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.

Organization

- Structure: 3V (lecture) + 1Ü (exercises)
  - 6 EECS credits
  - Part of the area "Applied Computer Science"
- Place & Time
  - Lecture: Tue 14:15 - 15:45 UMIC 025
  - Lecture/Exercises: Thu 14:15 - 15:45 UMIC 025
- Exam
  - Planned as oral exam
  - We'll propose a list of dates towards the end of the semester

Course Webpage

Tentative Schedule

Date	Topic	Content	Slides	Related Material
08.04.14	*no class (Introduction week)*	-		
10.04.14	*no class (Introduction week)*	-		
15.04.14	**Introduction**	What is Tracking?	pdf, fullpage	
17.04.14	*Exercise 0*	*Intro Matlab*		Matlab resources Presentation Exercise Content
22.04.14	**Background Modeling**	MoG Background Model, Online Adaptation, Non-parametric Models		
24.04.14	**Template based Tracking**	LK Tracking, fast template matching, Affine LK, Line Tracking, Model based Tracking		
29.04.14	**Color based Tracking**			
01.05.14	*no class (May 1st)*	-		

Thursday: Matlab tutorial

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

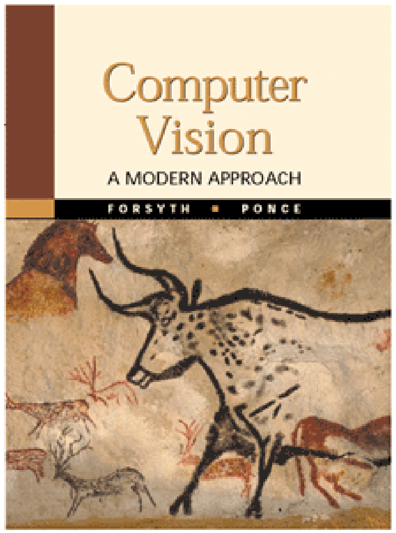
Exercises and Demos

- 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
  - Mies-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…





Slide credit: Kristen Grauman

Images and Video Are Everywhere…

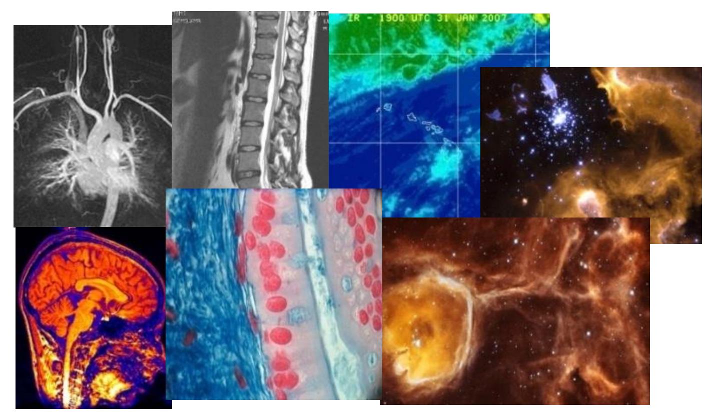
Personal photo albums

Movies, news, sports

Internet services

Surveillance and security

Mobile and consumer applications

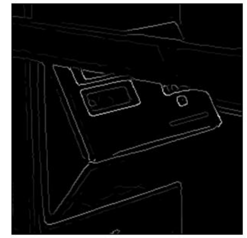


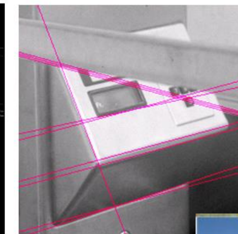
Medical and scientific images

Slide adapted from Svetlana Lazebnik

Computer Vision I Covered the Basics...

