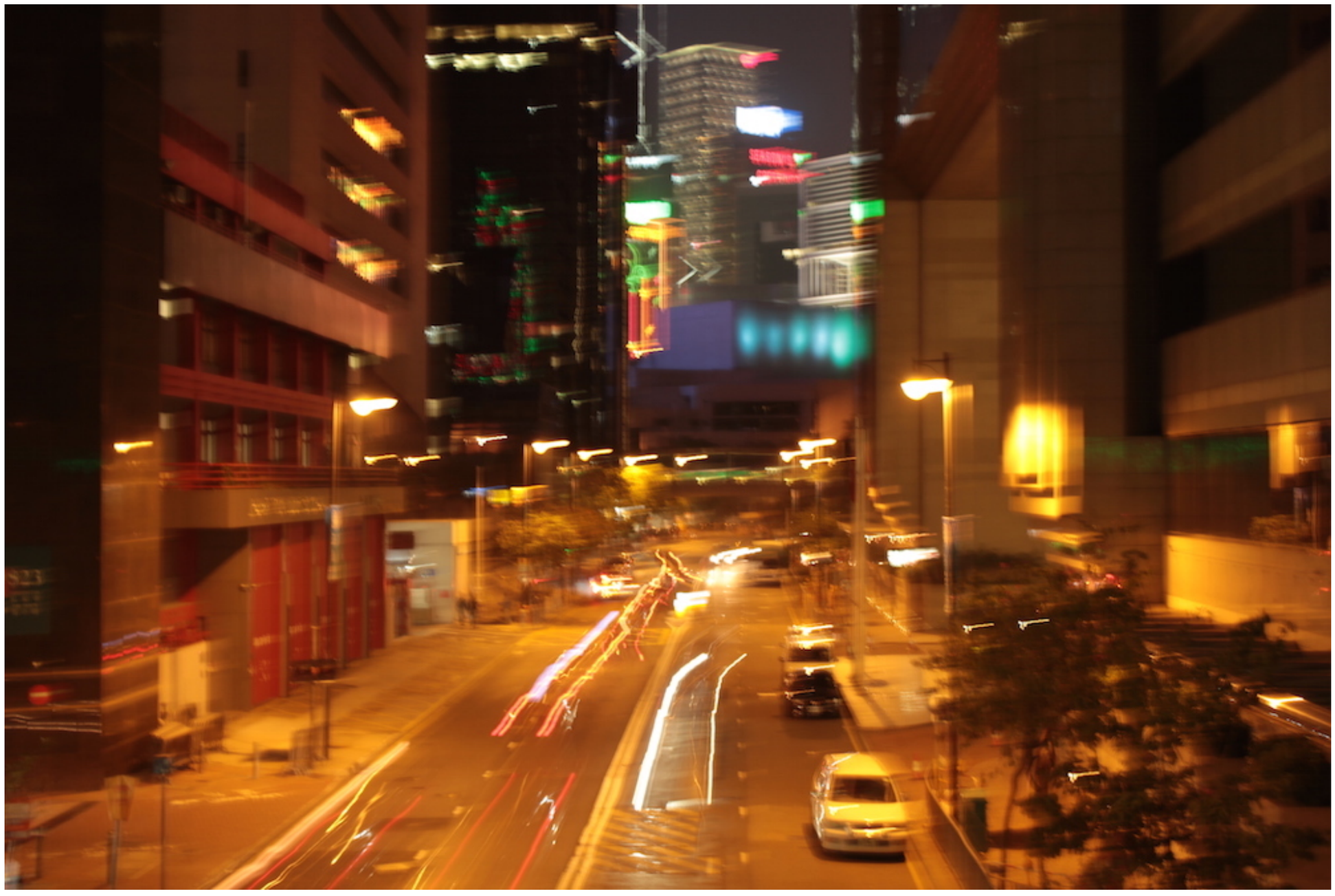
Best Practices: Blurry Photos

- Blurry photos are caused by:
    - Movements in the scenes or of the camera; i.e., motion blur.
    - Camera is out-of-focus
- These photos MUST be avoided!
- They cause issues during reconstruction and/or degrade the final result!

