3D Vision

Marc Pollefeys and Daniel Barath

Spring 2024

3D Vision

- Understanding geometric relations
  - between images and the 3D world
  - between images
- Obtaining 3D information describing our 3D world
  - from images
  - from dedicated sensors

3D Vision

- Extremely important in robotics, self-driving cars and AR / VR
  - Visual navigation
  - Sensing / mapping the environment
  - Obstacle detection, …
- Many further application areas
- A few examples …

ETH

Google Tango

(officially discontinued, lives on as ARCore)

Google Tango

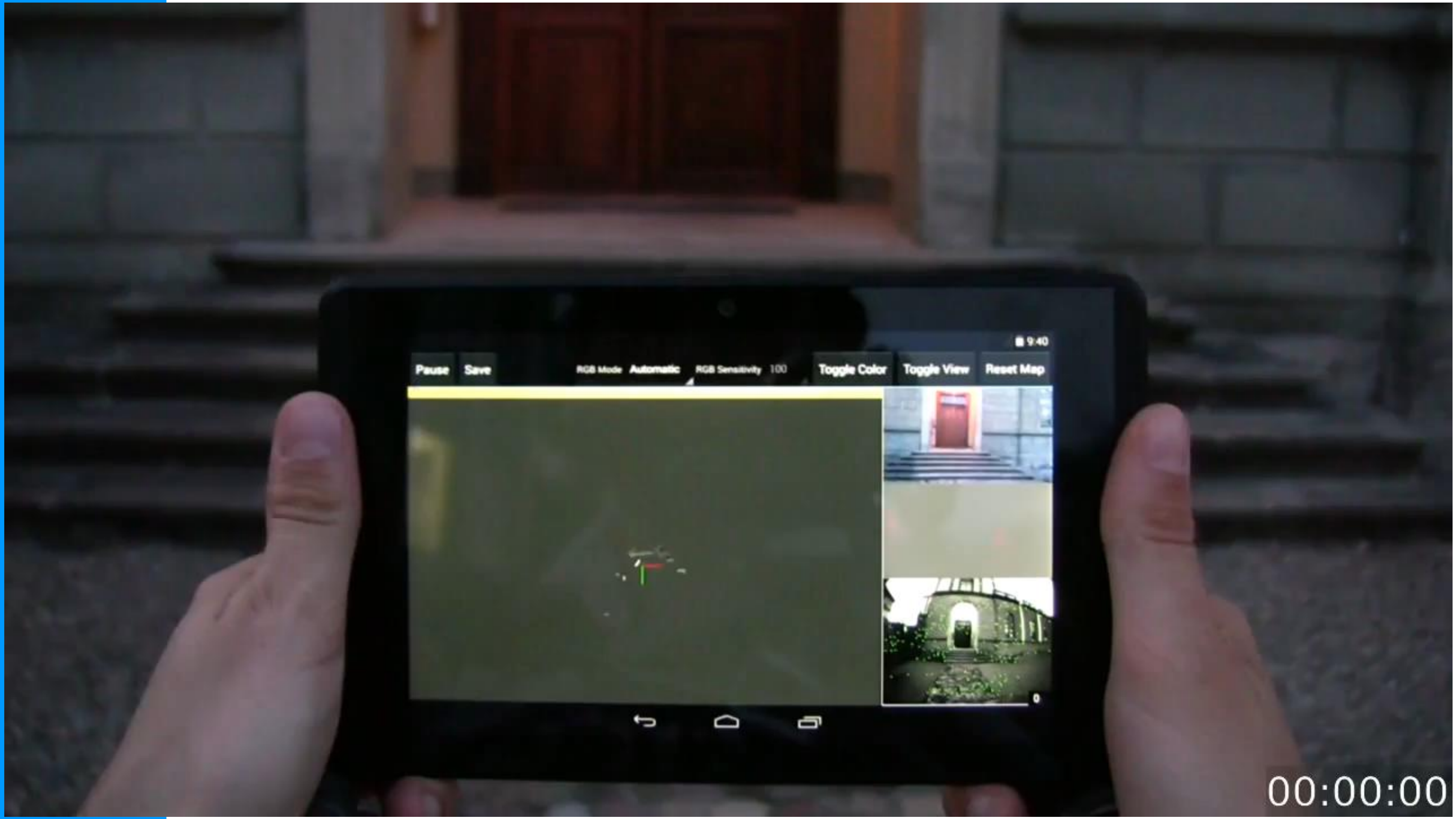


Image-Based Localization

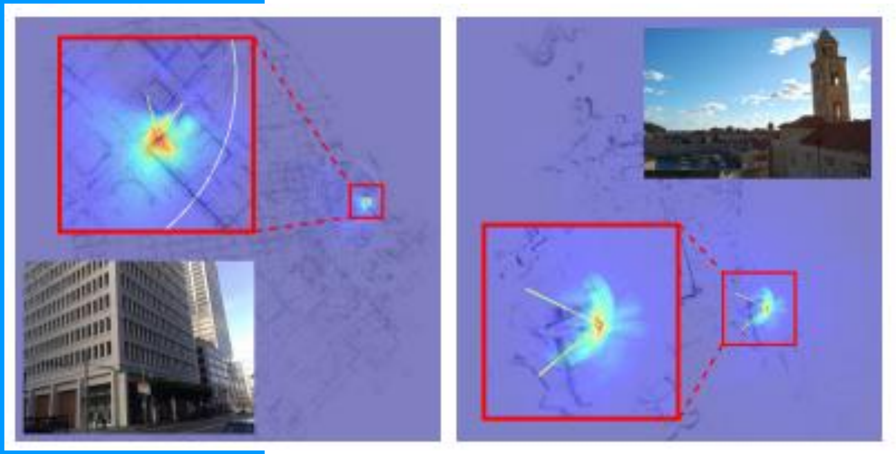
Large-scale, Real-Time
Visual-Inertial Localization

Simon Lynen, Torsten Sattler,
Mike Bosse, Joel Hesch,
Marc Pollefeys and Roland Siegwart

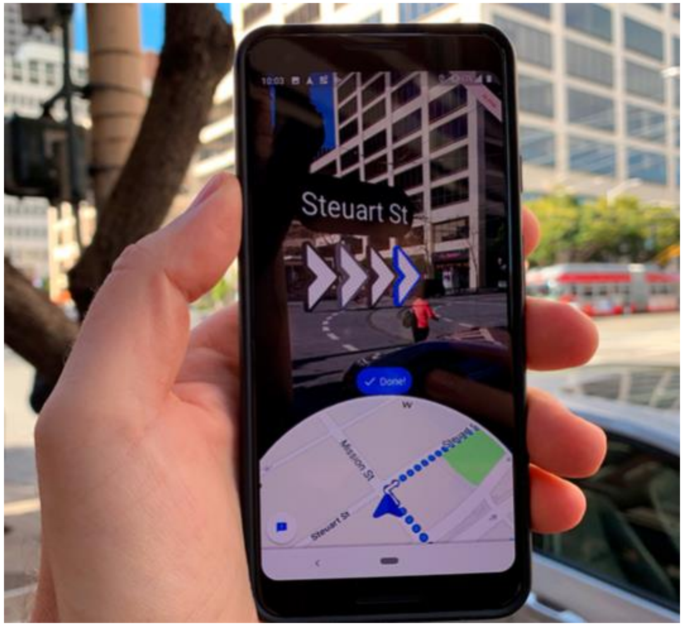
Geo-Tagging Holiday Photos

(Li et al. ECCV 2012)

Image-based localization

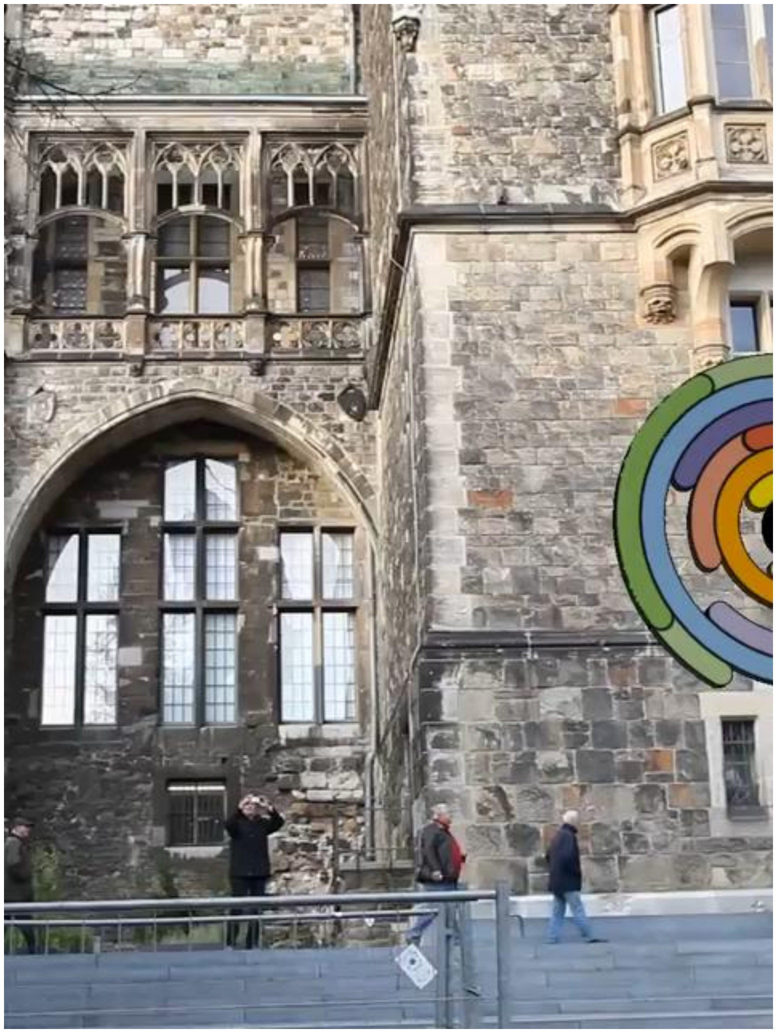


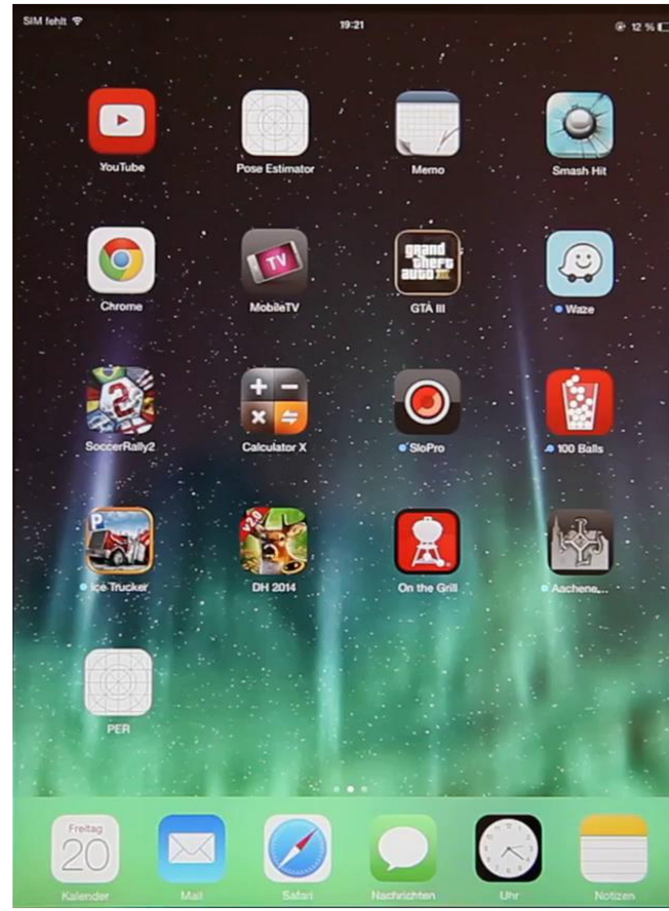
Zeisl et al ICCV2015



Google Maps AR mode

Augmented Reality





(Middelberg et al. ECCV 2014)