- Image Processing Basics

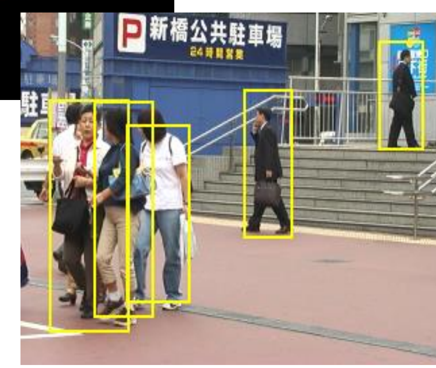




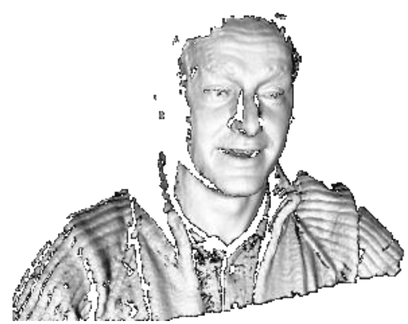
- Segmentation

- Local Features & Matching

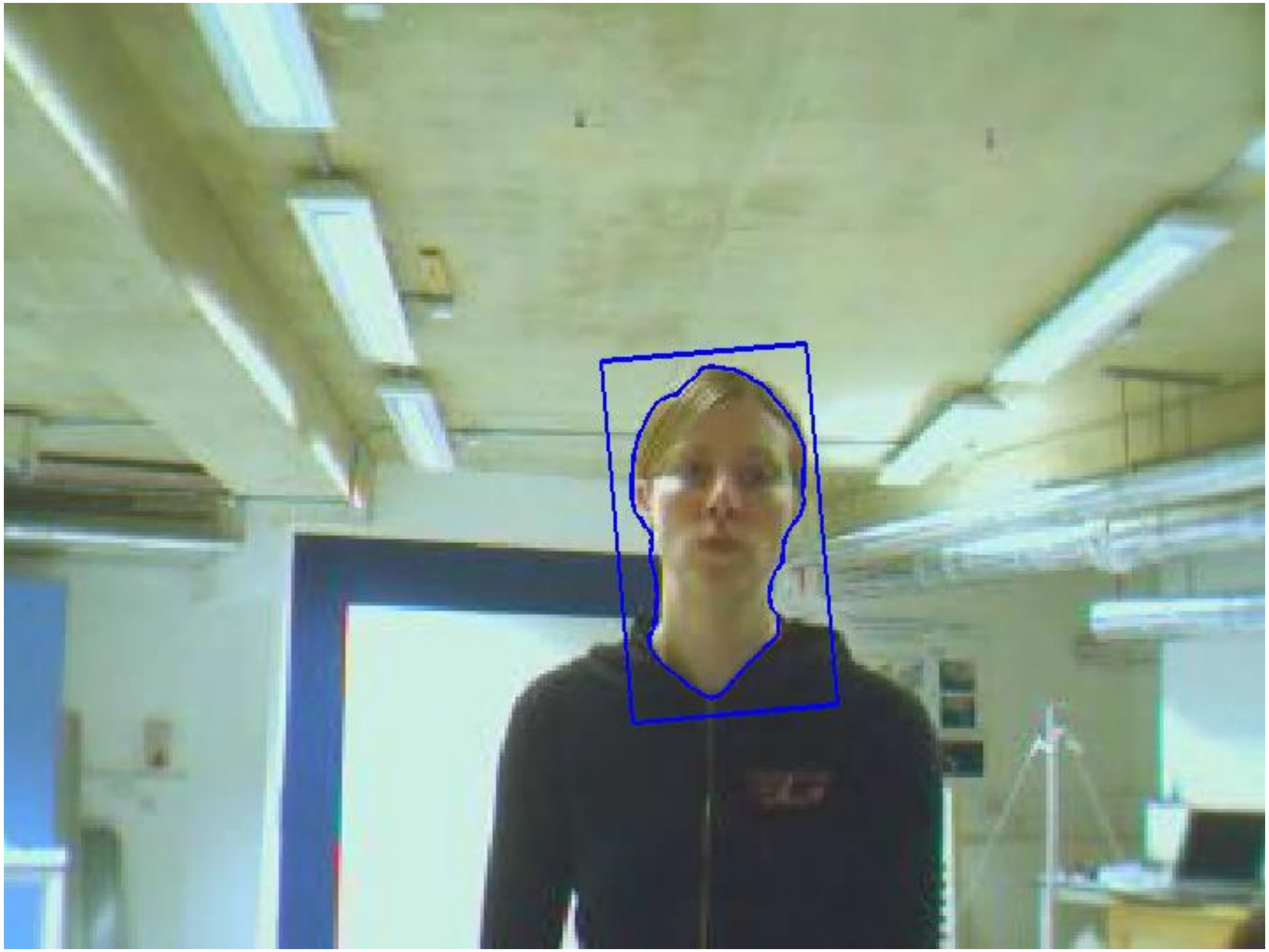
- Object Recognition and Categorization



- 3D Reconstruction



Computer Vision II is all about Motion

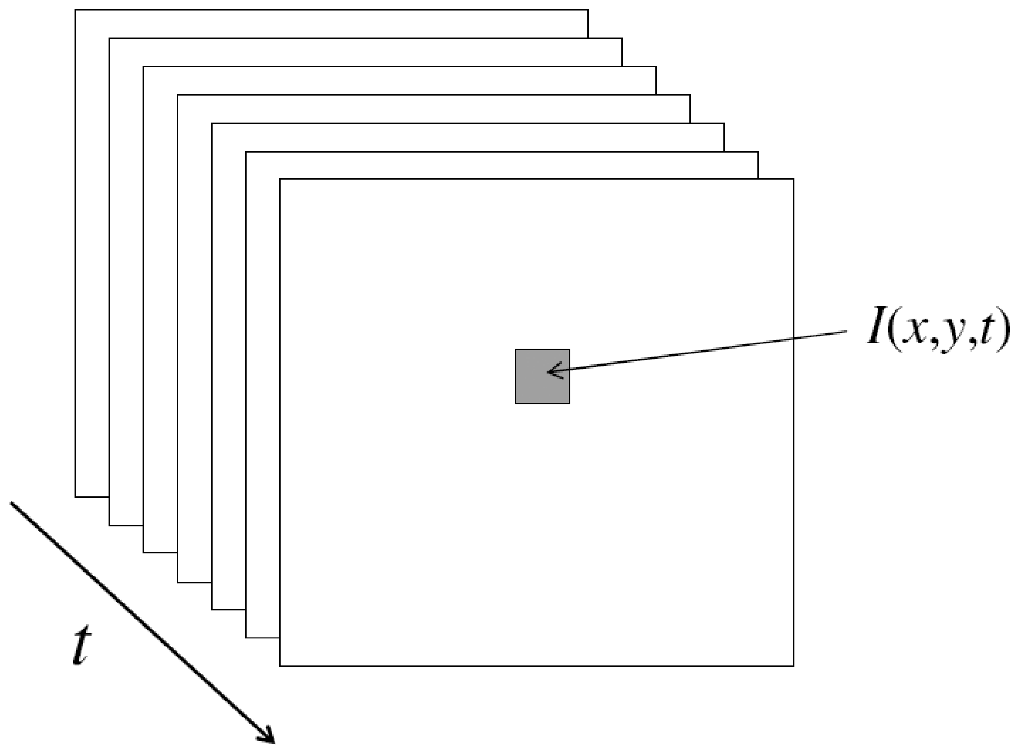


How can we track an object's motion over time?

[Bibby & Reid, ECCV'08]

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 Is Tracking?

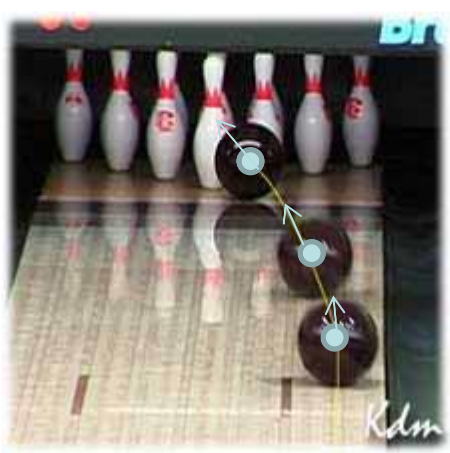
- **Goal**
  - *Estimate the number and state of objects in a region of interest*
- **State**
  - We are using the term state to describe a vector of quantities that characterize the object being tracked.

E.g.	$[x,\ y]$	(location)
	$[x,\ y,\ dx,\ dy]$	(location + velocity)
	$[x,\ y,\ appearance\text{-}params]$	(location + appearance)

  - Because our observations will be noisy, estimating the state vector will be a statistical estimation problem.

What Is Tracking?