Blurry Photos



Blurry Photos



Best Practices: Self-Occlusions

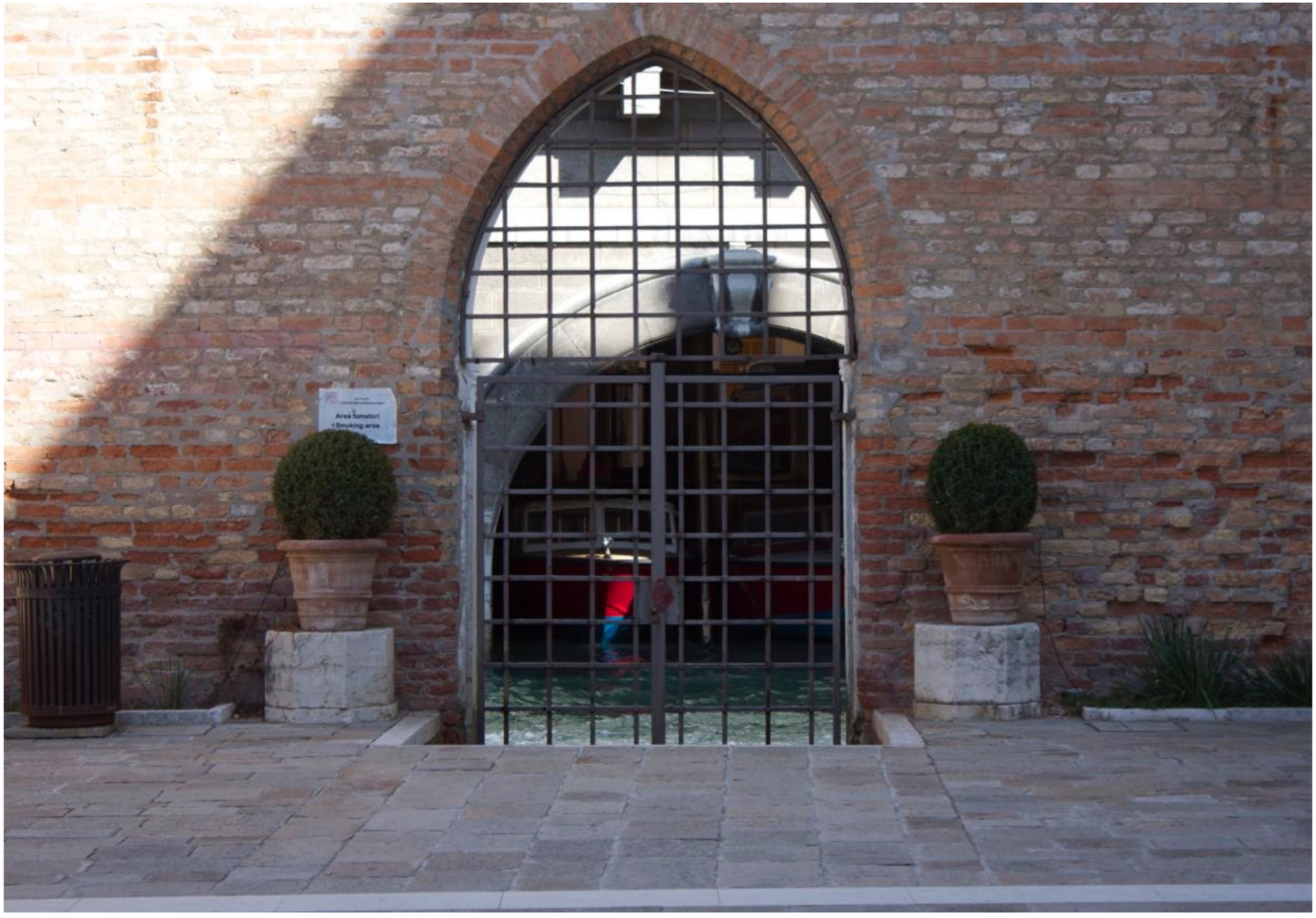
- Self-occlusions have to be treated with care!