Large-Scale Structure-from-Motion

Rome dataset

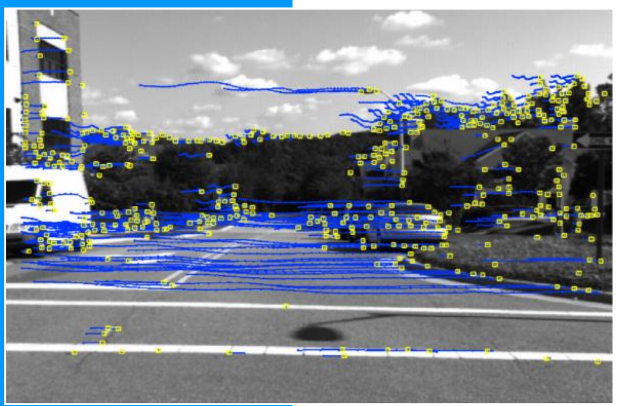
74,394 images

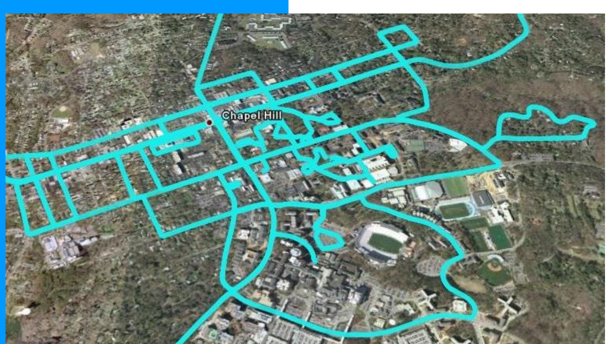
Video credit: Johannes Schönberger

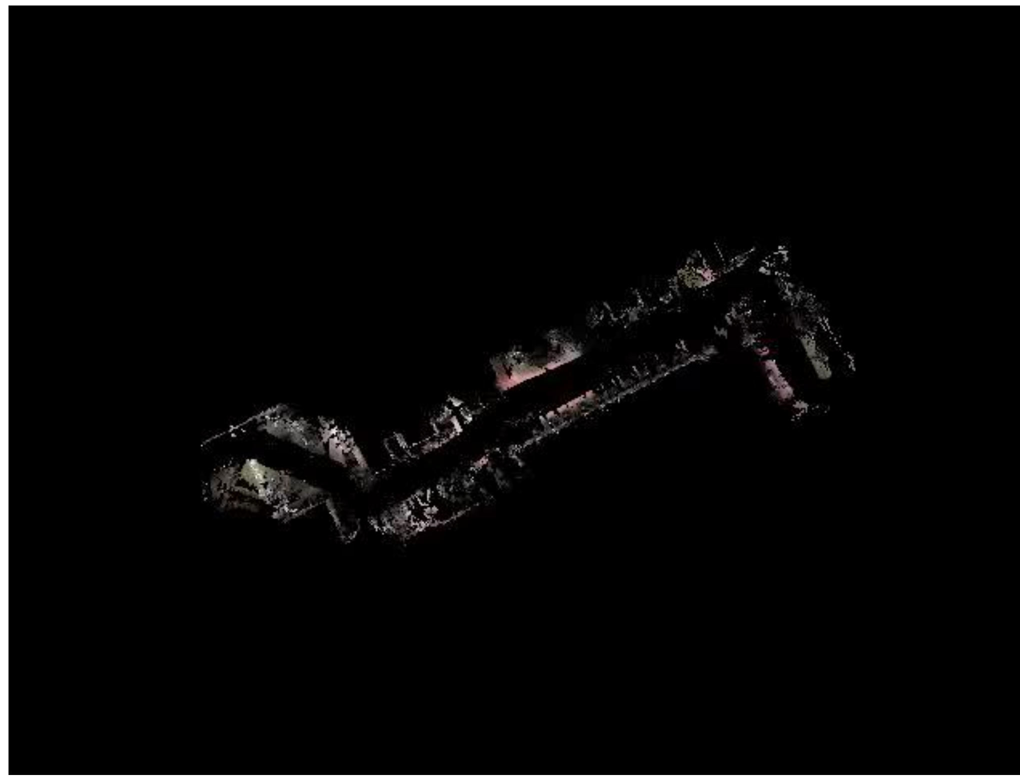
Virtual Tourism

ETH

3D Urban Modeling

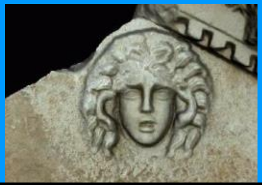






UNC/UKY UrbanScape project

3D Urban Modeling

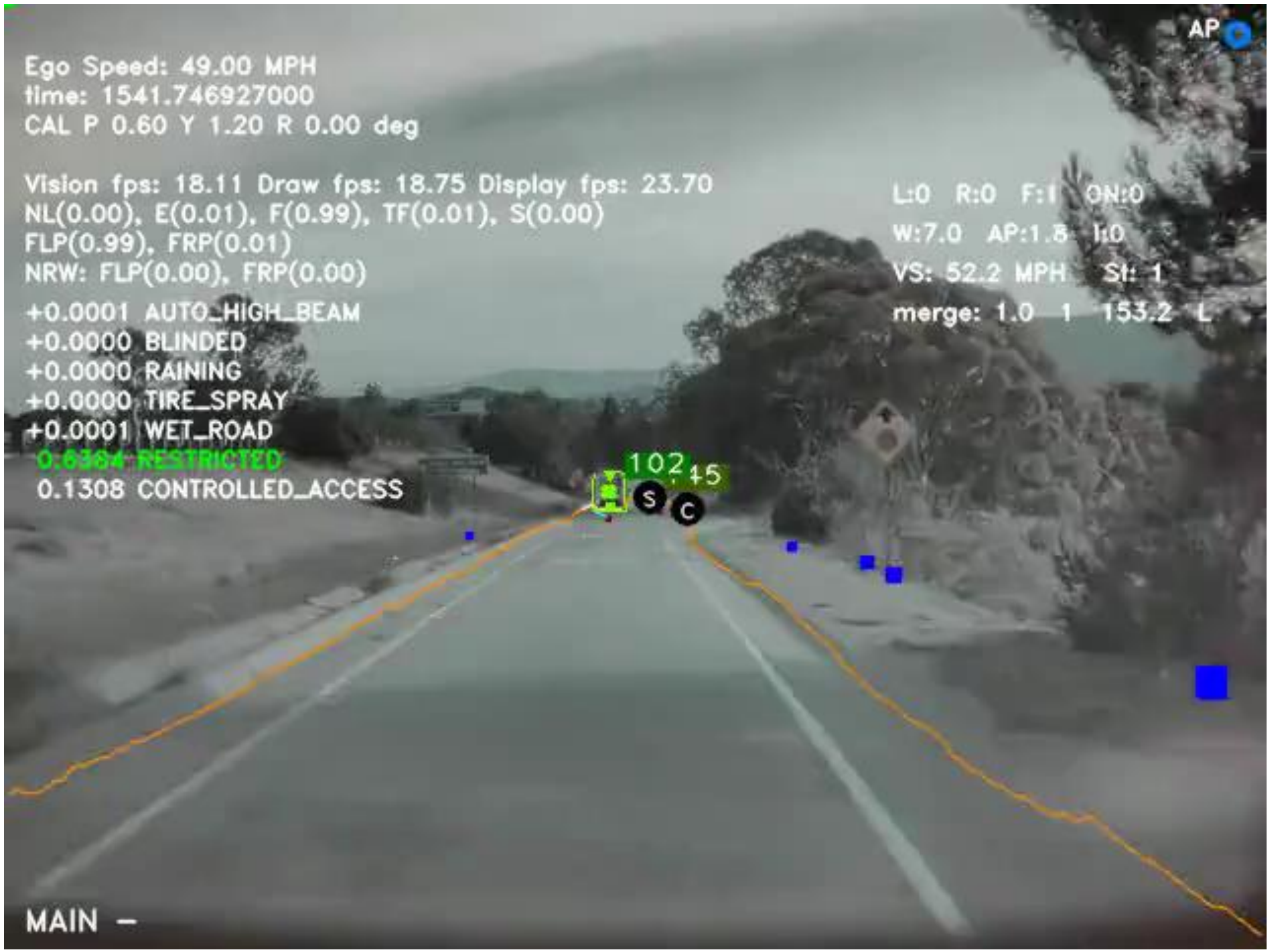



CapturingReality.com

Mobile Phone 3D Scanner

Mobile Phone 3D Scanner

Self-Driving Cars



Self-Driving Cars

AutoVision

3D Vision for Autonomous Vehicles

DSO NATIONAL LABORATORIES

ETH zürich

NUS National University of Singapore

ETH

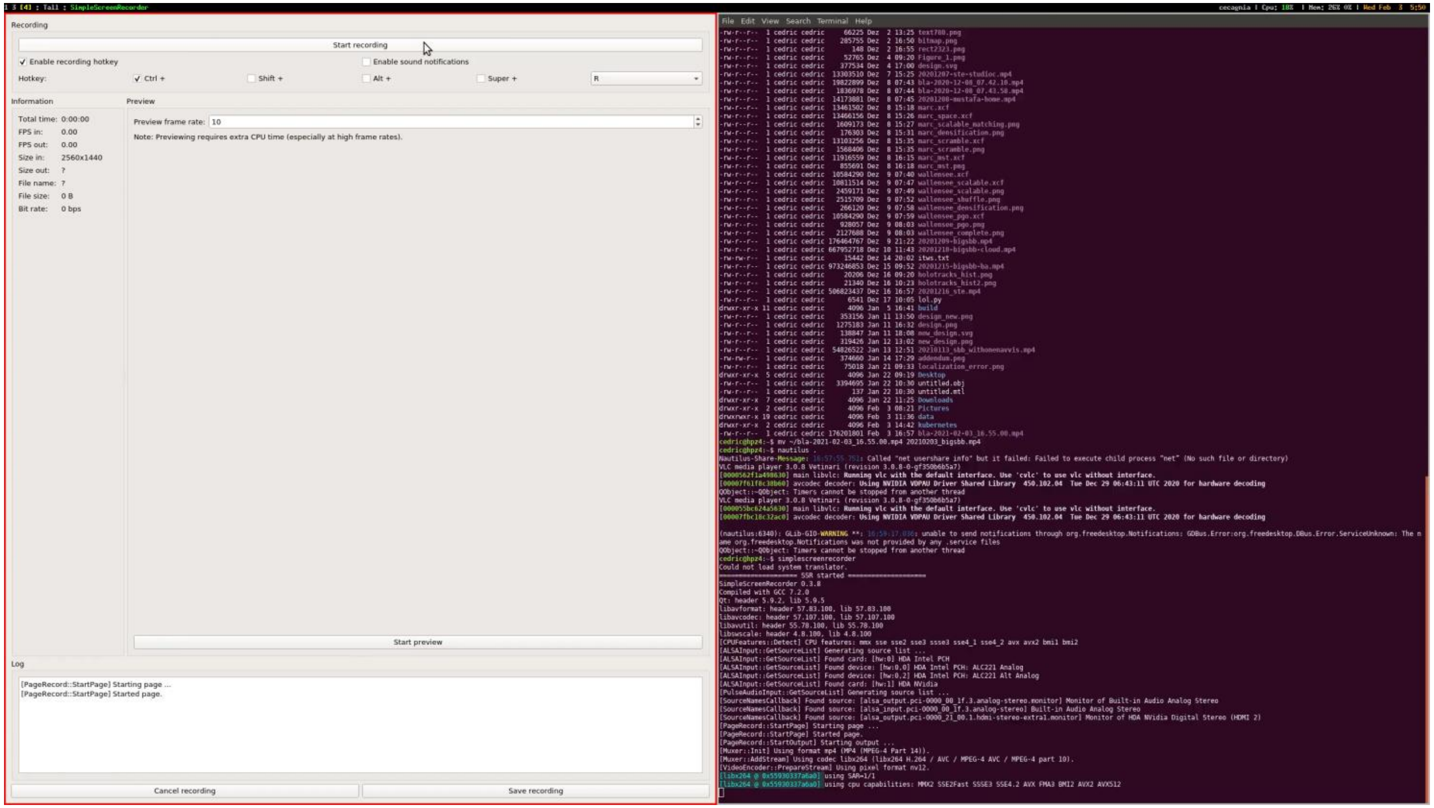