- Goal
  - *Estimate the number and state of objects in a region of interest*
- Objects
  - We will look at a large variety of objects to track.
  - They can be given by a user or detected automatically.
  - Most interesting are people.



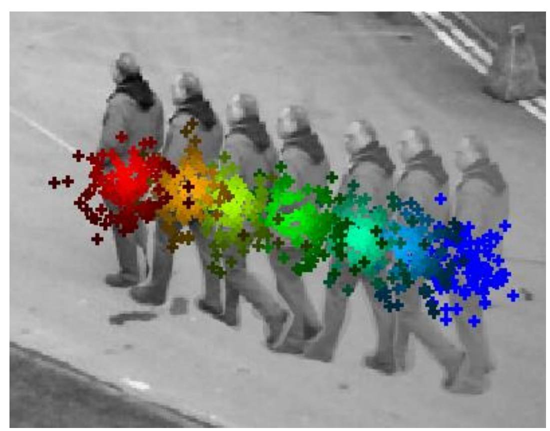


Image sources: Kristen Grauman, Michael Breitenstein, Vittorio Ferrari

What Is Tracking?

- Goal
  - *Estimate the number and state of objects in a region of interest*
- 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 $1, 2, ..., t\text{-}1$. (*Can’t look into the future!*)
  - Usually we want to be as efficient as possible, even “real-time”.

⇒ These concerns lead naturally to recursive estimators.

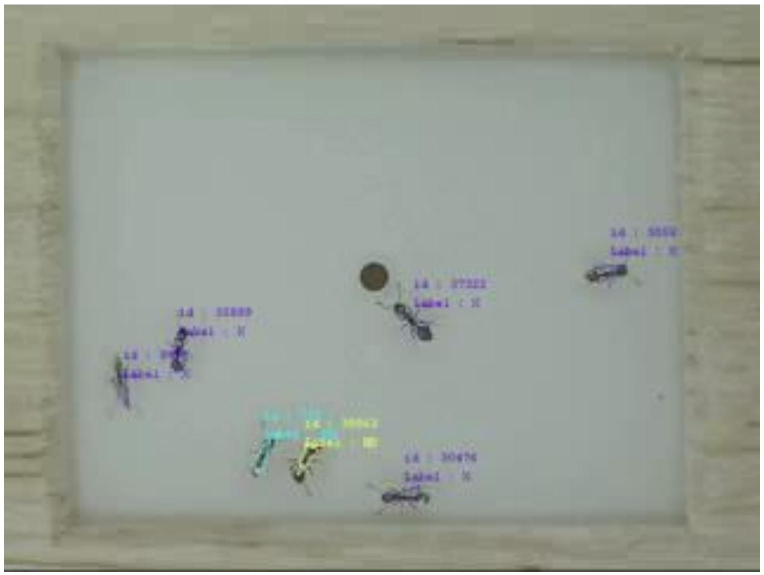
Types of Tracking

- **Single-object tracking** focuses on tracking a single target in isolation.

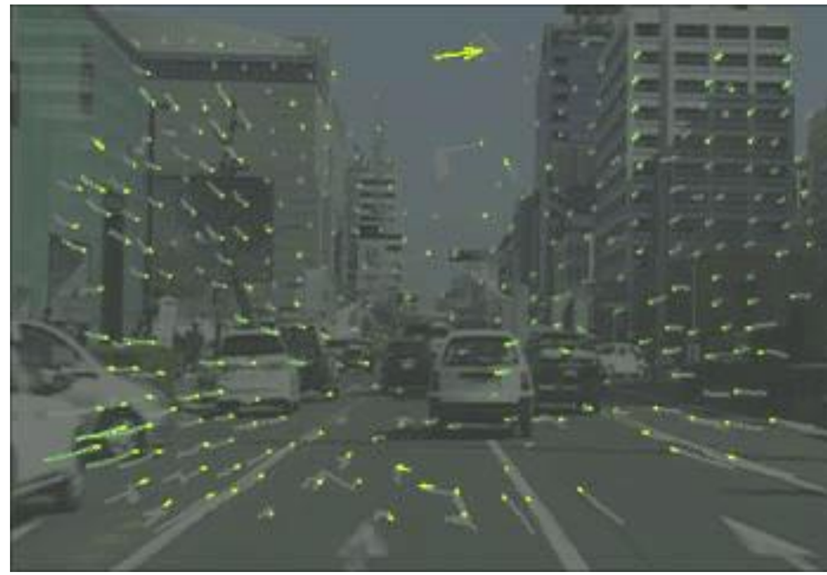
[Z. Kalal, K. Mikolajczyk, J. Matas, PAMI'10]

Types of Tracking

- **Multi-object tracking** tries to follow the motion of multiple objects simultaneously.