- We have to cover all self-occluded parts.

Best Practices: Lighting Conditions

Cloudy days are ideal because lighting is stable!

Best Practices:
Lighting Conditions



Avoid moving shadows!

Capturing People

- To capture faces, hands, etc, we have to:
  - Make people to stay still; i.e., find a sit!
  - Use markers (e.g., checker boards) to increase features.
  - Make them wear patternful clothes.

Software

Software

Free Software:

- Regard 3D
- Multi-View Environment (MVE)
- COLMAP

Commercial Software:

- 3DZephyr
- Autodesk ReCap
- Agisoft Metashape
- RealityCapture

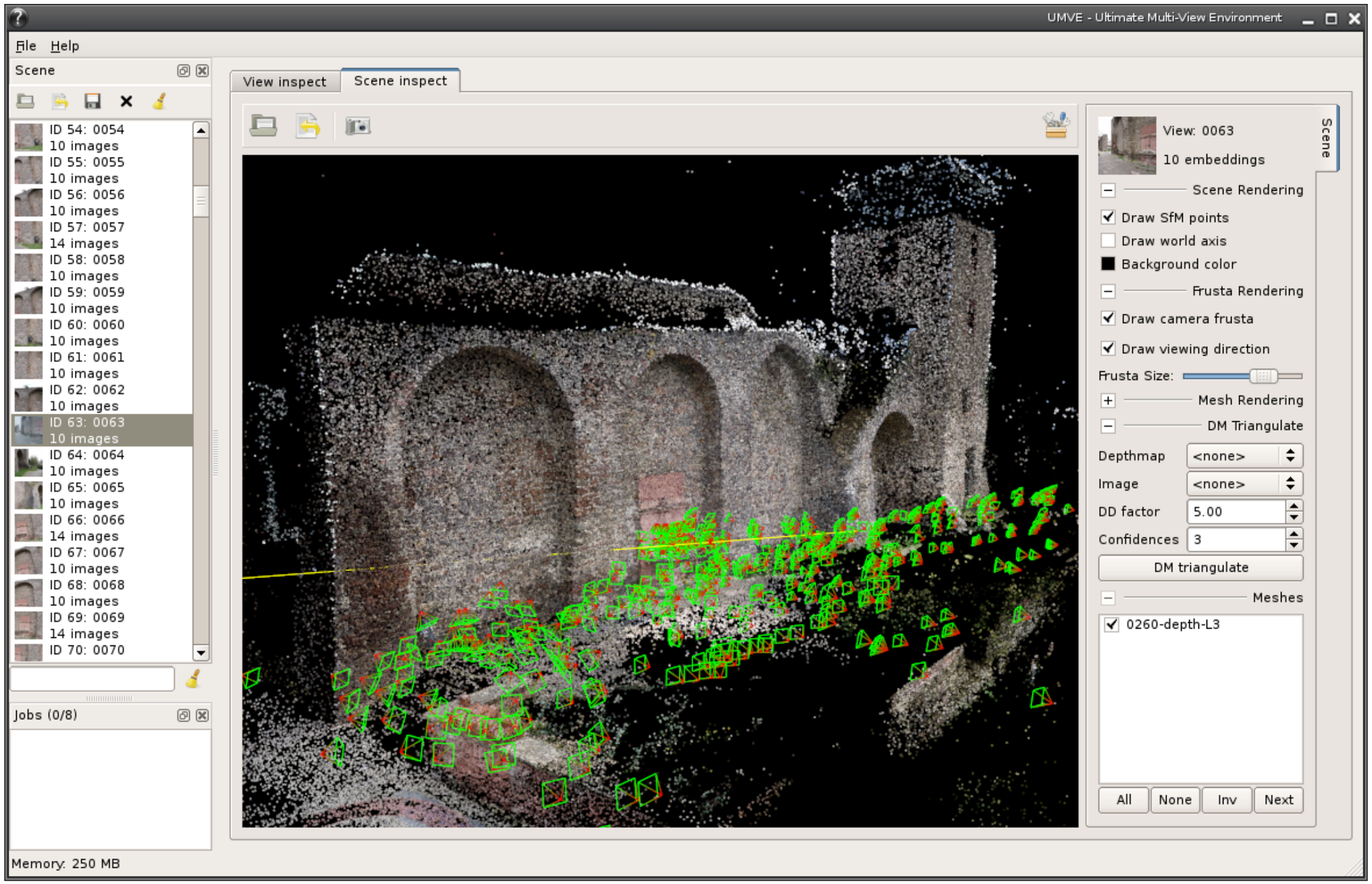
Software: Regard3D

- GUI application that integrates different open-source libraries and tools (MVE, OpenMVG, and CMVS/PMVS).
- Complete pipeline; from the Structure-from-Motion to the Surface Reconstruction.
- Open-source (C++).
- This tool works locally.
- http://www.regard3d.org/index.php

Software:
Multi-View Environment

- End-to-end pipeline for image-based geometry reconstruction: Structure-from-Motion, Multi-View Stereo, and Surface Reconstruction.