Micro Aerial Vehicles

ETH

Mixed Reality

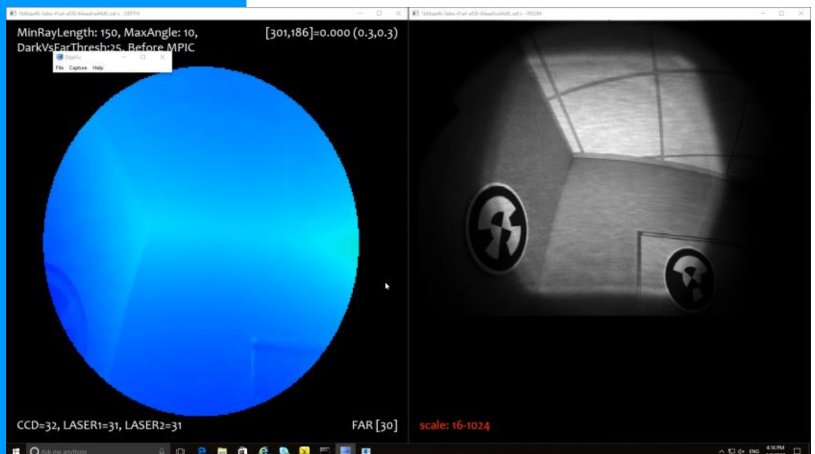
Microsoft HoloLens 2

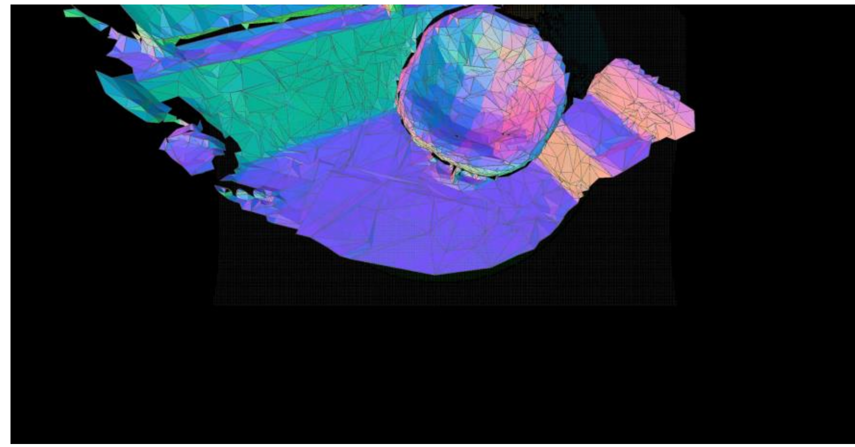
Large-scale visual mapping

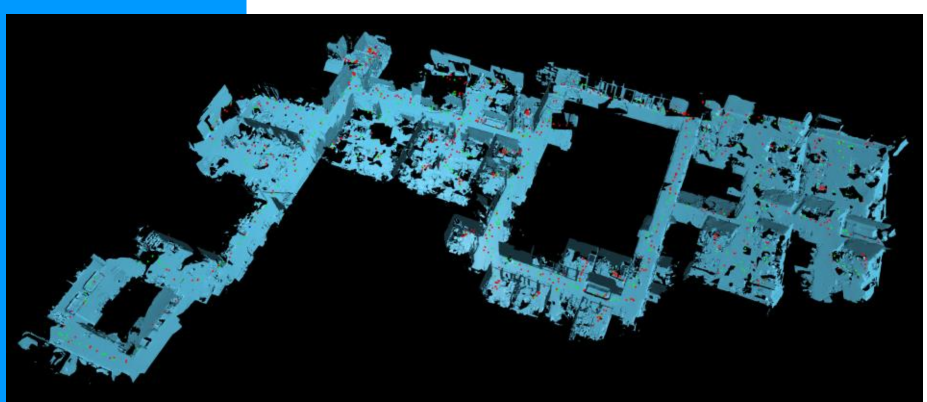


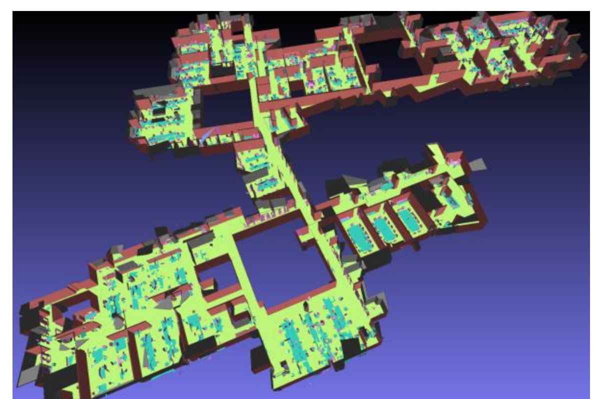
Mixed Reality

depth sensing and surface reconstruction









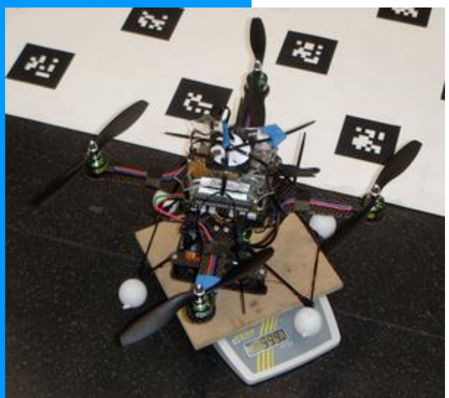
Virtual Reality

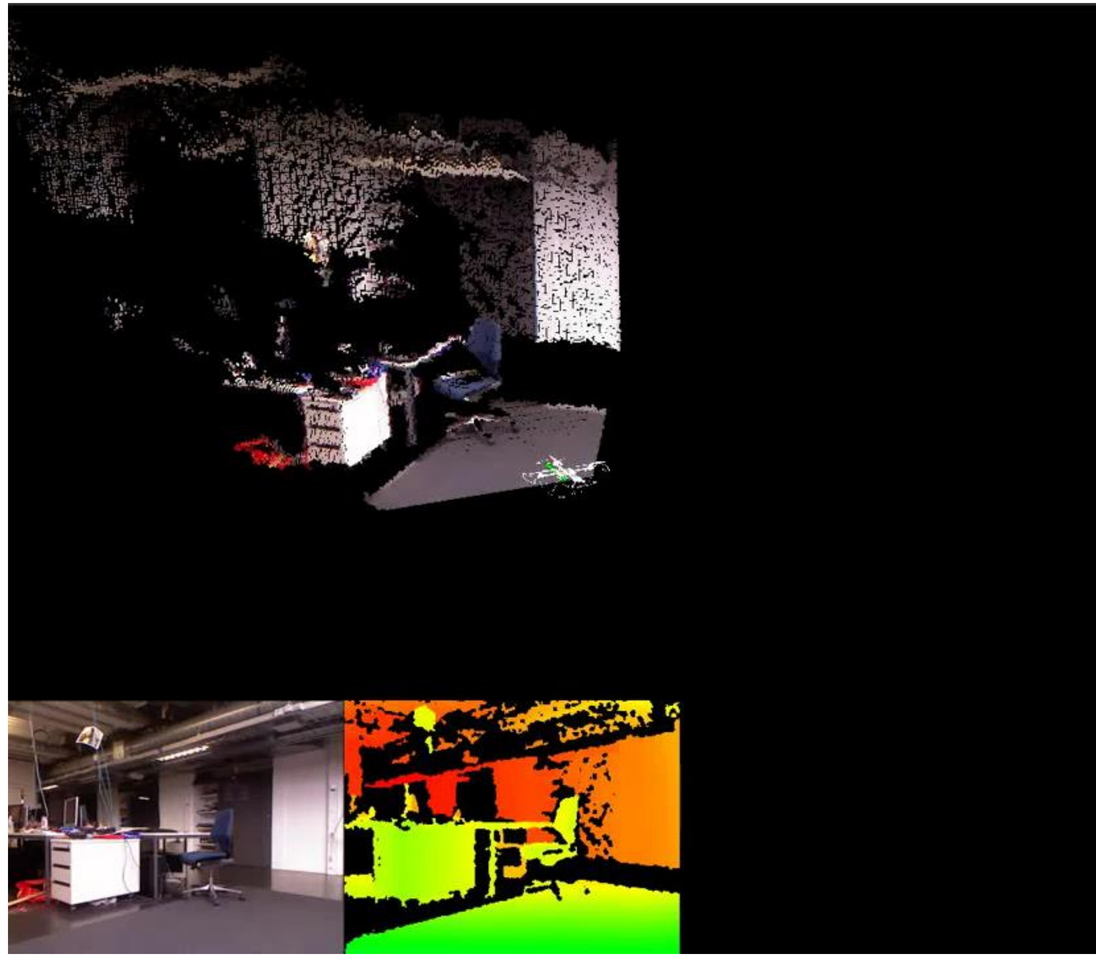
Raw Kinect Output: Color + Depth

http://grouplab.cpsc.ucalgary.ca/cookbook/index.php/Technologies/Kinect

Human-Machine Interface

Autonomous Micro-Helicopter Navigation





Use Kinect to map out obstacles and avoid collisions

ETH

Dynamic Reconstruction

DynamicFusion:

Reconstruction & Tracking of Non-rigid Scenes in *Real-Time*

Richard Newcombe, Dieter Fox, Steve Seitz

Computer Science and Engineering,
University of Washington

ETH

Performance Capture

High-Quality Passive Facial Performance Capture using Anchor Frames

Thabo Beeler, Fabian Hahn, Derek Bradley, Bernd Bickel,
Paul Beardsley, Craig Gotsman, Robert. W. Sumner, Markus Gross

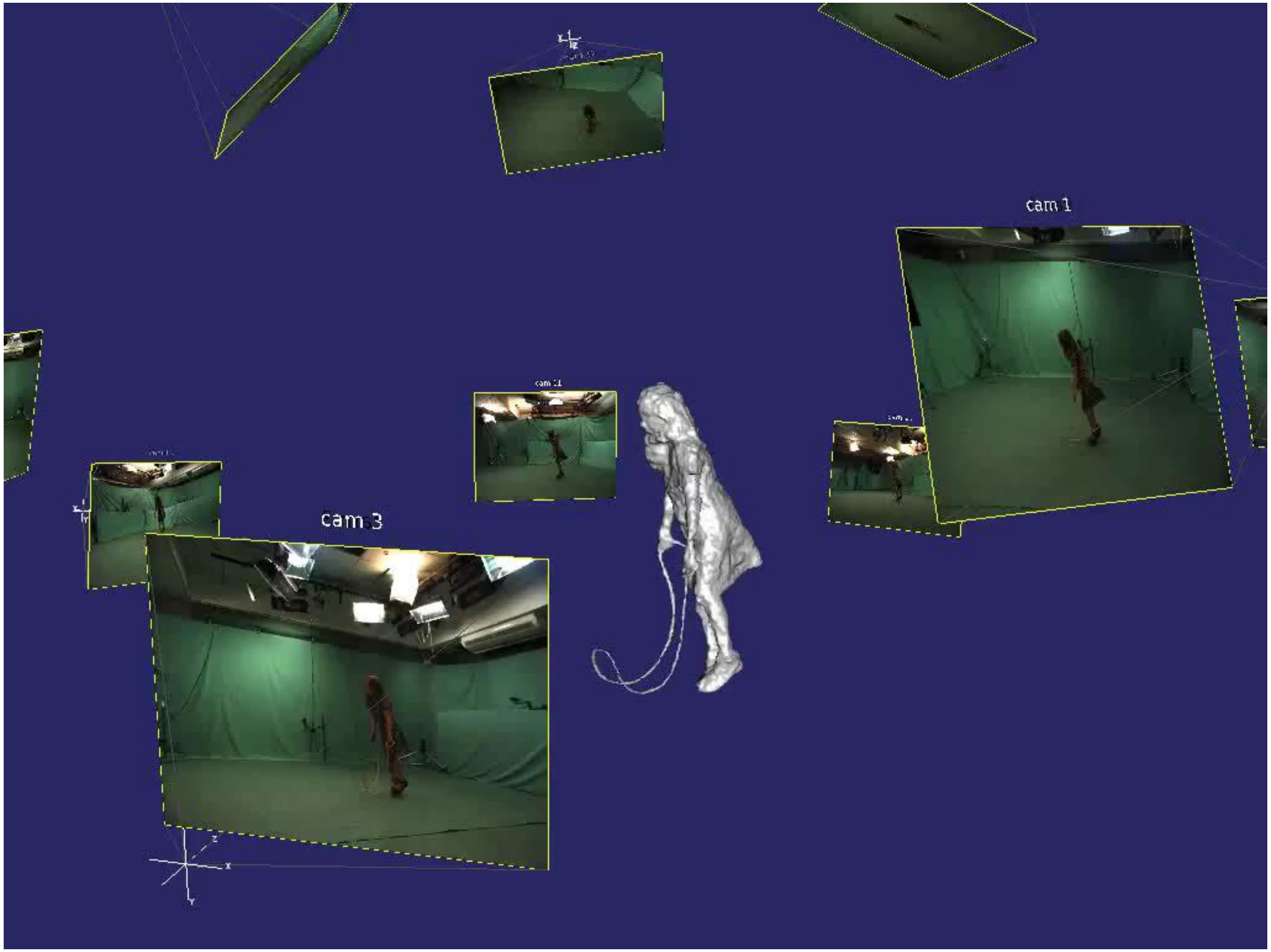
Disney Research, Zurich

ETH Zürich

Technion
Israel Institute of Technology

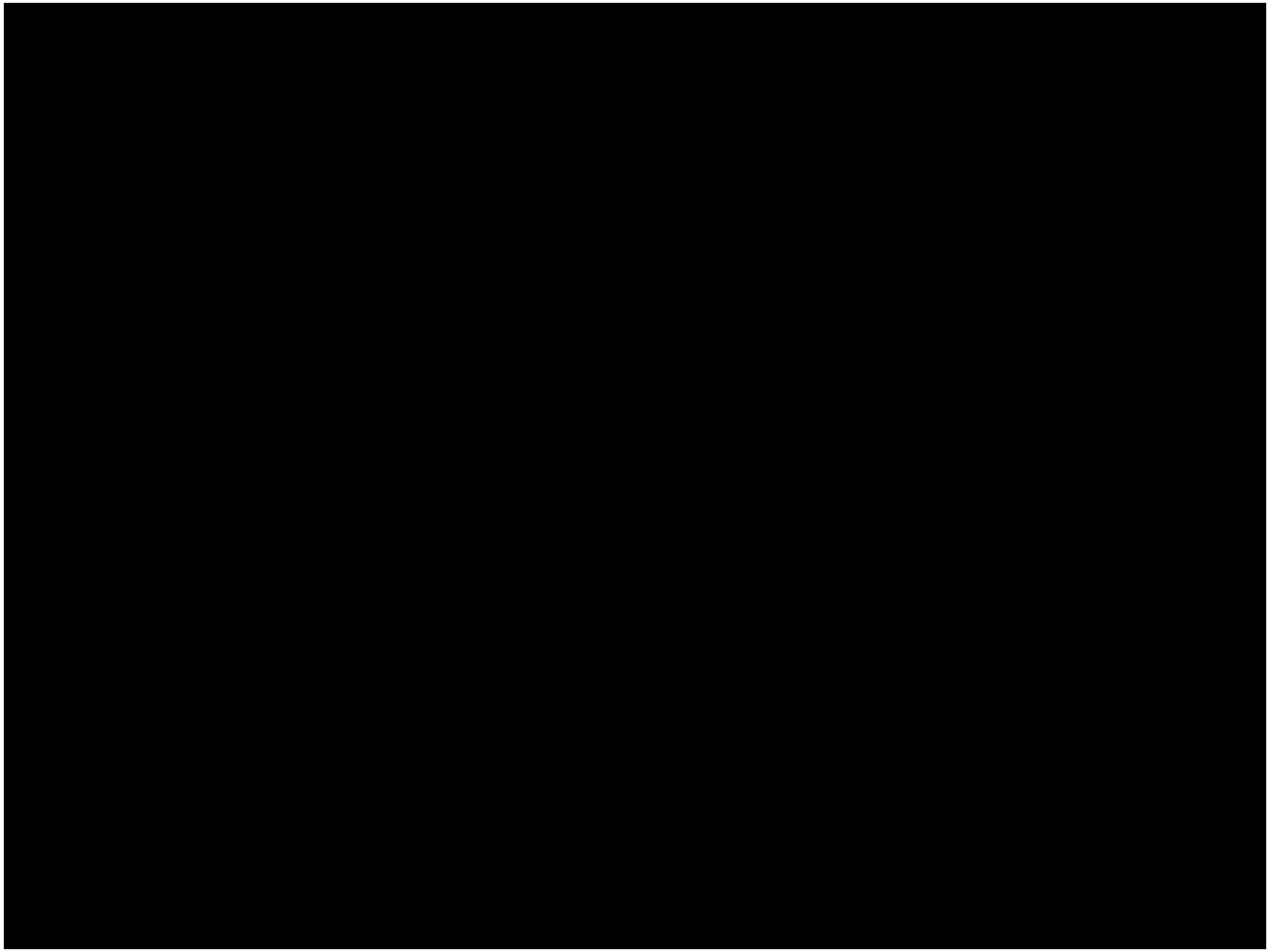
ETH

Performance Capture



(Oswald et al. ECCV 14)

Motion Capture



Interactive 3D Modeling

(Sinha et al. Siggraph Asia 08)

collaboration with Microsoft Research (and licensed to MS)

ETH

Scanning Industrial Sites

as-build 3D model of off-shore oil platform

Scanning Cultural Heritage

ETH

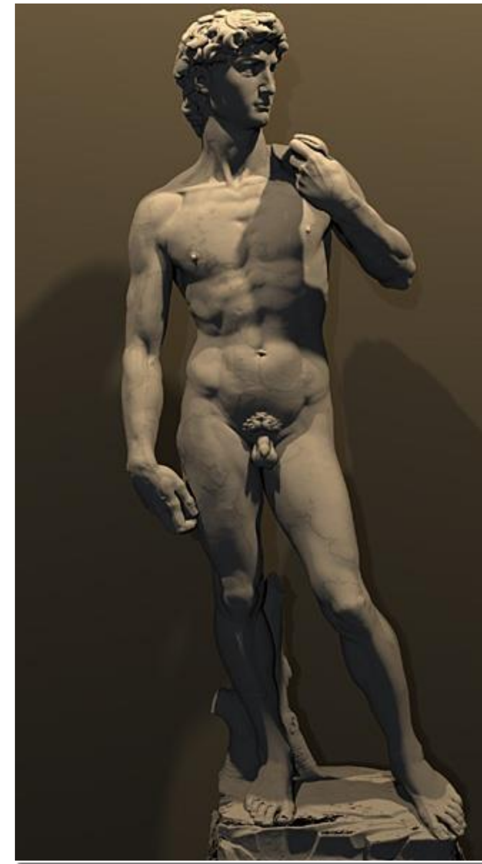
Some more scanner examples

- https://www.navvis.com/m6
- https://shop.leica-geosystems.com/learn/reality-capture/blk2go

