Computer Vision II - Lecture 1

Introduction

15.04.2014

Bastian Leibe
RWTH Aachen
http://www.vision.rwth-aachen.de
leibe@vision.rwth-aachen.de

Organization

• Lecturer
  ➢ Prof. Bastian Leibe (leibe@vision.rwth-aachen.de)

• Teaching Assistants
  ➢ Stefan Breuers (breuers@vision.rwth-aachen.de)
  ➢ Umer Rafi (rafi@vision.rwth-aachen.de)

• Course webpage
  ➢ http://www.vision.rwth-aachen.de/teaching/
    → Computer Vision2
  ➢ Slides will be made available on the webpage
  ➢ There is also an L2P electronic repository

• Please subscribe to the lecture on the Campus system!
  ➢ Important to get email announcements and L2P access!

Language

• Official course language will be English
  ➢ If at least one English-speaking student is present.
  ➢ If not... you can choose.

• However...
  ➢ Please tell me when I’m talking too fast or when I should repeat
    something in German for better understanding!
  ➢ You may at any time ask questions in German!
  ➢ You may turn in your exercises in German.
  ➢ You may answer exam questions in German.

Course Webpage

http://www.vision.rwth-aachen.de/teaching/

Exercises and Demos

• Exercises
  ➢ Typically 1 exercise sheet every 2 weeks (Matlab based)
  ➢ Hands-on experience with the algorithms from the lecture.
  ➢ Send in your solutions the night before the exercise class.

• Teams are encouraged!
  ➢ You can form teams of up to 3 people for the exercises.
  ➢ Each team should only turn in one solution.
  ➢ But list the names of all team members in the submission.
Textbooks

- No single textbook for the class.
- Some basic material is covered in the following book:
  
  D. Forsyth, J. Ponce  
  Computer Vision - a Modern Approach  
  Prentice Hall, 2002  

(available in the library's "Handapparat")

- We will mostly give out research papers
  - Tutorials for basic techniques
  - State-of-the-art research papers for current developments

Computer Vision II

- We will build upon the basics from previous lectures
  - Computer Vision I
  - Machine Learning

- However,
  - If you haven’t heard those lectures yet, you may still attend and benefit from this lecture.
  - But please look at the available online material from those lectures to get additional background on the basic techniques.
  - I will regularly point out what background to repeat.

How to Find Us

- Office:
  - UMIC Research Centre
  - Miles-van-der-Rohe-Strasse 15, room 124

- Office hours
  - If you have questions to the lecture, come to Stefan/Umer or me.
  - My regular office hours will be announced (additional slots are available upon request)
  - Send us an email before to confirm a time slot.

Questions are welcome!

Why Computer Vision?

Cameras are all around us...

Images and Video Are Everywhere...

- Personal photo albums
- Movies, news, sports
- Internet services
- Surveillance and security
- Mobile and consumer applications
- Medical and scientific images

Computer Vision I Covered the Basics...

- Image Processing Basics
- Segmentation
- Local Features & Matching
- Object Recognition and Categorization
- 3D Reconstruction
Computer Vision II is all about Motion

How can we track an object’s motion over time?

B. Leibe

Motion Requires Video

- A video is a sequence of frames captured over time
- Our image data is a function of space \((x, y)\) and time \((t)\)

What Is Tracking?

- Goal
  - Estimate the number and state of objects in a region of interest
- Number
  - 1: Single-target tracking
  - 0 or 1: Detection and tracking
  - N: Multi-target detection and tracking

What distinguishes tracking from “typical” statistical estimation (or machine learning) problems?

- Typically a strong temporal component is involved.
- Estimating quantities that are expected to change over time (thus, expectations of the dynamics play a role).
- Interested in current state \(S_t\) for a given time step \(t\).
- Usually assume that we can only compute information seen at previous time steps \(1, 2, ..., t-1\).
- Usually want to be as efficient as possible, even “real-time”.

\[\Rightarrow\] These concerns lead naturally to recursive estimators.
Types of Tracking

- Single-object tracking focuses on tracking a single target in isolation.
- Multi-object tracking tries to follow the motion of multiple objects simultaneously.
- Articulated tracking tries to estimate the motion of objects with multiple, coordinated parts.
- Active tracking involves moving the sensor in response to motion of the target. Needs to be real-time!