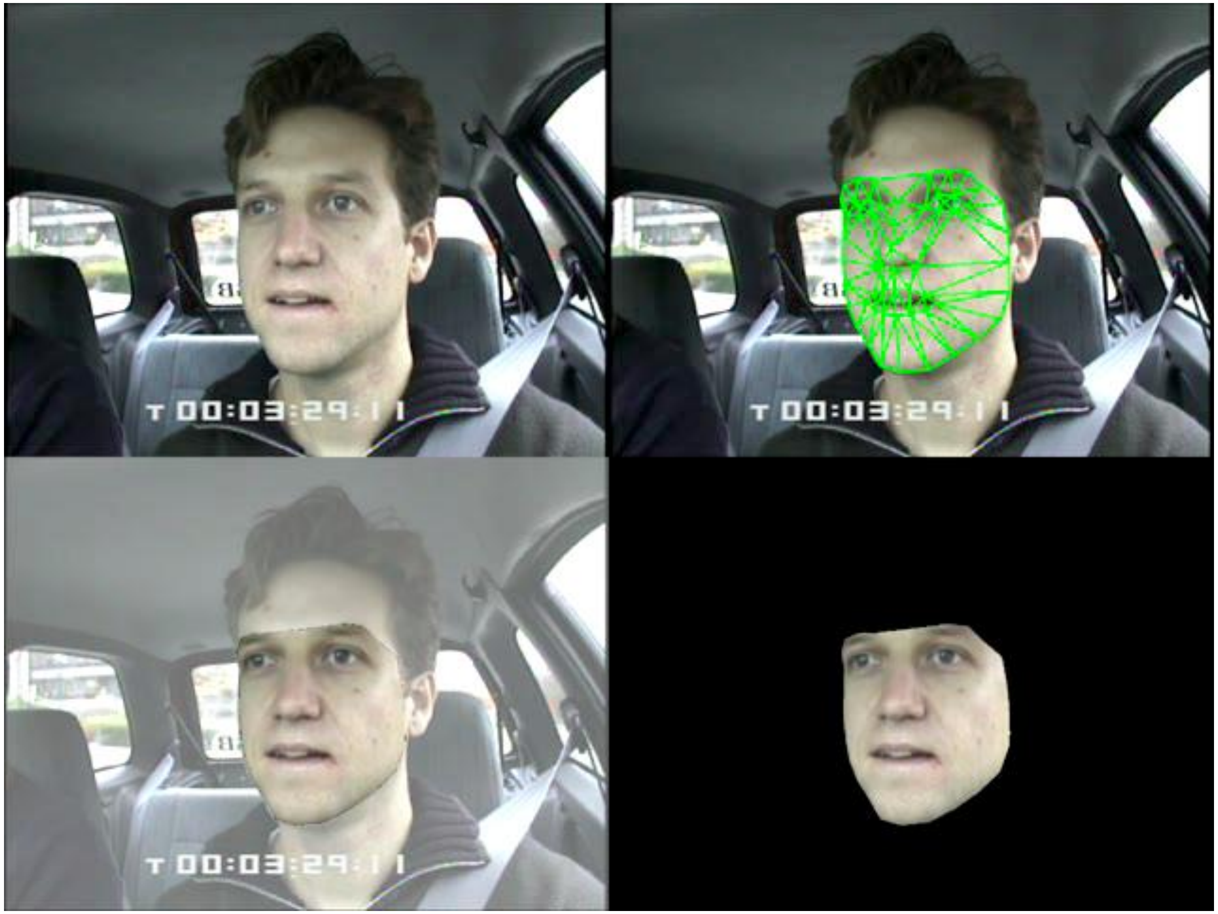
Ant behavior, courtesy of Georgia Tech biotracking



"Objects" can be corners, and tracking gives us optical flow.

Types of Tracking

- **Articulated tracking** tries to estimate the motion of objects with multiple, coordinated parts



Types of Tracking

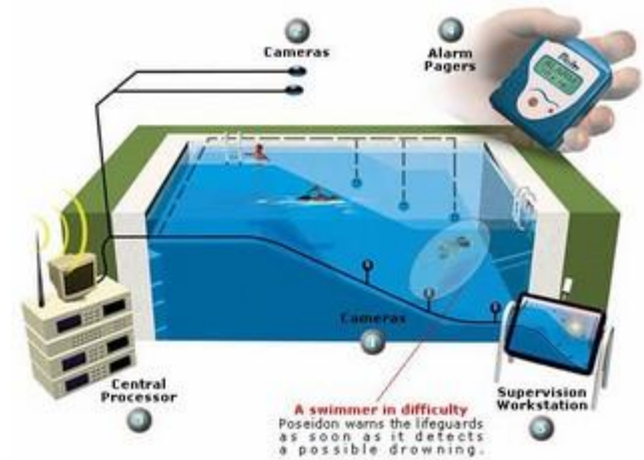
- **Active tracking** involves moving the sensor in response to motion of the target. Needs to be real-time!

Slide credit: Robert Collins

Applications: Safety & Security

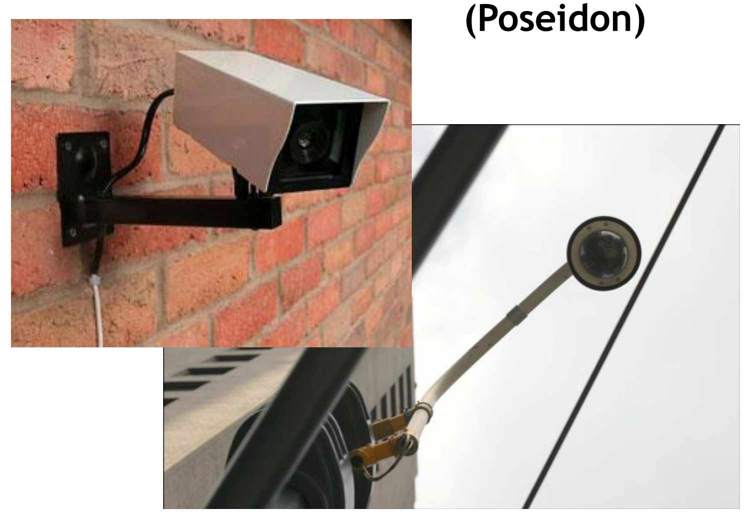
Autonomous robots

Driver assistance



Monitoring pools
(Poseidon)

Pedestrian detection
[MERL, Viola et al.]



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 Carribean*, Industrial Light and Magic
(Source: S. Seitz)

Slide adapted from Svetlana Lazebnik, Kristen Grauman

Why Are There So Many Papers on Tracking?

- **Because what kind of tracking "works" depends on problem-specific factors...**

