Application of Affine Transformations for 3D Computer Vision

Levente Hajder

Faculty of Informatics, Eötvös Loránd University

National Research, Development and Innovation Office Hungary

PROGRAM FINANCED FROM THE NRDI FUND

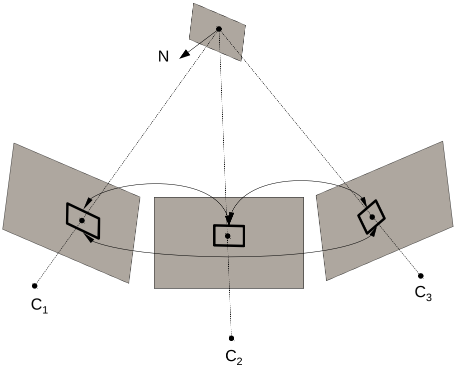
Introduction

- Cameras are essential in visual perception.
- Estimation of extrinsic camera parameters, i.e. visual odometry, is a basic problem in both
  - computer vision and
  - robotics.
- Our work deals with the utilization of affine transformations for stereo pose estimation,
  - instead of using only point correspondences.
- Possible application area: autonomous driving.
  - Autonomous cars are equipped with digital cameras.
  - Autonomous driving is very popular.

Project Plan

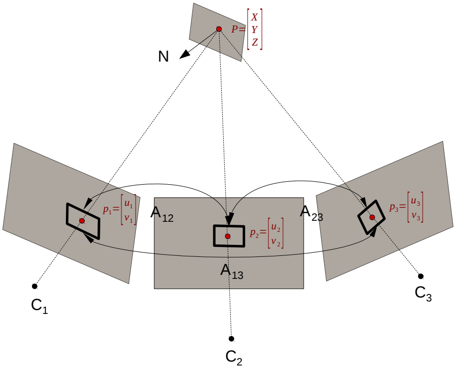
- Goal: Develop algorithms for 3D Computer Vision.
- Principal application area: vision systems for autonomous vehicles.
- Plan of research:
  - **First semester.** *Visual odometry using affine transformations.*
  - **Second semester.** *Estimation of affine transformations between images.*
  - **Third semester.** Segmentation of planar surfaces.
  - **Fourth semester.** Special surfaces: horizontal and vertical planes.
  - **Fifth semester.** Segmentation of moving objects.
  - **Sixth semester.** Camera-based free-space detection.

Motivation: Three-view Geometry of a Surflet



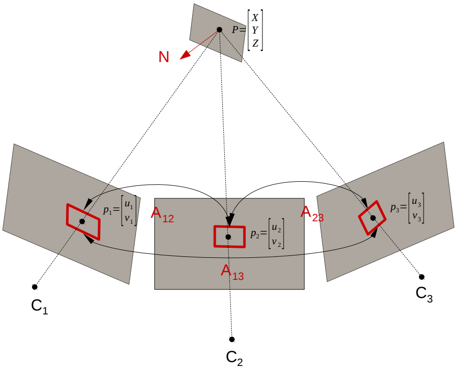
- Perspective (pinhole) camera applied

Point Correspondence (PC)-based Approaches



- Common approaches: 3D parameters estimated from patch centers

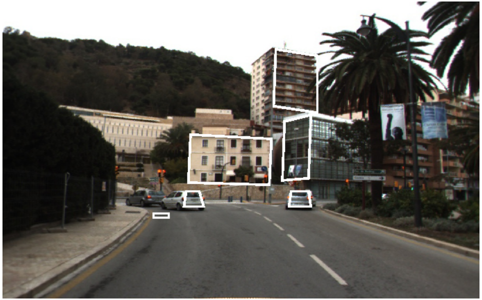
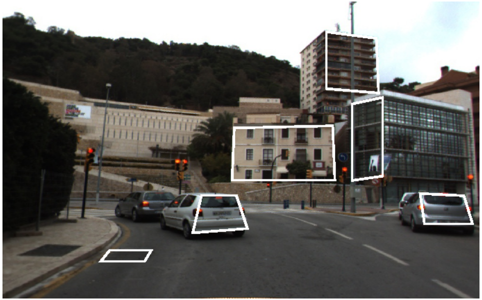
Affine Correspondence (AC)-based Approaches



- Novel Approach: Patch Deformation Considered.