Applications: Safety & Security

- Autonomous robots
- Driver assistance
- Monitoring pools (Poseidon)
- Pedestrian detection
- Surveillance

Applications: Vision-based Interfaces

- Games (Microsoft Kinect)
- Assistive technology systems
- Camera Mouse
- Boston College
Applications: Visual Special Effects

The Matrix

MoCap for Pirates of the Caribbean, Industrial Light and Magic (Source: S. Seitz)

Why Are There So Many Papers on Tracking?

• Because what kind of tracking “works” depends on problem-specific factors...

Factors: Discriminability

• How easy is it to discriminate one object from another?

Appearance models can do all the work
Constraints on geometry and motion become crucial

Factors: Frame Rate

Gradient ascent (e.g. mean-shift) works OK
Much harder search problem. Good data association becomes crucial.

Other Factors

• Single target vs. multiple targets
• Single camera vs. multiple cameras
• On-line vs. batch mode
• Do we have a good generic detector? (e.g., faces, pedestrians)
• Does the object have multiple parts?
• ...

Elements of Tracking

• Detection
  Find the object(s) of interest in the image.
Elements of Tracking

- Detection
  - Find the object(s) of interest in the image.
- Association
  - Determine which observations come from the same object.

Outline of This Lecture

- Single-Object Tracking
  - Background modeling
  - Template based tracking
  - Color based tracking
  - Contour based tracking
  - Tracking by online classification
  - Tracking-by-detection
- Bayesian Filtering
- Multi-Object Tracking
- Articulated Tracking

Applications: Visual Surveillance

Background Modeling

- Learning a statistical model of background appearance

Template Tracking

- Lucas-Kanade registration applied to tracking \(\Rightarrow\) KLT
Color-based Tracking

- Mean-Shift Tracking


Model-based Tracking

- Tracking lines on the object given a 3D model


Applications: Tracking Faces for AR

- Flexible models for an entire class of objects


Geometrically Constrained Level-Set Tracking

- Encode geometric constraints into region tracking
- Constrained homography transformation model
  - Translation on the ground plane
  - Rotation around the ground plane normal

⇒ Input for high-level tracker with car steering model.

Contour-based Tracking

- Level Set Contour Tracking

H. Grabner, M. Grabner, H. Bischof, *Real-time Tracking via On-line Boosting*, BMVC'06.
Outline of This Lecture

- Single-Object Tracking
- Bayesian Filtering
  - Kalman Filters, EKF
  - Particle Filters
- Multi-Object Tracking
- Articulated Tracking

Image sources: Andreas Ess

Multi-Person Tracking


Applications: Tracking Sports Players

- Automatic player tracking for sports scene analysis
  > Several companies active in this area...

Applications: Pedestrian Safety in Cars

Predicting Behavior of “Dynamic Obstacles”
Applications: Mobile Robot Navigation

source

link to the video

Most Recent Version (Demo available)

- Kinect-based head-worn setup
  - Person detection + Tracking + Visual odometry + GP estimation
  - Result: 20-35 fps on single CPU core (no GPU involved!)
  - 15 fps with additional far-range detector (on the GPU)

Mobile Tracking in Densely Populated Settings

- Tracking based on stereo depth only, no detector verification

Analyzing Person-Object Interactions

Outline of This Lecture

- Single-Object Tracking
- Bayesian Filtering
  - Kalman Filters, EKF
  - Particle Filters
- Multi-Object Tracking
  - Multi-hypothesis data association
  - MHT, JPDAF, MCMCDA
  - Network flow optimization
- Articulated Tracking
  - GP body pose estimation
  - Model-based tracking, AAMs
  - Pictorial Structures

Articulated Person Tracking

- Tracking and interpreting detailed body motion.

Combination: Articulated Multi-Person Tracking

- Multi-Person tracking
  - Recovers trajectories and solves data association
- Articulated Tracking
  - Estimates detailed body pose for each tracked person

Applications: Facial Animation Transfer

Face/Off: Live Facial Puppetry

PaperID 102

T. Weise, S. Bouaziz, H. Li, M. Pauly: Realtime Performance-based Facial Animation, SIGGRAPH 2011

Commercialized by faceshift
https://www.youtube.com/watch?v=85m40eF3C9E

Questions?