- Command line applications, but most features are also available from our user interface UMVE.
- Open-source (C++).
- This tool works locally.
- External tool for high resolution texturing:
  - https://github.com/simonfuhrmann/mve

Software:
Multi-View Environment



Software: COLMAP

- End-to-end pipeline for image-based geometry reconstruction:
  - Structure-from-Motion.
  - Multi-View Stereo.
- Command line and graphical user interface.
- Open-source (C++).
- This tool works locally.
- External tool for high resolution texturing:
  - https://colmap.github.io/

Software: 3DZephyr

- It is a local software that requires an one time fee:
  - Free version (50 photos limit).
  - $149 Lite version.
  - $3,900 Pro version.
- It creates high-quality 3d meshes.

Software: Autodesk ReCap

- It was a web service by Autodesk:

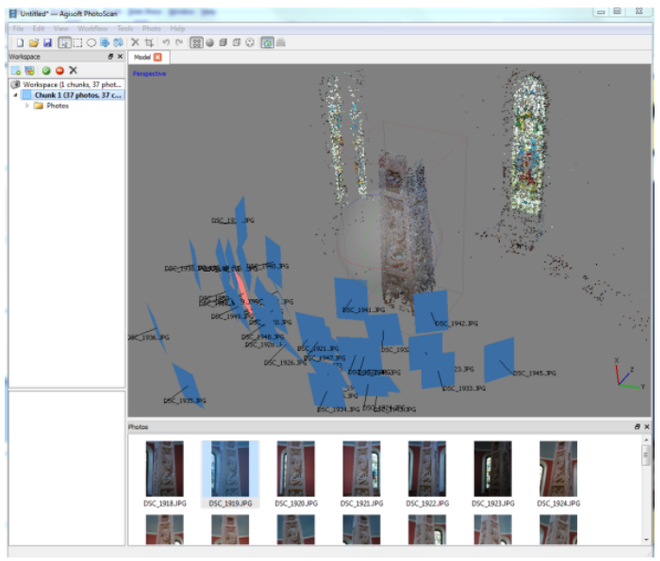
  https://www.autodesk.com/products/recap/overview
- Main advantages:
  - It is robust; it works even with challenging photo sequences.
  - It creates high quality results.
- It requires a subscription per year; roughly $360.
- It exports 3D models with color.