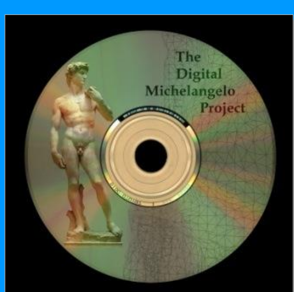
ETH

Cultural Heritage

Stanford's Digital Michelangelo



Digital archive
Art historic studies



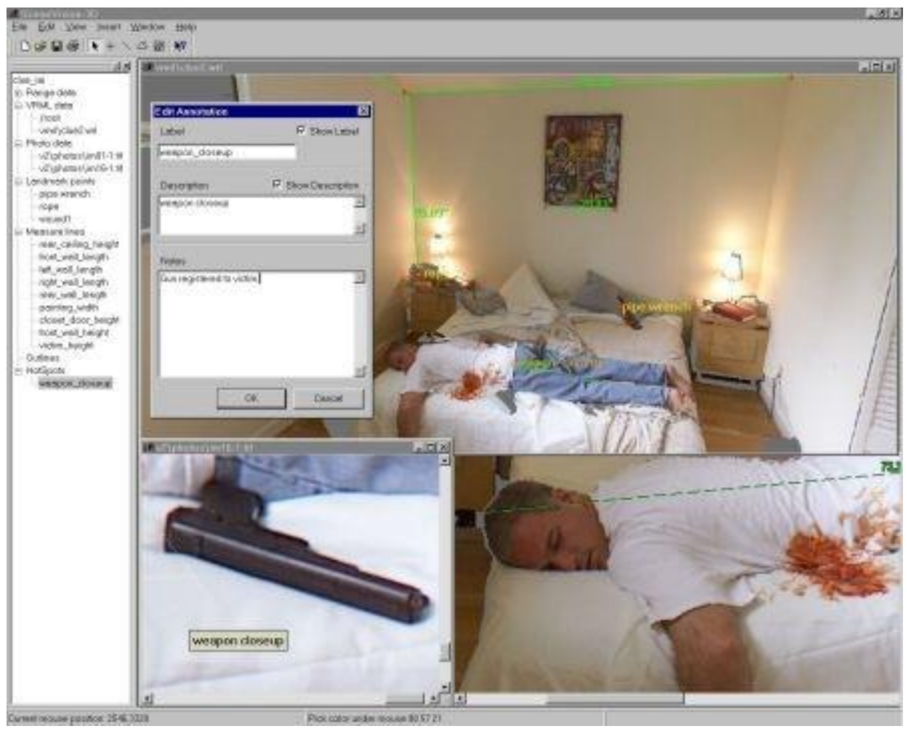
Archaeology


accuracy ~1/500 from DV video (i.e. 140kb jpegs 576x720)

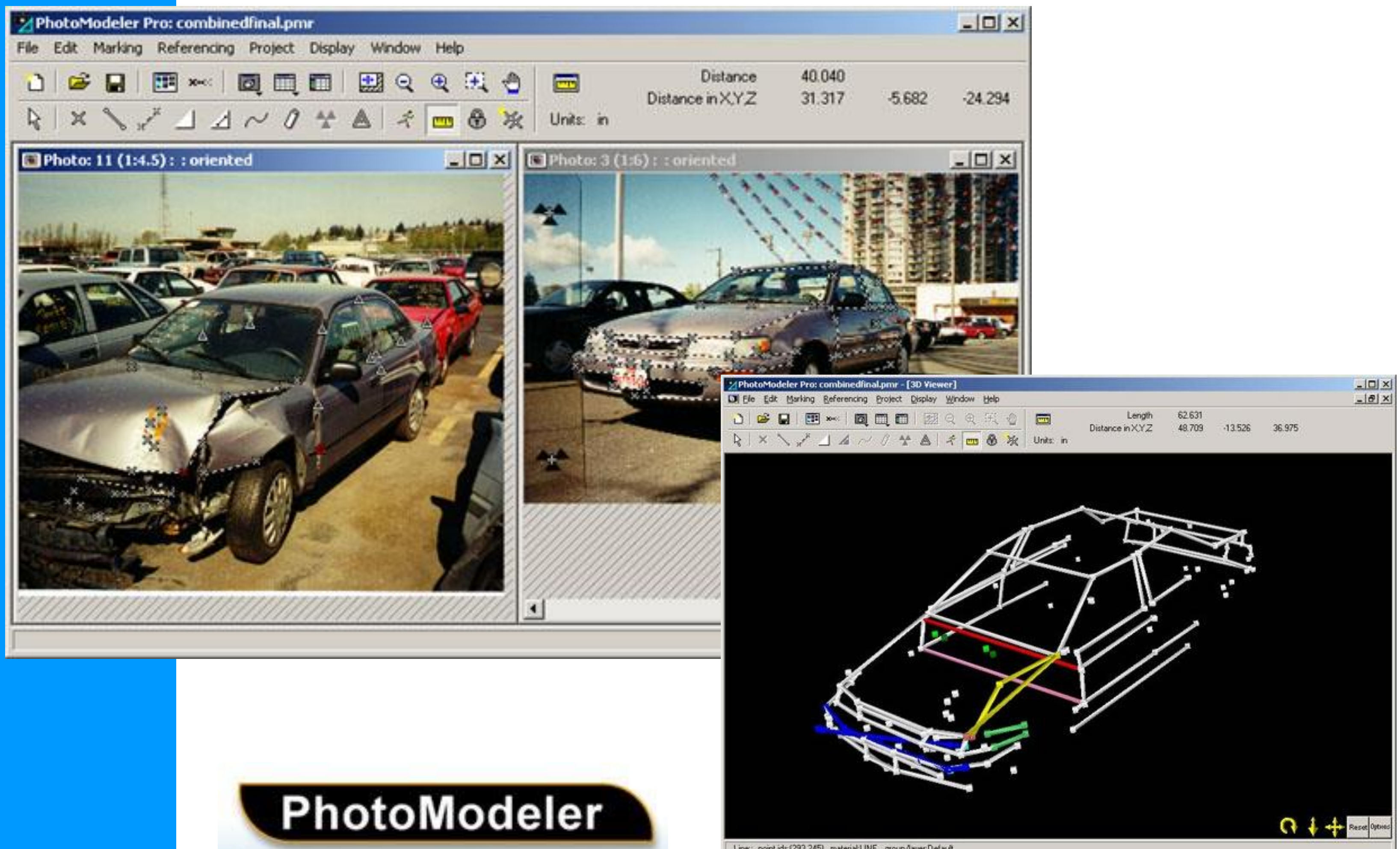
ETH

Forensics

- Crime scene recording and analysis



Forensics



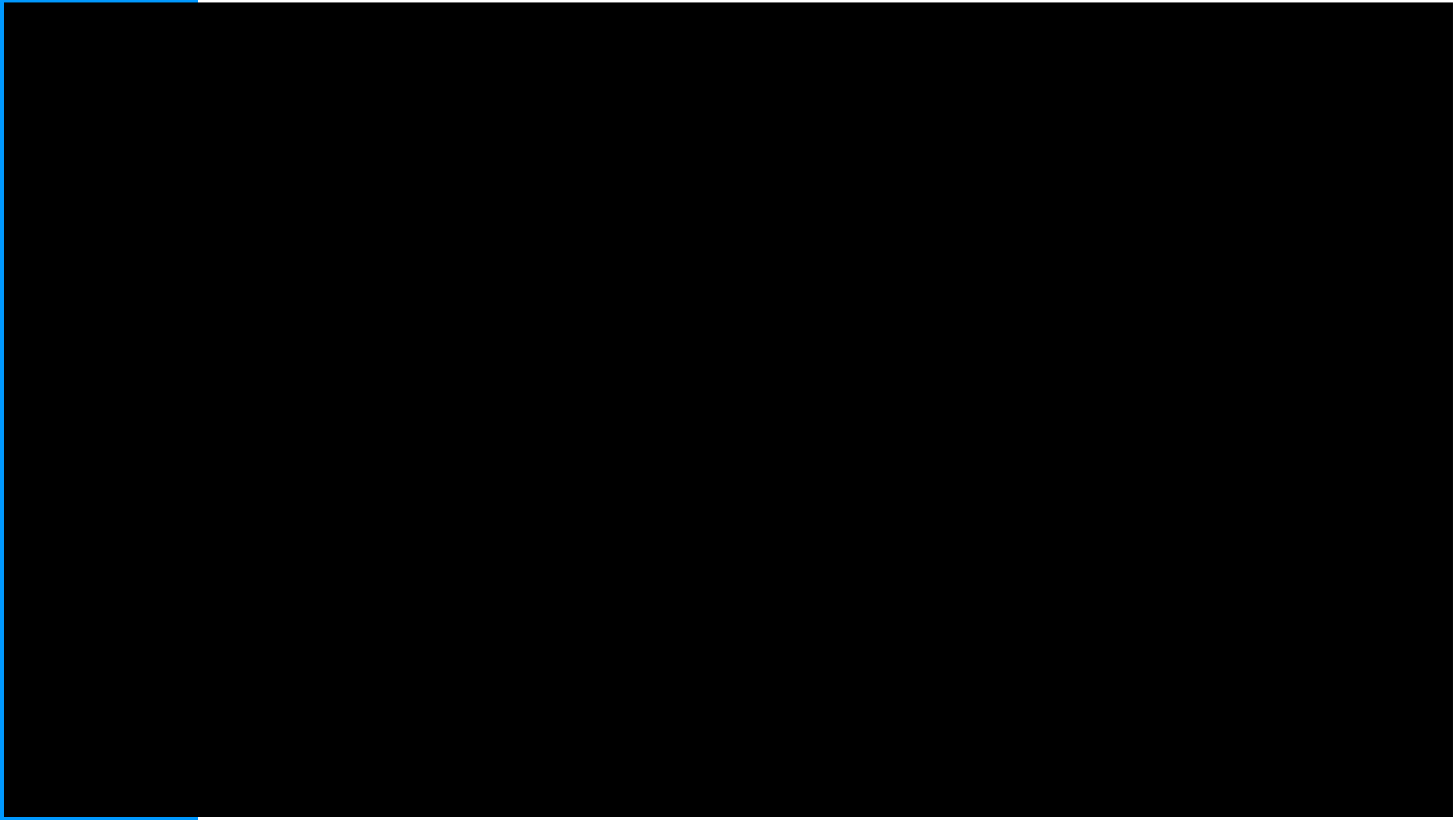
PhotoModeler

ETH

Sports

ETH

Surgery



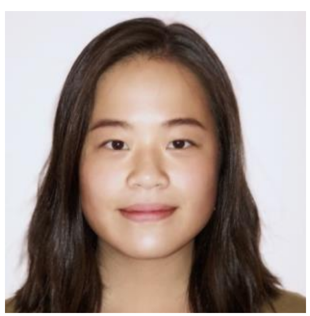
3D Vision Course Team

Marc Pollefeys
CNB G 105
marc.pollefeys@inf.ethz.ch

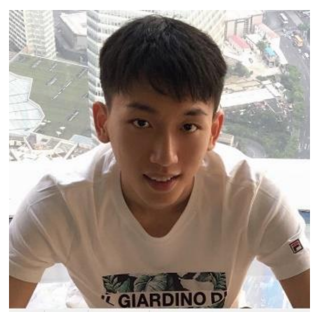
Daniel Barath
CNB G 103.2
daniel.barath@inf.ethz.ch