image source: Microsoft

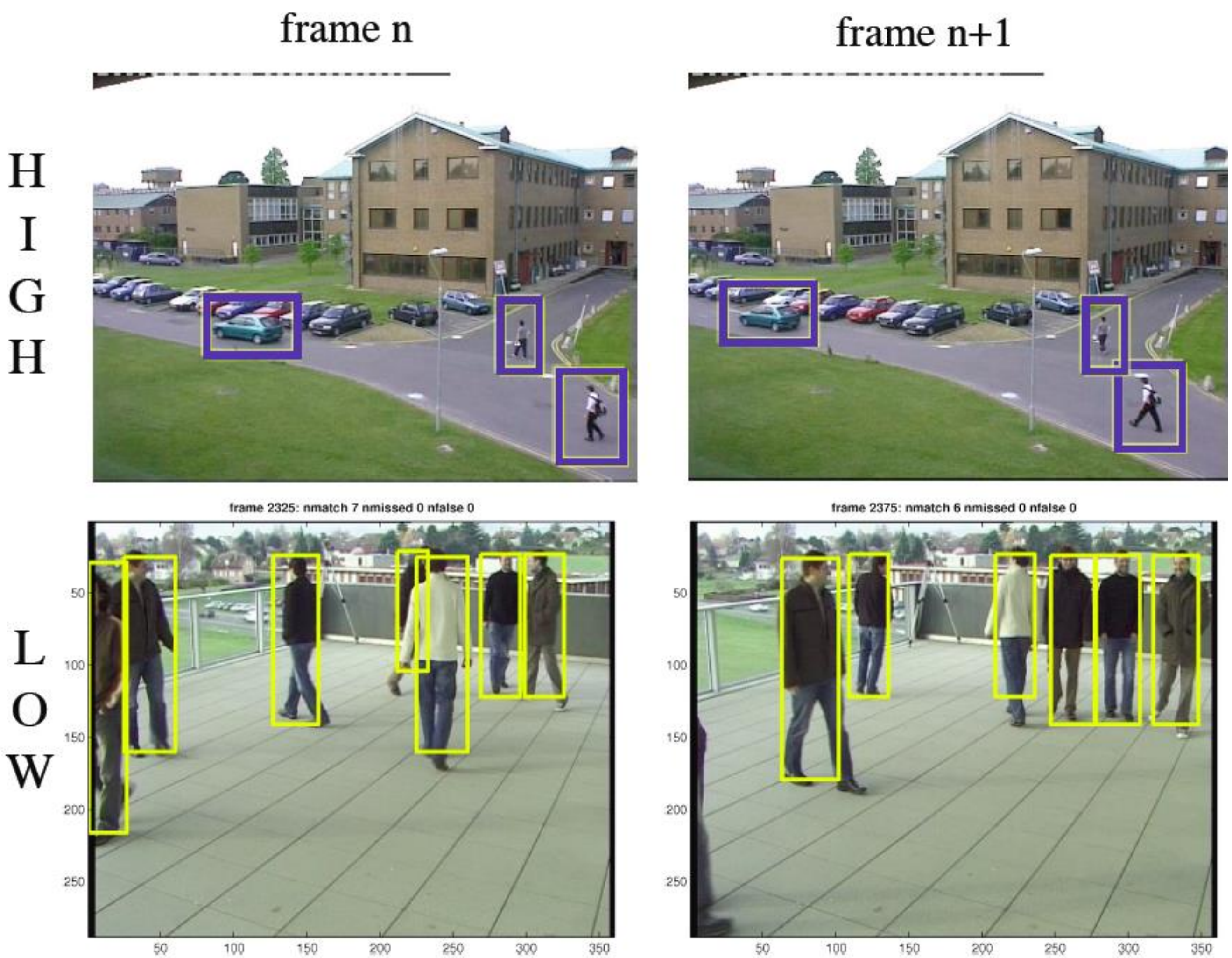
Factors: Discriminability

- 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

t=1 t=2 ... t=20 t=21

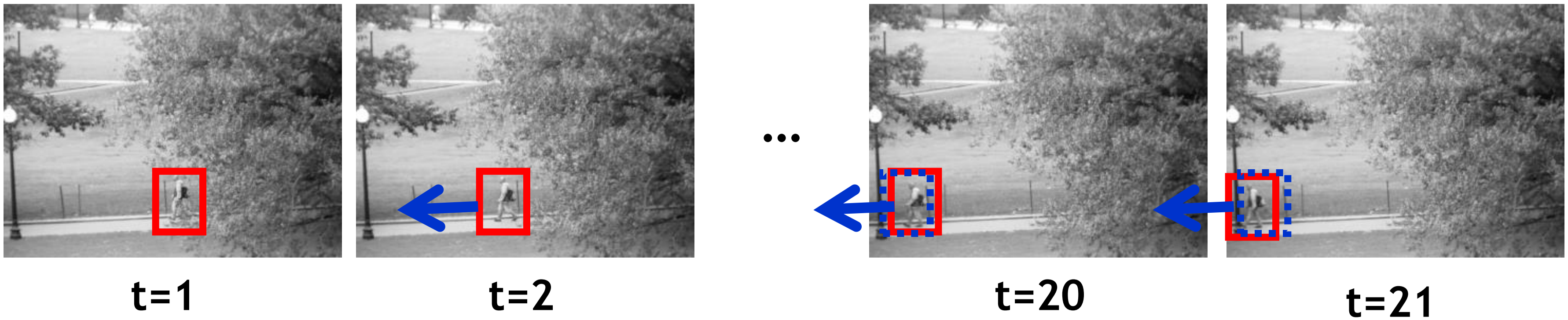
- **Detection**
  - Find the object(s) of interest in the image.

Elements of Tracking

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

Image credit: Kristen Grauman

Elements of Tracking

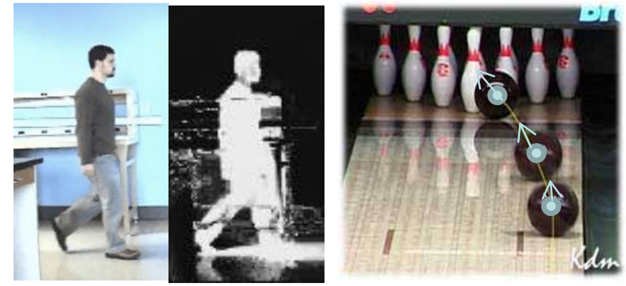


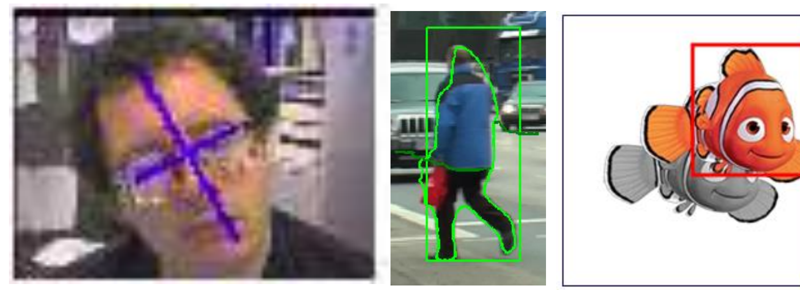
- Detection
  - Find the object(s) of interest in the image.
- Association
  - Determine which observations come from the same object.
- Prediction
  - Predict future motion based on the observed motion pattern.
  - Use this prediction to improve detection and data association in later frames.