$\rightarrow$ 3D motion / surflet normal partially decoded in affinities.

Motivation: Example for AC-utilization

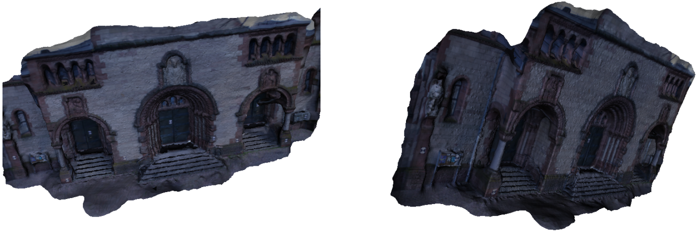
 

- Two frames taken by a car-mounted camera. It is trivial that size of patches depends on the distance of the objects.

Literature Overview

- PC-based solution dominates the literature, e.g. [Hartley&Zissermann].
- ACs can be applied for estimating
  - Surface normals: [Köser PhD 2009], [Barath&Hajder CVWW2014+VISAPP2015]
  - Homographies: [Barath&Hajder PRL 2016]
  - Fundamental matrix:
    - from 3 ACs [Bentolila & Francos CVIU 2014]
  - Essential matrix (relative pose):
    - from 2 ACs: [Raposo et al. CVPR 2016]
    - from 2ACs, focal length also estimated: [Barath et al. CVPR 2017]
  - Camera parameters [Eichhardt & Hajder ICPR 2016] [Eichhardt & Chetverikov ECCV 2018]
  - Structure and Motion (SfM) [Eichhardt & Hajder ICCV WS 2017]

Application #1: Perspective Structure from Motion



- Affine transformation can be detected for each images pair $\rightarrow$ more information available from affine transformations.
  - $\rightarrow$ Occlusion can be handled.
- Spatial points„ surface normals and camera parameters jointly estimated.
  - $\rightarrow$ Numerical method, so-called 'bundle adjustment', is adapted .

$\rightarrow$ Ivan Eichhardt, Levente Hajder: Computer Vision Meets Geometric Modeling: Multi-view Reconstruction of Surface Points and Normals using Affine Correspondences. International Conference on Computer Vision, M3D Workshop, 2017.

Applications #2: Plane Fitting via Homography Estimation

- Transformation between correponding planes in images represented by 3×3 homography matrix
- Homography can be estimated using affine correspondences.
- Proposed algorithms:
  - General case: homography estimation from two affine correspondences (ACs)
  - Only a single AC required if epipolar geometry is known.

  $\rightarrow$ Epipolar geometry can be estimated both from point and affine correspondences.

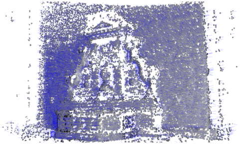
$\longrightarrow$ Daniel Barath, Levente Hajder: A theory of point-wise homography estimation. Pattern Recognition Letters 94: 7-14 (2017)

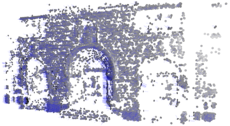
Applications #2: Plane Fitting via Homography Estimation

- A single affine correspondence yields a plane-plane homography $\rightarrow$ a hypothesis for a plane obtained.
- Feature points corresponding to the homography can be segmented.
- It can be inserted into a multi-model fitting framework.

$\rightarrow$ Daniel Barath, Jiri Matas, Levente Hajder Multi-H: Efficient recovery of tangent planes in stereo images. British Machine Vision Conference 2016.

Application #3: Surface Normal Estimation



- Normal vectors can be computed if
  - $\rightarrow$ cameras are calibrated;
  - $\rightarrow$ point correspondences and
  - $\rightarrow$ and related affine transformation are given.