Zuoyue Li
CAB G 85.2
li.zuoyue@inf.ethz.ch

Fangjinhua Wang
CNB G 100.9
fangjinhua.wang@inf.ethz.ch



Jiaqi Chen
CAB G 86.3
jiaqi.chen@inf.ethz.ch



Shaohui Liu
CNB G 102.2
shaohui.liu@inf.ethz.ch

Boyang Sun
UNH F 3
boyang.sun@inf.ethz.ch

Course Objectives

- To understand the concepts that relate images to the 3D world and images to other images
- Explore the state of the art in 3D vision
- Implement a 3D vision system/algorithm

Learning Approach

- Introductory lectures:
  - Cover basic 3D vision concepts and approaches.
- Further lectures:
  - Short introduction to topic
  - Paper presentations (you)
    (seminal papers and state of the art, related to your projects)
- 3D vision project:
  - Choose topic, define scope (by week 4)
  - Implement algorithm/system
  - Presentation/demo and paper report

Grade distribution

- Paper presentation & discussions: 25%
- 3D vision project & report: 75%

ETH

Materials

Slides and more

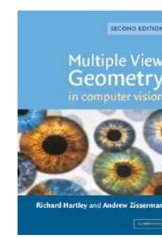
http://www.cvg.ethz.ch/teaching/3dvision/

Also check out on-line “shape-from-video” tutorial:

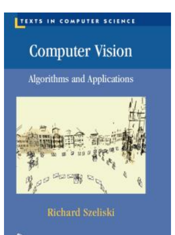
http://www.cs.unc.edu/~marc/tutorial.pdf

http://www.cs.unc.edu/~marc/tutorial/

Textbooks:



- Hartley & Zisserman, Multiple View Geometry



- Szeliski, Computer Vision: Algorithms and Applications

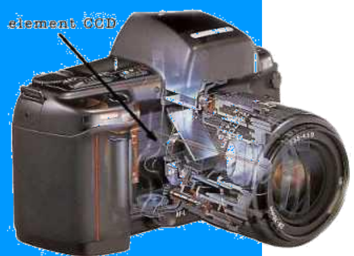
Schedule

Feb 19	**Introduction**
Feb 26	Geometry, Camera Model, Calibration
Mar 4	Guest lecture + Features, Tracking / Matching
Mar 11	**Project Proposals by Students**
Mar 18	3DV conference
Mar 25	Structure from Motion (SfM) + papers
Apr 1	Easter break
Apr 8	Dense Correspondence (stereo / optical flow) + papers
Apr 15	Bundle Adjustment & SLAM + papers
Apr 22	**Student Midterm Presentations**
Apr 29	Multi-View Stereo & Volumetric Modeling + papers
May 6	3D Modeling with Depth Sensors + papers
May 13	Guest lecture + papers
May 20	Holiday
May 27	**Student Project Demo Day = Final Presentations**

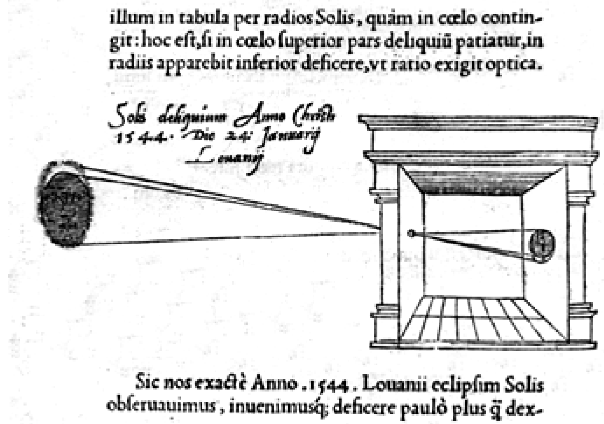
Fast Forward

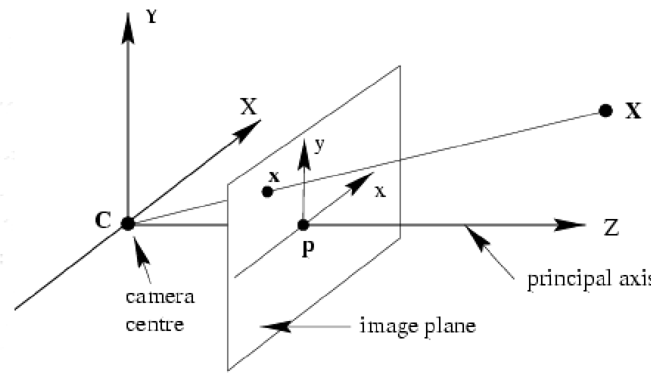
- Quick overview of what is coming…

Camera Models and Geometry



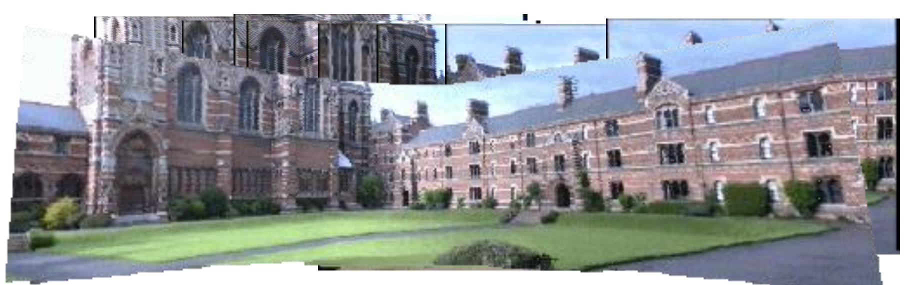
Pinhole camera

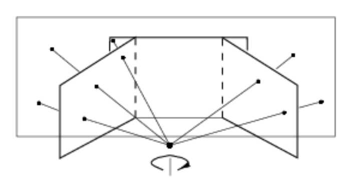




$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & & p_x \\ & f_y & p_y \\ & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \end{bmatrix} \begin{bmatrix} \mathtt{R} & \mathtt{t} \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \text{ or } \lambda \mathtt{x} = \mathtt{PX}$$

Geometric transformations in 2D and 3D

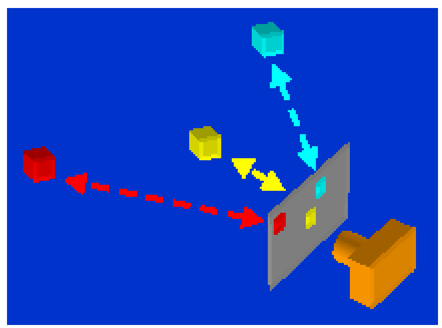


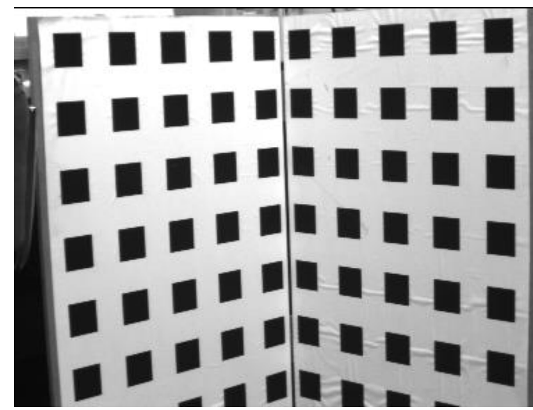


ETH

Camera Calibration

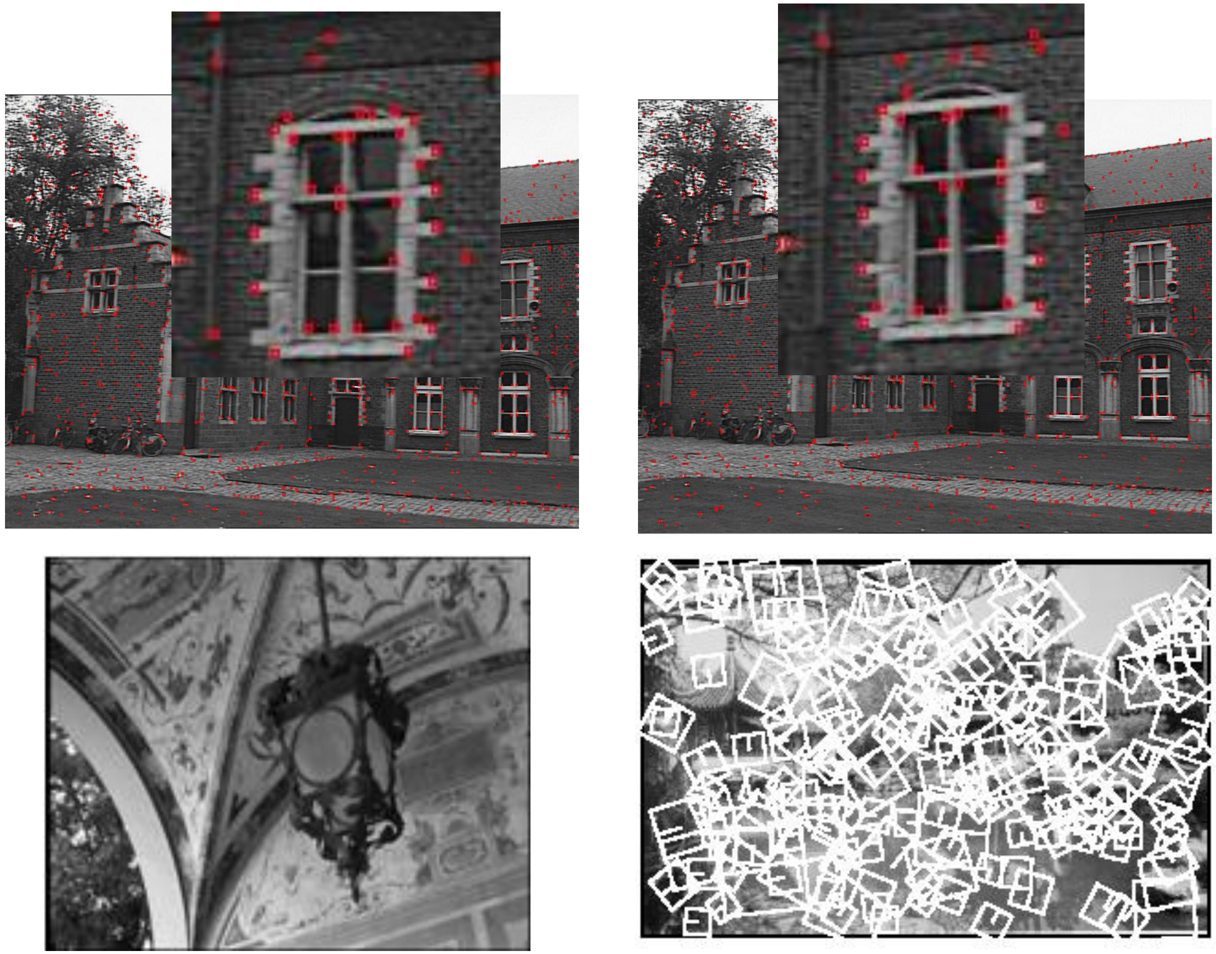
- Know 2D/3D correspondences, compute projection matrix





also radial distortion (non-linear)