Image credit: Kristen Grauman

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

Image sources: Kristen Grauman, Gary Bradsky, Esther Horbert, Helmut Grabner

Background Modeling

- **Learning a statistical model of background appearance**

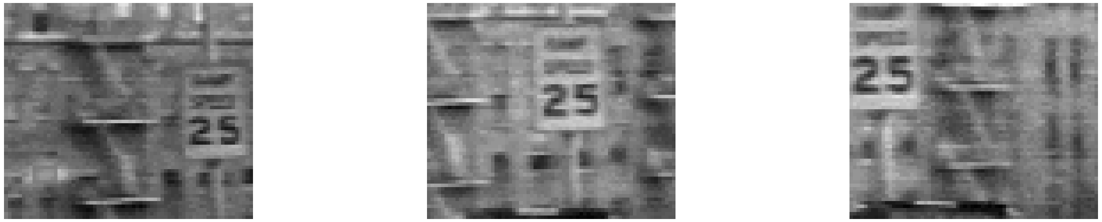
A. Elgammal, D. Harwood, L.S. Davis, Non-parametric Model for Background Subtraction, ECCV 2000.

Applications: Visual Surveillance

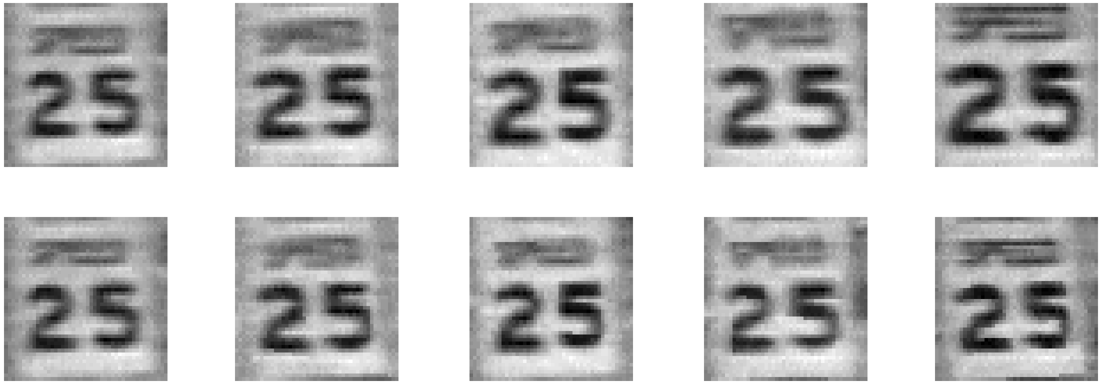
Video by Daniel Roth, ETH Zurich

Template Tracking

- Lucas-Kanade registration applied to tracking ⇒ KLT



Video sequence



Tracked template

J. Shi and C. Tomasi. Good Features to Track. CVPR 1994.

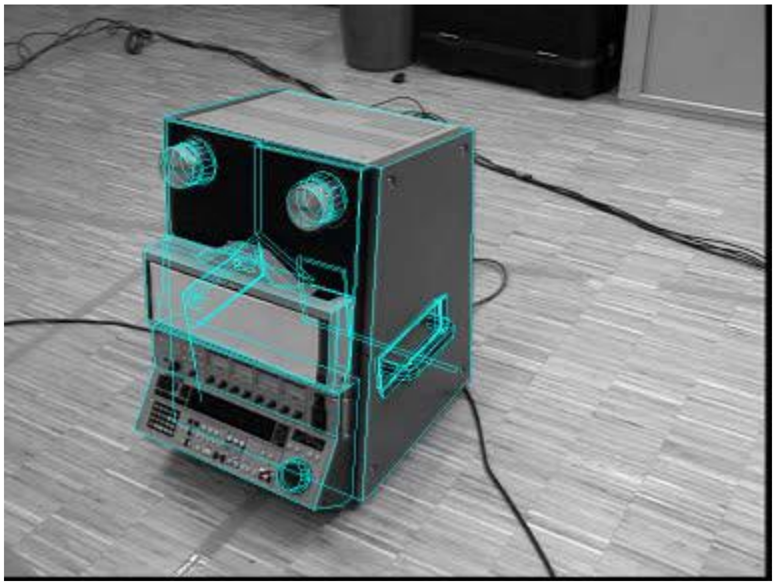
Color-based Tracking

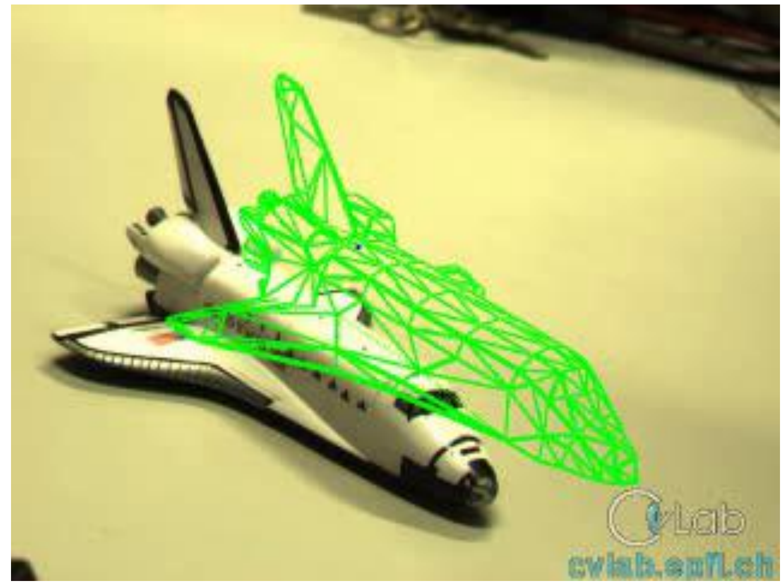
- **Mean-Shift Tracking**

D. Comaniciu, V. Ramesh, P. Meer. Kernel-Based Object Tracking, PAMI, Vol. 25(5), pp. 564-575, 2003.

Model-based Tracking

- Tracking lines on the object given a 3D model





L. Vacchetti, V. Lepetit and P. Fua, Stable Real-Time 3D Tracking Using Online and Offline Information, PAMI, Vol. 26(10), pp. 1385-1391, 2004.

