

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, selfdriving cars and AR / VR
 - Visual navigation
 - Sensing / mapping the environment
 - Obstacle detection, ...
- Many further application areas
- A few examples ...





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







Large-Scale Structure-from-Motion

Rome dataset

74,394 images

Video credit: Johannes Schönberger



Virtual Tourism

Photo Tourism Exploring photo collections in 3D

Noah Snavely Steven M. Seitz Richard Szeliski University of Washington Microsoft Research

SIGGRAPH 2006





3D Urban Modeling







UNC/UKY UrbanScape project





3D Urban Modeling













Mobile Phone 3D Scanner







Mobile Phone 3D Scanner

Face Scan Demo

www.aquilaviz.com





Self-Driving Cars

Ego Speed: 49.00 MPH time: 1541.746927000 CAL P 0.60 Y 1.20 R 0.00 deg

Vision fps: 18.11 Draw fps: 18.75 Display fps: 23.70 NL(0.00), E(0.01), F(0.99), TF(0.01), S(0.00) FLP(0.99), FRP(0.01) NRW: FLP(0.00), FRP(0.00)

+0.0001 AUTO_HIGH_BEAM +0.0000 BLINDED +0.0000 RAINING +0.0000 TIRE_SPRAY +0.0001 WET_ROAD

0.1308 CONTROLLED_ACCESS

L:0 R:0 F:1 ON:0 W:7.0 AP:1.8 It0 VS: 52.2 MPH St: 1 merge: 1.0 1 153.2 L



Self-Driving Cars

Autovision for Autonomous Vehicles











Micro Aerial Vehicles

71

Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich





Mixed Reality



Microsoft HoloLens 2





Large-scale visual mapping

cecapital (pu; 181 Nes; 281 001 Nes Pet 01 5:50							
Recording					File Edit View Search Terminal Help		
		Start recording			The for a contraction to the contraction of the con		
✓ Enable recording hotkey		Enable sound	रे notifications		- fm-ff- 1 (edf12 (edf12 = 148 042 2 / 16:55 m/ (f)(2:23).jmg - fm-ff 1 (edf12 (edf12 = 126) fm (f)(2:23).jmg - fm-ff 1 (edf12 (edf12 = 126) fm (f)(2:23).jmg		
Hotkey:	√ Ctrl + Shift +	Alt +	Super +	B *	rer-r-r- 1 certac certa: 37754 bez 4 37:00 obs1001.cm rer-r-r- 1 certac certa: 337554 bez 4 37:00 obs1001.cm rer-r-r-r- 1 certac certa: 330557 \$400120*15*te=studior.ap4		
					- Pu-FF- 1 Cepric Cepric 1982/2090 Deg 8 07-35 Dia-2020-12-00 07.42.18.0p4 - Pu-FF 1 Cepric Cepric 1300/070 Deg 8 07-44 Dia-2020-1-00 07.42.18.0p4		
Information	Preview				-rw-rr- 1 cerric cerric 14173081 Dez 8 07-65 20201200-mix1aTa-tooe.mp4 -rw-rr 1 cerric cerric 13412502 Dez 8 15:10 aurc.xcf		
Total time: 0:00:00	Preview frame rate: 10			0	- nwr-r-r- 1 certric certric 11466156 Dez B 15:26 marc_space.xcf -nwr-r-r- 1 certric certric 106973 Dez B 15:27 marc_scalable_matching.png		
FPS in: 0.00	Note: Previewing requires extra CPU time (espe-	ially at high frame rates).			-n+r-r- 1 cedric cedric 176030 Dez B 15:31 marc (densification png -n+r-r- 1 cedric cedric 1100250 Dez B 13:35 marc (sranble.scf		
Size in: 2560x1440					-rwr-r 1 cedric cedric 1568000 Dez 8 15:35 marc scramble.pm -rwr-r 1 cedric cedric 13058050 Dez 8 16:15 marc.mst.scf		
Size out: 7					invrrir-1 (addric addric) 85501 Dez 8 18618 marcinit.ng invrri-r-1 (addric addric) 95500 Dez 9 8744 wollense.xf		
File name: 7					-ne-rr- 1 cedric cedric 1081514 Dez 9 07:47 sallenses scalable.xcf		
File size: 0 B					- Nerr-r-1 Londriz cedriz 2515700 Dec 9 07:52 sullanse_hbuffle.png		
Bit rate: 0 ops					-net-r-r- 1 cedric cedric 100120 Deg 9 07:36 walleneer generat/watten.pm		
					-rw-r-r- 1 cerric cerric 122805/bez 9 08:03 willtense pp.pmg -rw-r-r 1 cerric cerric 1212880 bez 9 88:03 willtense complete.pmg		
					- "n+ "" - 1 codrig: codrig: 176464787 082 9 21:22 2/2012/01-bigsbb opt - "n+ "" - 1 codrig: codrig: 50752718 082 10 11:43 22001218-bigsbb cloud:ap4		
					- nu+nv+r- 1 cedric cedric 15442 Dez 14 20:02 ibsi.txt - nu+r 1 cedric cedric 97244655 Dez 15 09:22 2020/23-bigsbb-ba.mp4		
					-nurrr 1 cedric cedric 2008 Dez 16 09:20 holotrauba hist.png		
					- Pr		