Feature Tracking and Matching

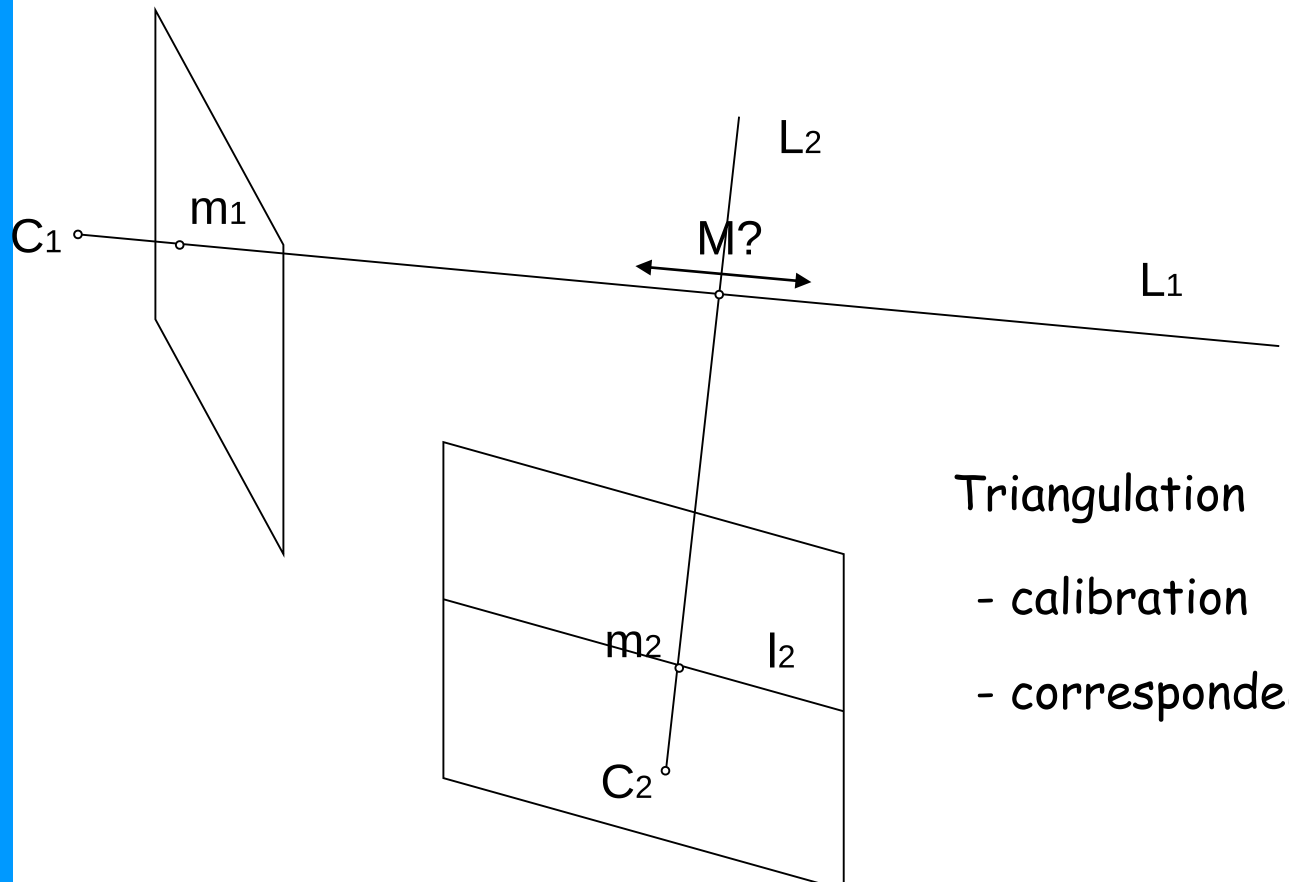


Harris corners, KLT features, SIFT features

key concepts: invariance of extraction, descriptors to viewpoint, exposure and illumination changes

ETH

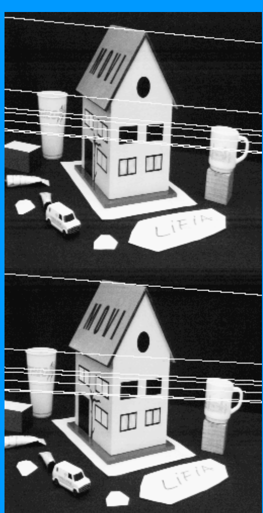
3D from Images



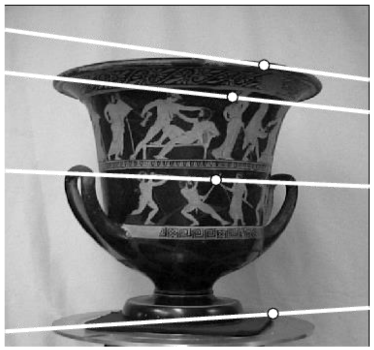
Triangulation

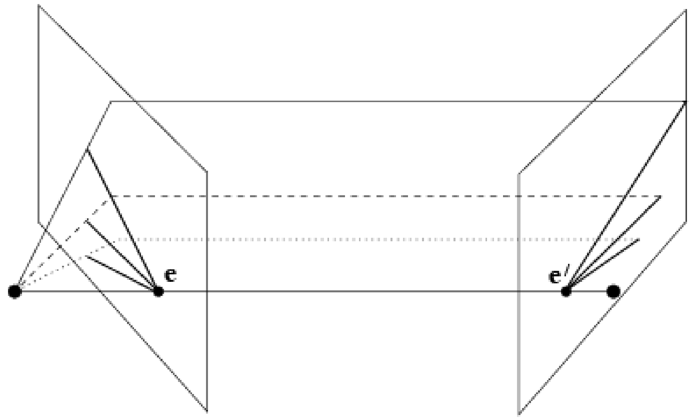
- calibration

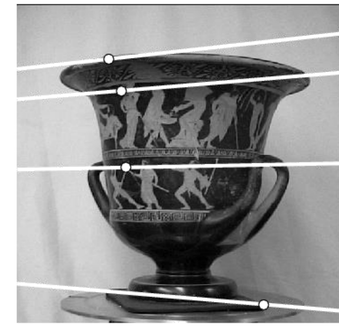
- correspondences



Epipolar Geometry







Fundamental matrix

$$x^\top F x = 0$$

$$F \leftrightarrow P, P'$$

Essential matrix

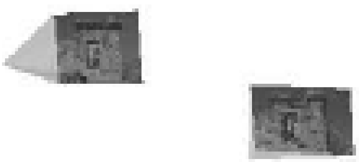
$$x^\top [t]_\times R x = 0$$

$$E \leftrightarrow P, P'$$

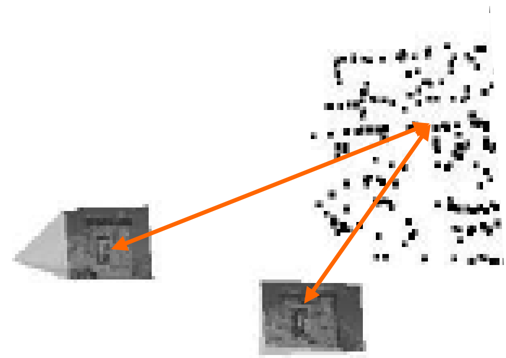
Also how to robustly compute from images

ETH

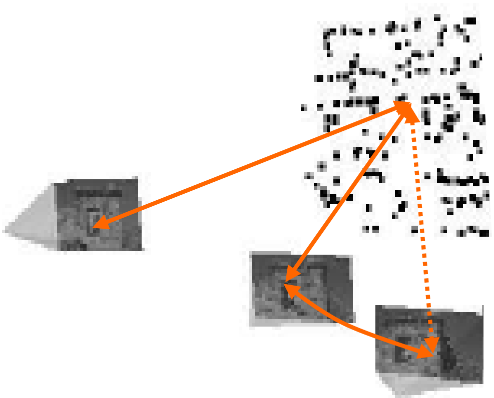
Structure from Motion



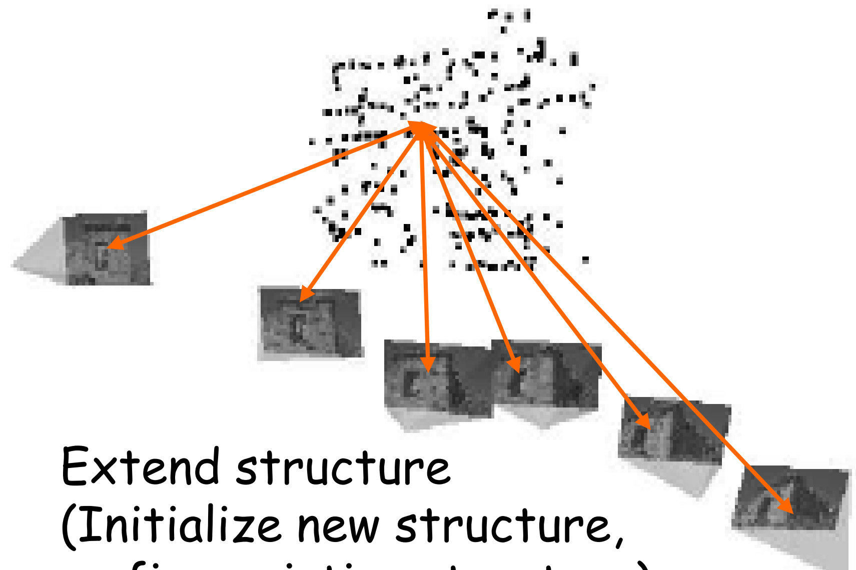
Initialize Motion
(P_1,P_2 compatibel with F)



Initialize Structure
(minimize reprojection error)