Applications: Tracking Faces for AR

- Flexible models for an entire class of objects

[L. Vacchetti, V. Lepetit, P. Fua, PAMI'04]

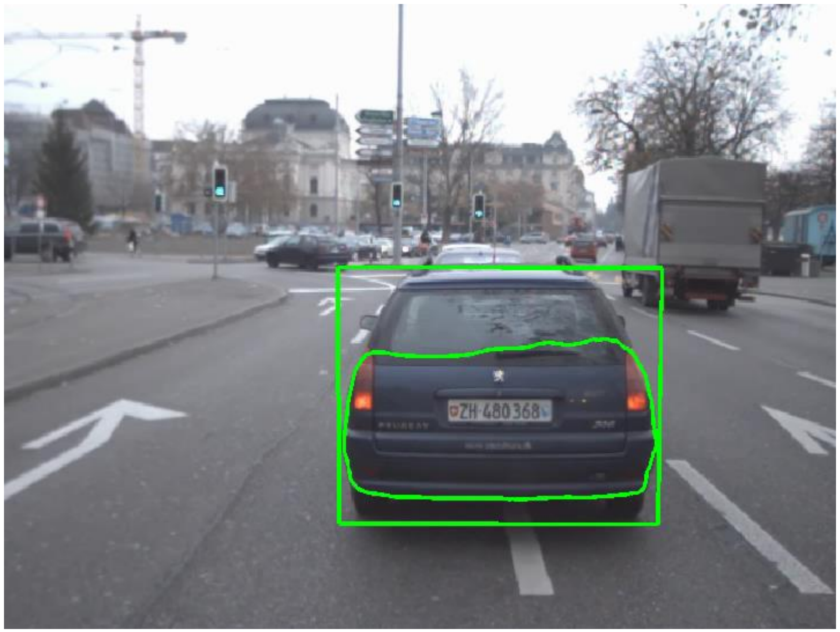
Contour-based Tracking

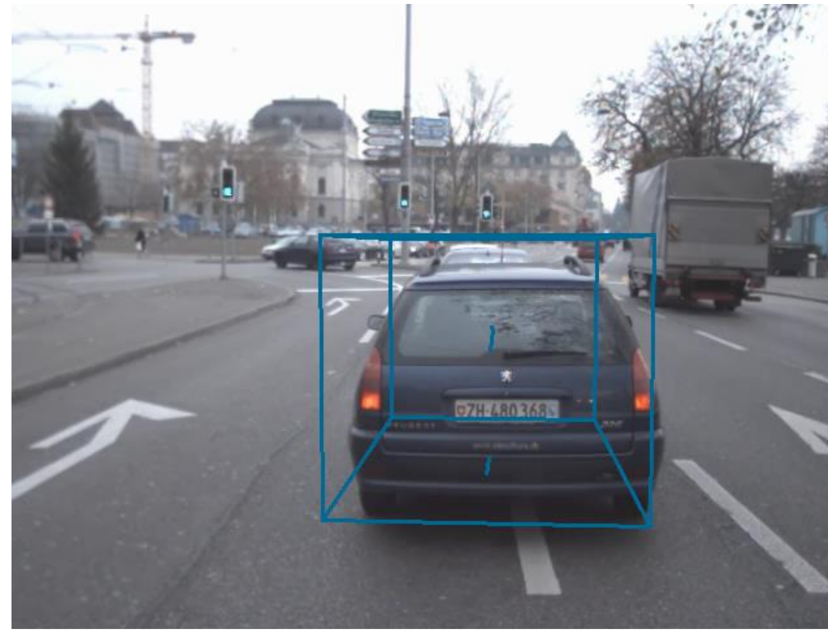
- **Level Set Contour Tracking**

C. Bibby, I. Reid, Robust Real-Time Visual Tracking using Pixel-Wise Posteriors, *ECCV'08*.

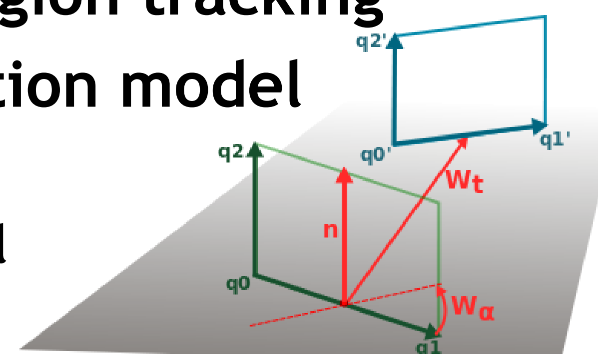
[Bibby & Reid, ECCV'08]