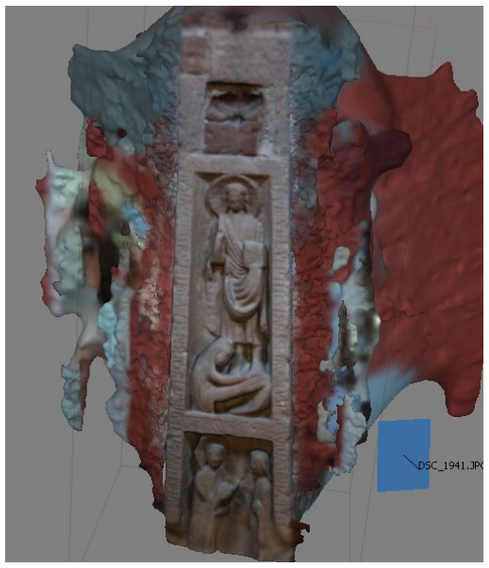
Software: Autodesk ReCap

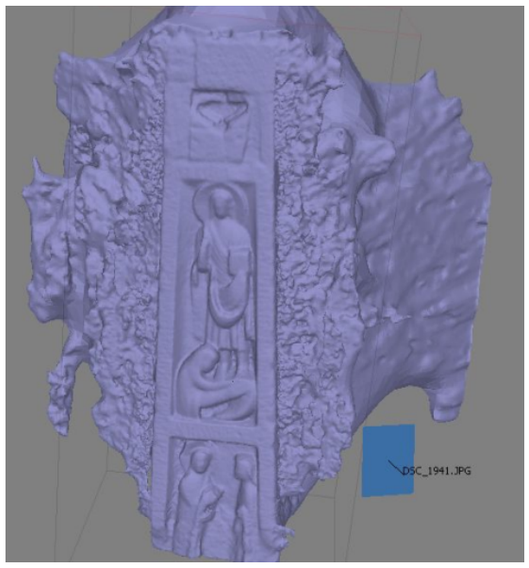
An example generated using the previous iteration of the Software

Software: Agisoft Metashape

- It is a local software by Agisoft that requires an one time fee:
  - $59 for students
  - $179 standard version
- It is fast and creates high quality models with color.







Software: RealityCapture

- It is a local software by Capture Reality with a subscription fee:
  - $10-20 subscriptions.
  - $3,750 forever.
- The fastest software (recently a real-time capture and reconstruction inside Unity), but very expensive.