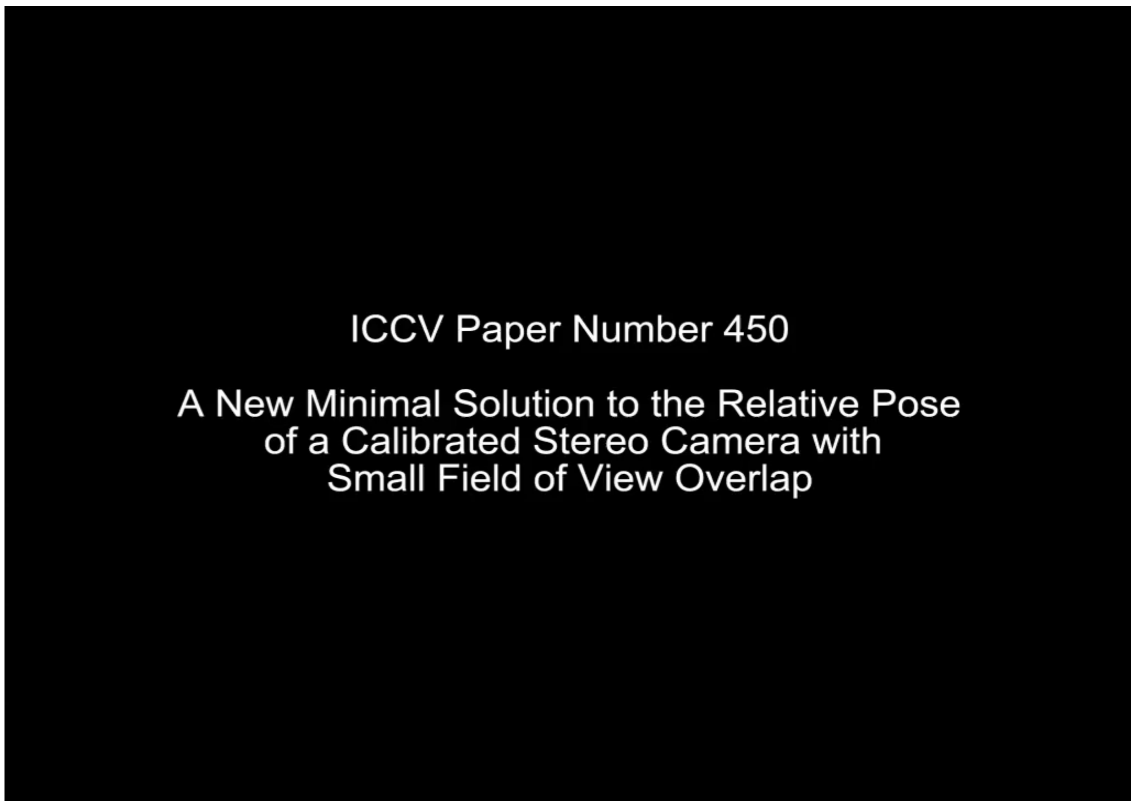
Extend motion
(compute pose through matches seen in 2 or more previous views)



Extend structure
(Initialize new structure, refine existing structure)

Visual SLAM

- Visual Simultaneous Navigation and Mapping

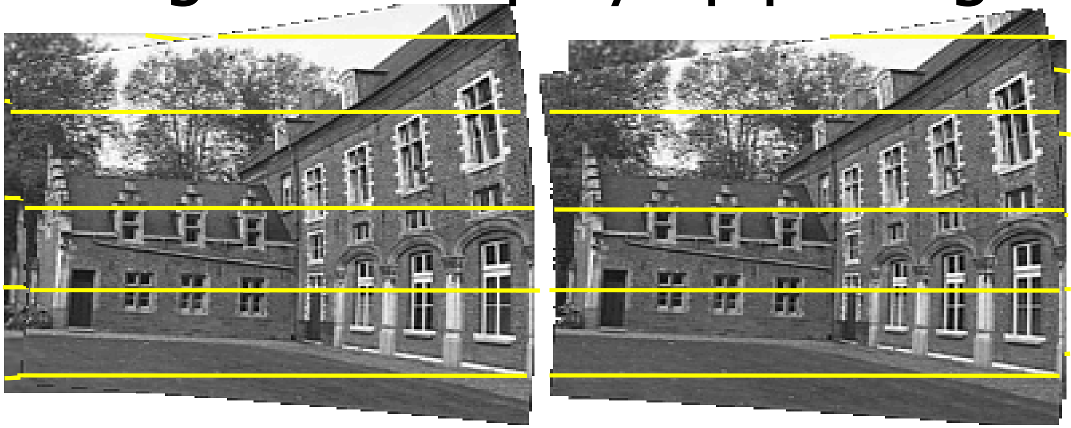


(Clipp et al. ICCV'09)

ETH

Stereo and Rectification

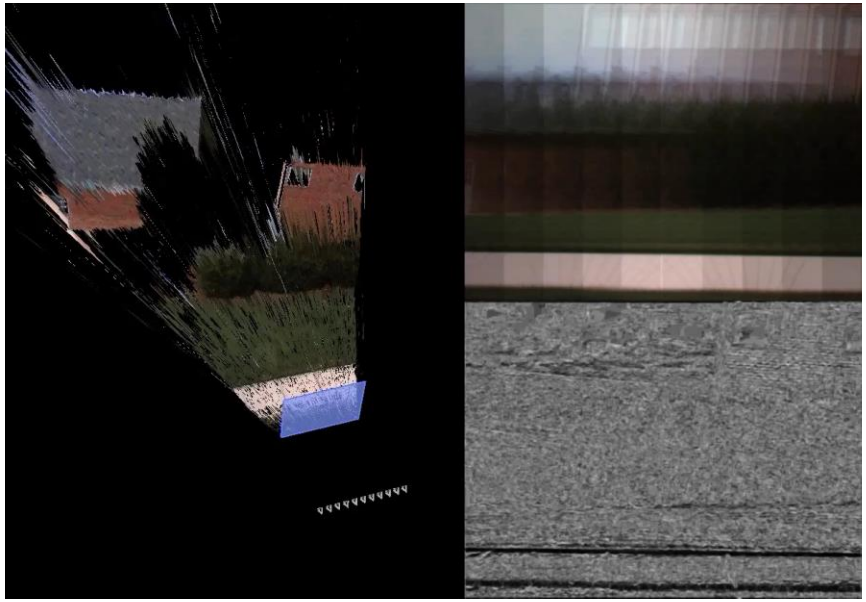
Warp images to simplify epipolar geometry

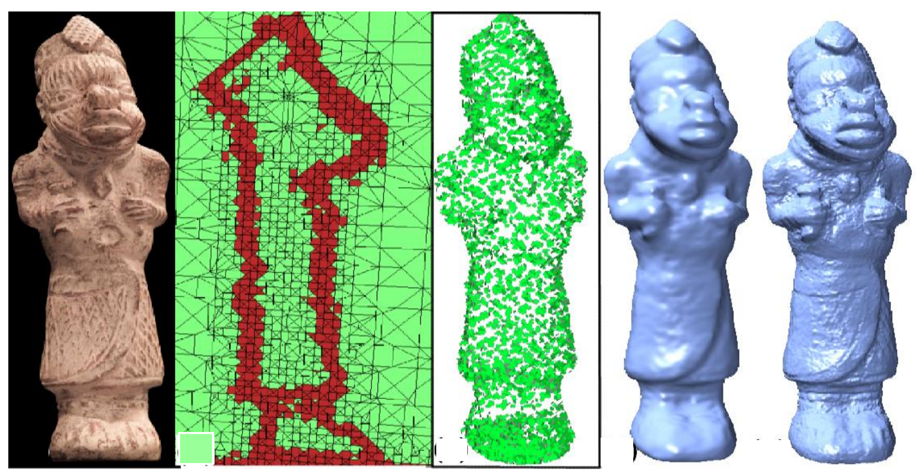


Compute correspondences for all pixels



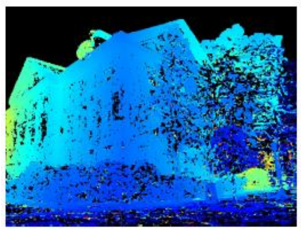
Multi-View Stereo

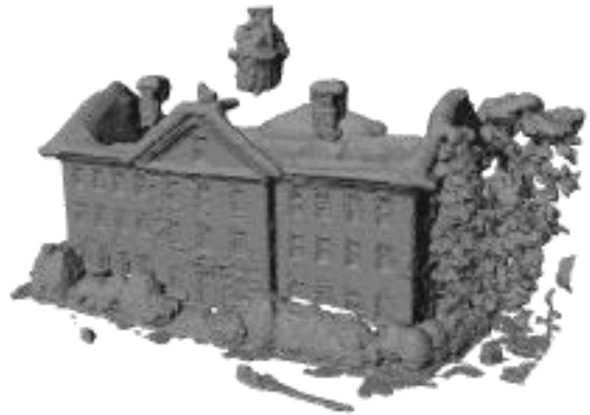




Joint 3D Reconstruction and Class Segmentation

(Haene et al CVPR13)

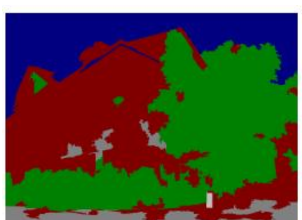


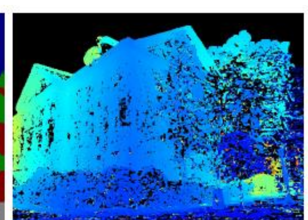


reconstruction only
(isotropic smoothness prior)

joint reconstruction
and segmentation
(ground, building, vegetation, stuff)

- Building
- Ground
- Vegetation
- Clutter

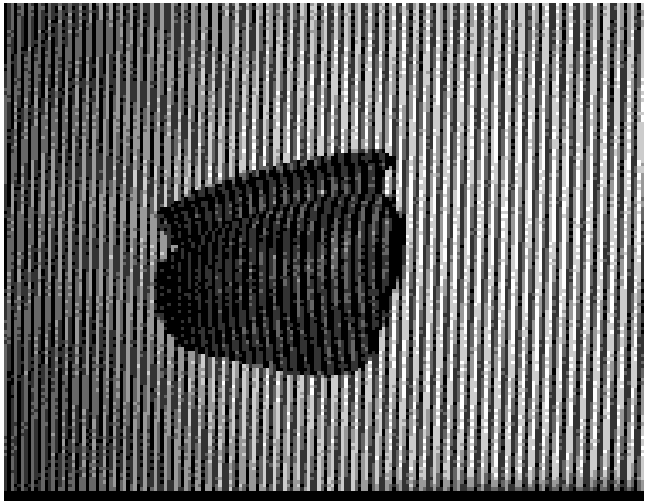




ETH

Structured Light

- Projector = camera
- Use specific patterns to obtain correspondences



ETH

Papers and Discussion

- Will cover recent state of the art
  - Each student team will present a paper (4-5min per team member), followed by discussion
  - “Adversary” to lead the discussion
  - Papers will be related to projects/topics
- Will distribute papers later (depending on chosen projects)

ETH

Projects and reports

- Project on 3D Vision-related topic
  - Implement algorithm / system
  - Evaluate it
  - Write a report about it
- 3 Presentations / Demos:
  - Project Proposal Presentation (week 4)
  - Midterm Presentation (week 8)
  - Project Demos (week 15)
- Ideally: Groups of 3-4 students

Course project example: Build your own 3D scanner!

Example: Bouguet ICCV'98

Project Topics

Your Own Project

Goal: Learn about the techniques presented in the lecture

Description:

Choose your own topic!

Available hardware:

- Google Tango Tablets
- Microsoft HoloLens
- GoPro Cameras
- Intel RealSense Sensor

Requirements / Tools:

Required: Related to 3D Vision / topics of the lecture

Supervisor:

We find one for you

Your Next Steps

- Find a group (ideally: groups of 3-4)
- Find a project (one of ours or your own)
- Topic subscription via doodle in a few days:
  - For questions contact us via the lecture Moodle (preferred) or contact Rémi per email
  - First come first serve!
  - Do **not** contact supervisors directly!
- After topic assignment: talk with your supervisor
- Write a project proposal
- Don’t worry: You’ll get reminders!

ETH