					drum vr. all control control dong bas a local both un		
					-rw-r-r- i cerric cerric i 353350 Jan 11 21530 design_env.eng -rw-r-r- i cerric cerric 127323 Jan 11 21532 design_eng		
					-rwir-i i cerric cerric 138847 Jan 11 Bi-08 now (dosign, svg -rwir-i-r i cerric cerric 13842 Jan 12 Bi-08 now (dosign, pug		
					-rw-rv-r- 1 cedric cedric 5482652 Jan 13 12:51 20(10)(3) (bb withoreavy)s.mp4 -rw-rv-r- 1 cedric cedric 37460 Jan 14 17:29 addinulus.mp4		
					nern-r-t- 1 cedric cedric 2003 3an 21 00:33 Constitution_error.psg		
					-ner-r-r- 1 cedric cedric 3394695 Jan 22 10:30 untitled.ebj		
					officer and a control control of and a control of the control of t		
					dnaxrxr-x 2 cedrac cedrac 4096 Feb 3 08:21 Pictures dnaxrxrx 19 cedrac cedrac 4096 Feb 3 10:10 data		
					dnxx+x+x 2 cedric cedric 4096 Feb 314:42 hubernetes -nx+r-r- 1 cedric cedric 1620(301 Feb 316:55 104:202) 16:55:00.mpt		
					codr1cd1pst-1-5 mv -/bla-2021-02-03_10.55.00.mp4 20230203_bigsbb.mp4 Codr1cd1pst-1-5 mut11us.		
					Hautile-Shara-Ressapr: [6:57:55:55]: Called "met usershare info" but it failed: Failed to execute child process "net" (No such file or directory) De meting name: 0 & Mathemari (and De Joséfendebase)		
					[000050711a40500] main libu(: Ramaing Vic with the default interface. Use 'cvlc' to use vic withest interface.		
					tower/contributers accord be stoped from another thread		
					VCC socia player J.O.B Vollari (revision J.D.B-0-gresouosia) [00005562434580] main labic: Rumning vic with the default interface. Use 'cvic' to use vic without interface.		
					[00007fbclBclBc32ac9] avcodec decoder: Using NVIDEA VDPAU Driver Shared Library 450-102.04 The Dec 29 06:43:13.UTC 2020 for hardware decoding		
					(nautlis:8340): GLB-610-WBRING **: 16:59:17 050: unable to send notifications through org.freedesktop.Notifications: GGBUS.Error.org.freedesktop.BBUS.Error.ServiceBinknown: The laws org.freedesktop.Butifications was not provided by any service files		
					000ječt:=00ject:Tamers cannot be stopped from another thread cedircDMP3-3 simplesreenresorder		
					Could not load system translator.		
					SimpleScreenRecorder 0.3.8		
					Ort header 5 9.2.1 LB 5.6.5		
					Libavional: mader 37:83:100, Lib 37:83:100 Libavione: head 7:107:100, Lib 57:107:100, Lib 57:107:100		
					Libracia: header 3.7.13.109. Ltb 3.7.67.800 Libracia: header 4.8.109 (Ltb 4.8.100		
		Start pr	VIEW		[AUSINgst:cetStore(s): CHP features: Hes size size) size) size() size() avx avx2 hes] hes2 [AUSINgst:cetStore(size)] Cenerating source list		
Log					[ALSATnput::detSourceList] Found card: [he:0] HDA Intel POH [ALSATnput::detSourceList] Found device: [hvi.00] HDA Intel POH; ALC22] Analog		
[PageRecord::StartPage] Starting page					IALSATnuvti(GESGUTCELIST)Faund Gevice: [hrv:0.2] HGA Intel PCH: ALC221 Alt Analog IALSATnuvti(GESGUTCELIST)Faund Card: [hrv:1] HGA NVidia		
[PageRecord::StartPage] Sta	[PageRecord::StartPage] Started page.				[Mulsahdiologut:detSourceList] Comercing source list [SourceAmmerch1]bod] Found source: [also putter to :0000 BD 17.3 avalop-sterem provider] Monitor of Ruilt-in Audio Sterem		
					[SourceManesCallback] Found source: [alss_input.pl:/0000_00] [1.3.amolog-streed_Built-In Audio Amolog Streed Built-In Audio Amolog S		
					(Poper enternation of the cases) output per per		
					IPageHead 1.1147Page1.1547146 page. IPageHead 1.5147Page1.5447146 page.		
					[Muxar:lnit] Using format mp4 (MP4 (MPEG-4 Part 14)). [Muxar:ldftrem] Using (mode: Llbix)64 (Llbix24 f MXC / MPEG-4 part 10).		
					VideoEncoder: PrepareStream: Using pixet format rv12. TTab/24 @ 055908727400T using SeP-/13		
	Cancel recording		Saye records	na	Tlabx264 @ 0x599903734600 wiing cpu capabilities: MMC2 SSE2Fast SSE3 SSE4.2 AVX FMA3 BM12 AVX2 AVX512		
			save records	·			