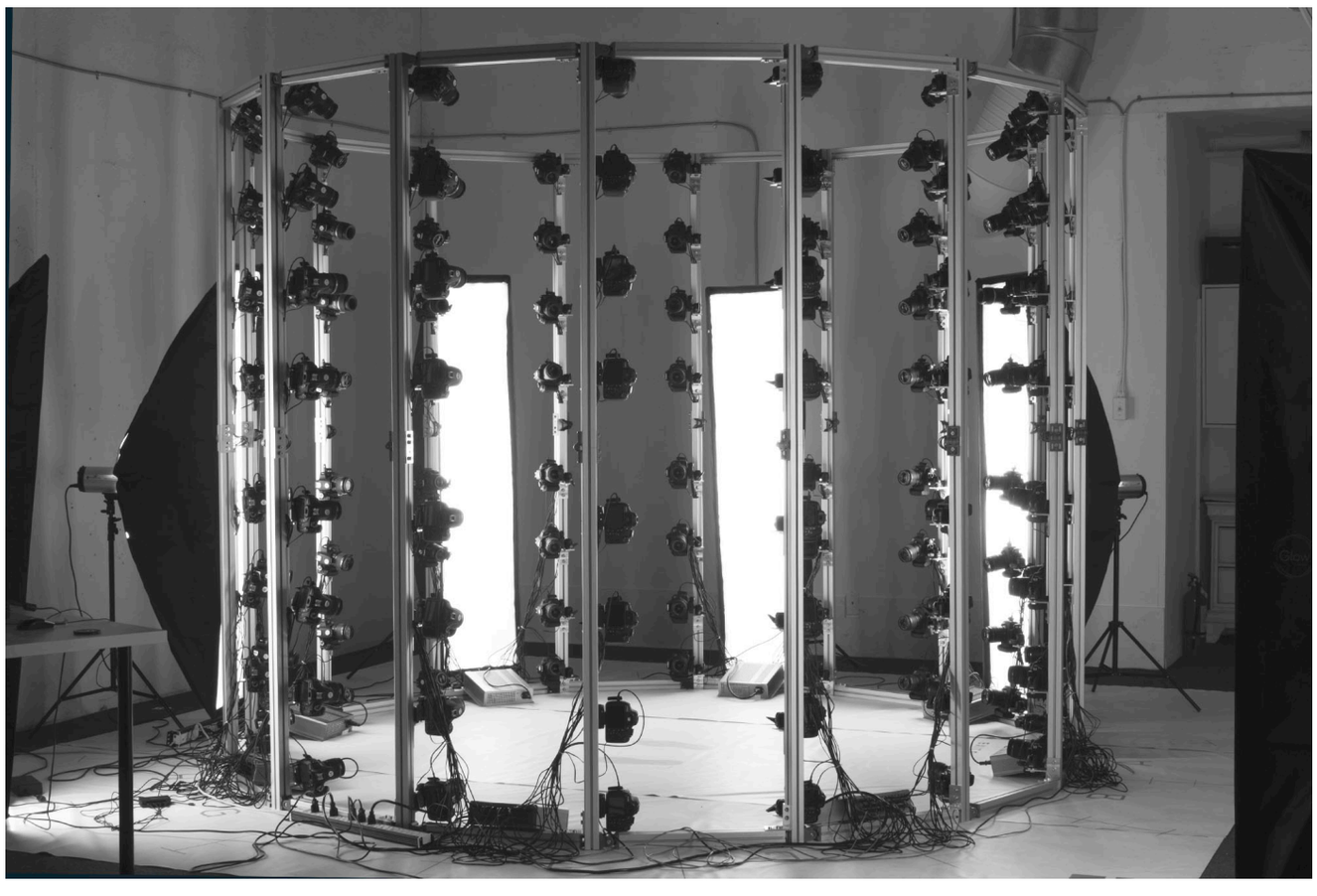
Why 3D from Photographs in this course?

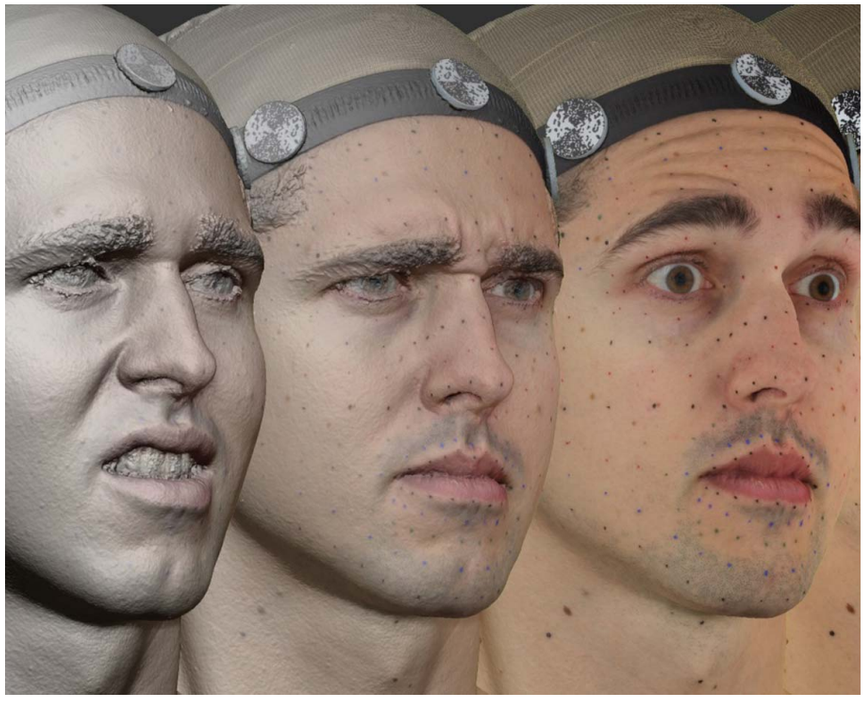
Why?