$\rightarrow$ Daniel Barath, József Molnár, Levente Hajder: Novel Methods for Estimating Surface Normals from Affine Transformations. VISIGRAPP (Selected Papers) 2015: 316-337

Theoretical Backgound: General Epipolar Geometry

- Essential matrix consists of the extrinsic camera parameters:
  - Translation vector without scale **t** (2 DoF)
  - Rotation matrix **R** (3 DoF)

$$\mathbf{E} = [\mathbf{t}]_x \, \mathbf{R}$$

- Relationship of fundamental and essential matrices:

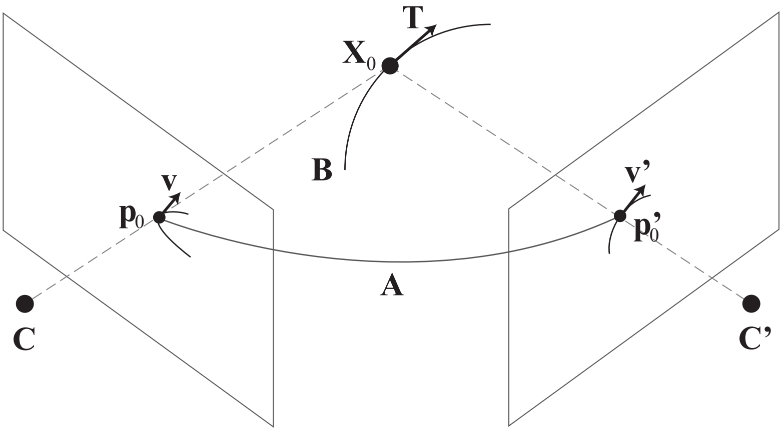
$$\mathbf{F} = \mathbf{K}_1^{-T} \mathbf{E} \mathbf{K}_2^{-1}$$

Epipolar geometry + Point Correspondences

- Each point pair yields one well-known equation for fundamental/essential matrices:
  - Point correspondences in images given in homogeneous form:
$$\mathbf{p}_1^T = [u_1 \quad v_1 \quad 1] \quad \mathbf{p}_2^T = [u_2 \quad v_2 \quad 1].$$
  - They should fulfill:
$$\mathbf{p}_1^T \mathbf{F} \mathbf{p}_2 = 0.$$

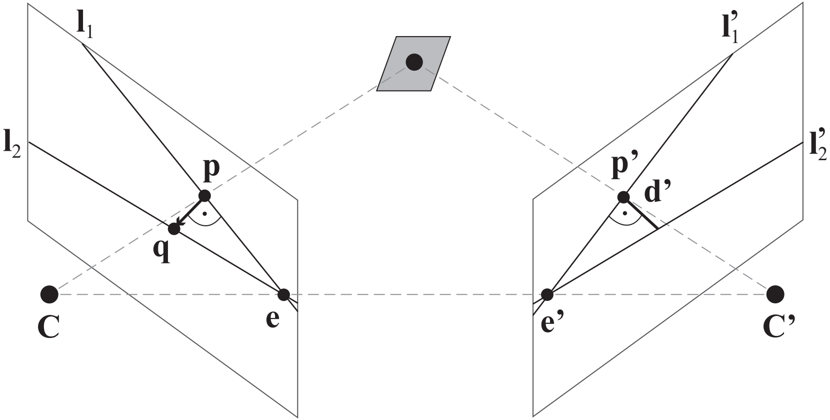
Epipolar geometry + Affine Transformations

- Point correspondences are locations, affine transformations determines
  - Directions of lines, and
  - Scale along these directions.