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

Control Humanoid Robot with Kinect





Autonomous Micro-Helicopter Navigation







Use Kinect to map out obstacles and avoid collisions





Dynamic Reconstruction

Dynamic Fusion:

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

Richard Newcombe, Dieter Fox, Steve Seitz

Computer Science and Engineering, University of Washington



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













ZUH Zürich







Performance Capture



(Oswald et al. ECCV 14)



Motion Capture







Interactive 3D Modeling

Interactive 3D Architectural Modeling from Unordered Photo Collections Paper # 0062

(Sinha et al. Siggraph Asia 08)



collaboration with Microsoft Research (and licensed to MS)



Scanning Industrial Sites





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







Scanning Cultural Heritage









Some more scanner examples

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





Cultural Heritage

Stanford's Digital Michelangelo









Digital archive Art historic studies





Archaeology



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





Forensics

Crime scene recording and analysis











Forensics









DiscoverEye













3D Vision Course Team



Marc Pollefeys **CNB G 105**



Daniel Barath CNB G 103.2 marc.pollefeys@inf.ethz.ch 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%



Materials

Slides and more <u>http://www.cvg.ethz.ch/teaching/3dvision/</u>

Also check out on-line "shape-from-video" tutorial: <u>http://www.cs.unc.edu/~marc/tutorial.pdf</u> <u>http://www.cs.unc.edu/~marc/tutorial/</u>

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



Geometric transformations in 2D and 3D







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



Epipolar Geometry







Fundamental matrixEssential matrix $x^{\top}Fx = 0$ $x^{\top}[t]_{\times}Rx = 0$ $F \leftrightarrow P, P'$ $E \leftrightarrow P, P'$

Also how to robustly compute from images





Structure from Motion





Initialize Motion $(P_1, P_2 \text{ compatibel with } F)$



Extend motion (compute pose through matches seen in 2 or more previous views) Extend structure (Initialize new structure, refine existing structure)

Initialize Structure

(minimize reprojection error)



Visual SLAM

Visual Simultaneous Navigation and Mapping

ICCV Paper Number 450

A New Minimal Solution to the Relative Pose of a Calibrated Stereo Camera with Small Field of View Overlap



(Clipp et al. ICCV'09)





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







Structured Light

- Projector = camera
- Use specific patterns to obtain correspondences







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)





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



3D Vision, Spring Semester 2018



Computer Vision and Geometry Lab



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!