- Fast and cheap acquisition of 3D model for 3D printing of replicas of anatomical parts.
- Full body 3D scanning
  - Healthcare
  - Research (ergonomic, biometrics, anthropology)
  - Size & Fit
  - 3D printed figurine

Full Body 3D Scanning

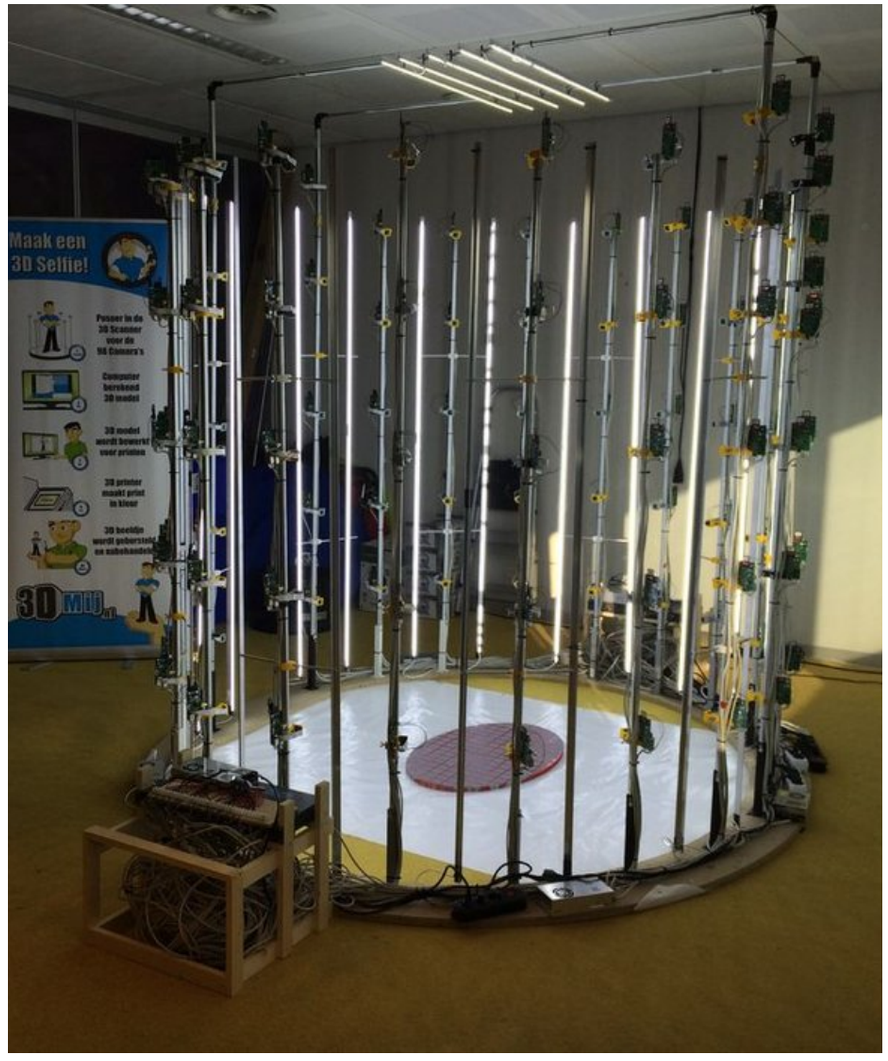
- 128X DSLR camera:
  - http://pixellighteffects.com/





Full Body 3D Scanning

- Raspberry PI with 100 camera module:
  - http://www.pi3dscan.com



that’s all folks!

Acknowledgements

- Some images on work by:
  - Dr. Matteo Dellepiane:
    - http://vcg.isti.cnr.it/~dellepiane/