Epipolar geometry + Affine Transformations

- Fundamental matrix determines the scale along the perpendicular direction of epipolar lines
- $\mathbf{A}^{-T}\left(\mathbf{F}^{T}\mathbf{p}_2\right)_{1:2} = -\left(\mathbf{F}\mathbf{p}_1\right)_{1:2}$, [Barath & Hajder CVPR 2017]



Workpackage #1: Estimation of Motion Trajectory of a Vehicle-mounted Camera

- Planar motion:
  - Road is flat.
  - Camera is mounted on the vehicle.
  - Image plane is perpendicular to the ground.
- Extrinsic camera parameters are special in this case:

$$\mathbf{t} = \begin{bmatrix} x \\ 0 \\ y \end{bmatrix} = \nu \begin{bmatrix} \cos\beta \\ 0 \\ \sin\beta \end{bmatrix}, \quad R = \begin{bmatrix} \cos\alpha & 0 & \sin\alpha \\ 0 & 1 & 0 \\ -\sin\alpha & 0 & \cos\alpha \end{bmatrix}$$

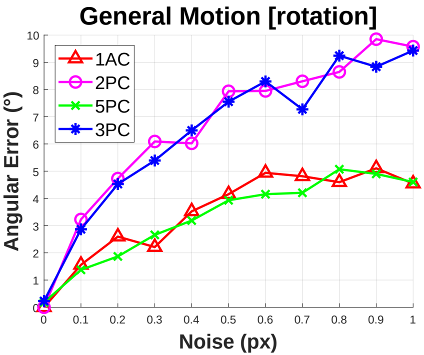
- Essential matrix:

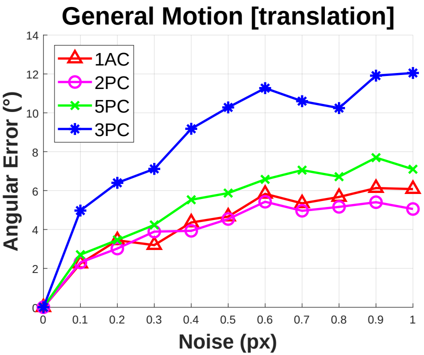
$$\mathbf{E} = [\mathbf{t}]_x \mathbf{R} \sim \begin{bmatrix} 0 & -\sin\beta & 0 \\ \sin(\alpha+\beta) & 0 & -\cos(\alpha+\beta) \\ 0 & \cos\beta & 0 \end{bmatrix}$$

Proposed methods

- Camera motion estimation estimation from one affine correspondences.
- Novelties:
  - State of the art methods require at least two point correspondences.
  - Robustification can be implemented very efficiently.
  - Proposed methods are very fast.
- Two methods introduced:
  1. Calibrated case
     - Camera intrinsic parameters, e.g. focal length, are known.
     - Optimal solution can be obtained via the roots of a six-order polynomial.
  2. Uncalibrated case
     - Rapid algorithm, optimality cannot be guaranteed.

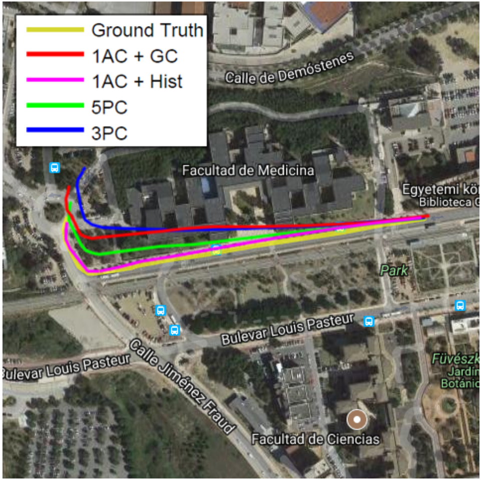
Synthesized Tests: General Motion





- Estimation error for the angles, general vehicle motion.

Real Tests: Malaga dataset



- Trajectories computed by forming many stereo pairs from a video.

Real Tests: Malaga dataset

- Robustly estimated trajectories. Vehicle speed retrieved from GPS.
- Can work real-time ($\sim$ 5 FPS) on modern GPUs. Affine matchers included.

