3D Computer Vision

Radim Šára     Martin Matoušek

Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

https://cw.felk.cvut.cz/doku.php/courses/a4m33tdv/
http://cmp.felk.cvut.cz
mailto:sara@cmp.felk.cvut.cz
phone ext. 7203

rev. October 3, 2017

Open Informatics Master’s Course
Part I

Course Overview
3D Vision is Not Just about 3D Point Clouds

- today, we have laser-based rangefinders (eg. LiDARs)
- figure: point cloud obtained from a vehicle with 4 LiDARs on its roof
- color = height
- this course focuses on obtaining such results (and more) by means of passive sensors
This Course: Structure from Motion & Surface Reconstruction

images + some knowledge about cameras → cameras in 3D + 3D points

36 of 237 images of a memorial

Typical phases of a processing pipeline:
1. finding sparse image correspondences
2. recovering camera poses
3. finding dense correspondences → 3D point clouds
4. surface reconstruction
Phase 1: Sparse Image Correspondences

image features, their descriptors, matches and correspondences

- matches $\sim$ visually similar
- correspondences $\sim$ visually similar and geometrically consistent (yellow)
- finding correspondences must cope with ambiguity
- 5 correspondences determine the relative orientation and translation direction between the ('calibrated') cameras
Phase 2: Semi-Dense Correspondences and 3D points (\( \equiv \) “Structure”)

- sensing depth from a single moving camera, 30 fps data stream
  - standard automotive wide-angle sensor – reversing camera
- moving videocamera \( \sim \) time constraint on image match evolution \( \sim \) ‘optical flow’

- standard term: SfM (Structure from Motion)
- problems with moving objects (wrong depth)
Phase 2: Recovering Camera Poses (="Motion")

- reversing camera on a car, 30fps; error against RT 3000 GPS system (red)  
  no fusion with GPS!

Scene I

Scene II

Some applications:
- visual odometry
- SLAM

- the drift is reduced if the correspondences linking camera pairs form a dense graph, not a chain like here

- 1 km, 5% accumulated drift
- measures elevation
- bad lighting conditions
- bad scene
Phase 3: Dense 3D Point Cloud by Stereovision

sometimes stereo seems easy

sometimes it is difficult

input images

Malmö Högskola, Centrum för teknikstudier

• the result is a dense 3D point cloud (color = range)

typically $10^6 - 10^9$ 3D points
Phase 4: Surface Reconstruction

cameras + point cloud + images $\rightarrow$ triangulated surface

- we will not explain surface reconstruction in this course
- but you will be able to use one of the popular algorithms
Thank You