Workpackage #2: Estimation of Affine Transformations between Images

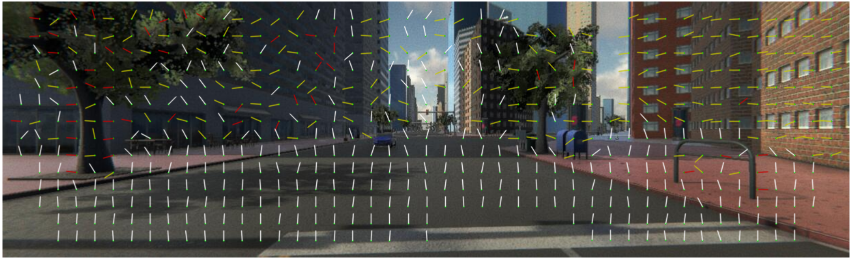
- For the aforementioned methods, affine transformations are the input.
  - Thus, results highly depend on quality of affine transformations.
- Two affine detector methods proposed:
  - An affine transformation can be estimated from one point correspondence and two directions if epipolar geometry is know,
  - or transformations are estimated from the optical flow.
- Both problems can be written as six-dimensional linear systems of equations.
  - Solution is straightforward.
  - Proposed algorithms are very fast.

3D Stereo Reconstruction Pipeline

- Affine frames used as input, containing
  - locations and
  - related affine transformations.
- Calibrated cameras are assumed.
- Pipeline stages:
  - Epipolar geometry estimated via point correspondences (ASIFT).
  - Affine transformations estimated by one of proposed algorithms.
  - 3D coordinates calculated by Hartley-Sturm triangulation.
  - Surface normal estimated by our method [CVWW2014].

Autonomous simulator : LG-SVL

- Unity-based multi-robot simulator developed by LG Electronics America R&D Center.
- Cars can be equipped with many sensors, e.g. cameras, LiDARs, radars.
- Normal vectors computed from affine transformations by proposed pipeline. Affine transformations computed from optical flow, applying one of the proposed methods.



Summary

- Exploiting affine transformations are overviewed. They can be applied for several 3D vision problems.
  - Surface normal estimation.
  - Plane fitting.
  - Full 3D reconstruction.
- Two novel sets of algorithms proposed:
  1. Camera motion estimation for vehicle-mounted cameras.
  2. Accurate estimation of affine transformations.

Publications

- Strongly connected papers
  - Levente Hajder, Dániel Baráth. *Least-squares Optimal Relative Planar Motion for Vehicle-mounted Cameras* – Accepted for International Conference on Robotics and Automation, 2020
  - Levente Hajder, Dániel Baráth. *Relative planar motion for vehicle-mounted cameras from a single affine correspondence* – Accepted for International Conference on Robotics and Automation, 2020
  - Nghia Le Minh, Levente Hajder. Affine Transformation from Fundamental Matrix and Two Directions. VISIGRAPP (4: VISAPP) 2020: 819-826.

Publications

- Other papers
  - Zoltán Pusztai, Gergő Gál, Levente Hajder. Quantitative Comparison of Affine Feature Detectors Based on Quadcopter Images. VISIGRAPP (Revised Selected Papers) 2019: 430-455.
  - Levente Hajder, Tekla Tóth, Zoltán Pusztai. Automatic Estimation of Sphere Centers from Images of Calibrated Cameras. VISIGRAPP (4: VISAPP) 2020: 490-497.
  - Tekla Tóth, Zoltán Pusztai, Levente Hajder. *Automatic LiDAR-Camera Calibration of Extrinsic Parameters Using a Spherical Target*– Accepted for International Conference on Robotics and Automation, 2020
  - Antal Hiba, Levente Sántha Tamás, Zsedrovits Levente Hajder, and Ákos Zarándy. Onboard Visual Horizon Detection for Unmanned Aerial Systems with Programmable Logic ELECTRONICS, 9 (4). 2020

Thank you for your attention.

hajder@inf.elte.hu