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

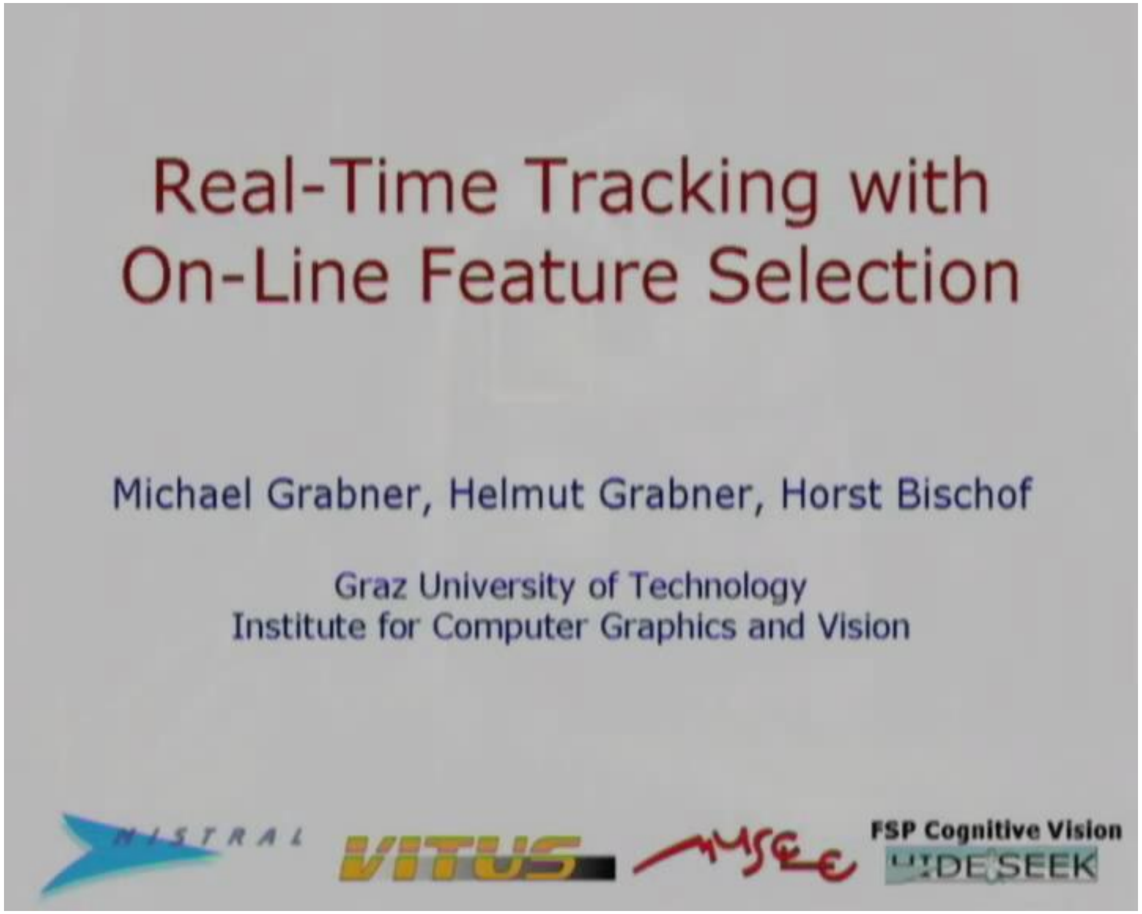


$\Rightarrow$ Input for high-level tracker with car steering model.

[E. Horbert, D. Mitzel, B. Leibe, DAGM'10]

Tracking by Online Classification

- **Learning an object model using Online Boosting**

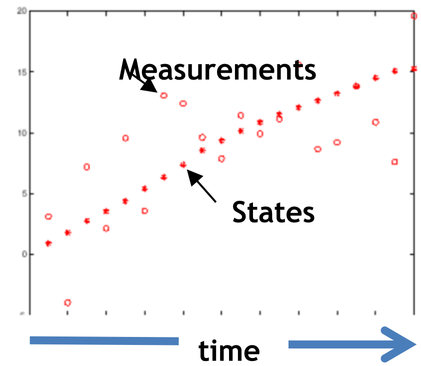


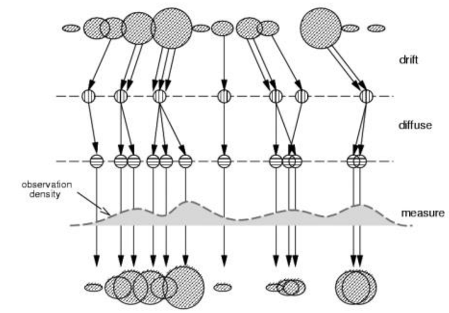
H. Grabner, M. Grabner, H. Bischof. [Real-time Tracking via On-line Boosting], BMVC'06.

[H. Grabner, M. Grabner, H. Bischof, BMVC'06]

Outline of This Lecture

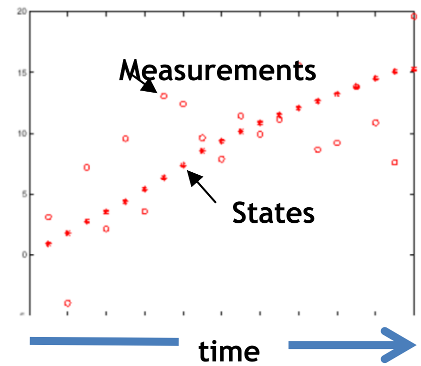
- Single-Object Tracking
- Bayesian Filtering
  - Kalman Filters, EKF
  - Particle Filters
- Multi-Object Tracking
- Articulated Tracking

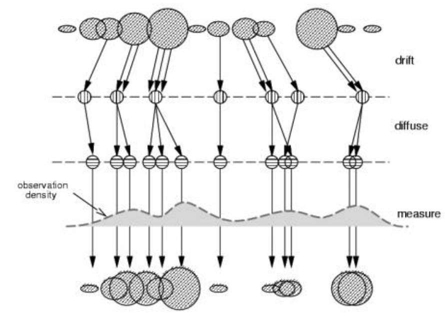


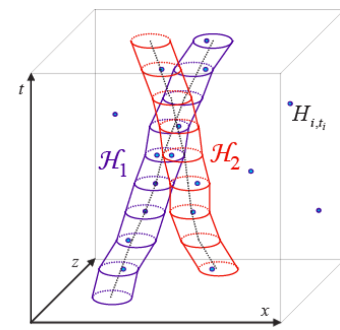


Outline of This Lecture

- 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







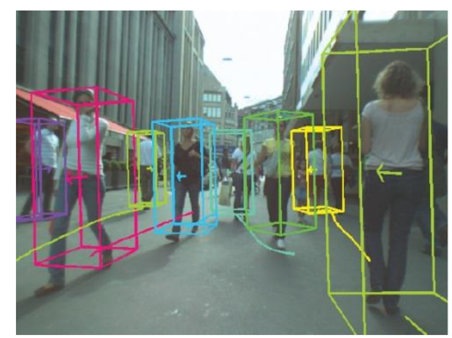


Image sources: Andreas Ess

Multi-Person Tracking

M. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, L. Van Gool. Online Multi-Person Tracking-by-Detection from a Single, Uncalibrated Camera, PAMI, Vol.33(9), pp. 1820-1833, 2011.

[Breitenstein, Reichlin, Leibe, Koller-Meier, Van Gool, ICCV'09; PAMI'10]

Applications: Tracking Sports Players

- **Automatic player tracking for sports scene analysis**
  - **Several companies active in this area…**

[Breitenstein, Reichlin, Leibe, Koller-Meier, Van Gool, ICCV'09; PAMI'10]

Applications: Pedestrian Safety in Cars

[Ess, Leibe, Schindler, Van Gool, PAMI'09]

Predicting Behavior of "Dynamic Obstacles"

SEVENTH FRAMEWORK PROGRAMME

EUROPA

[Ess, Leibe, Schindler, Van Gool, ICRA'09]

Applications: Mobile Robot Navigation

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)

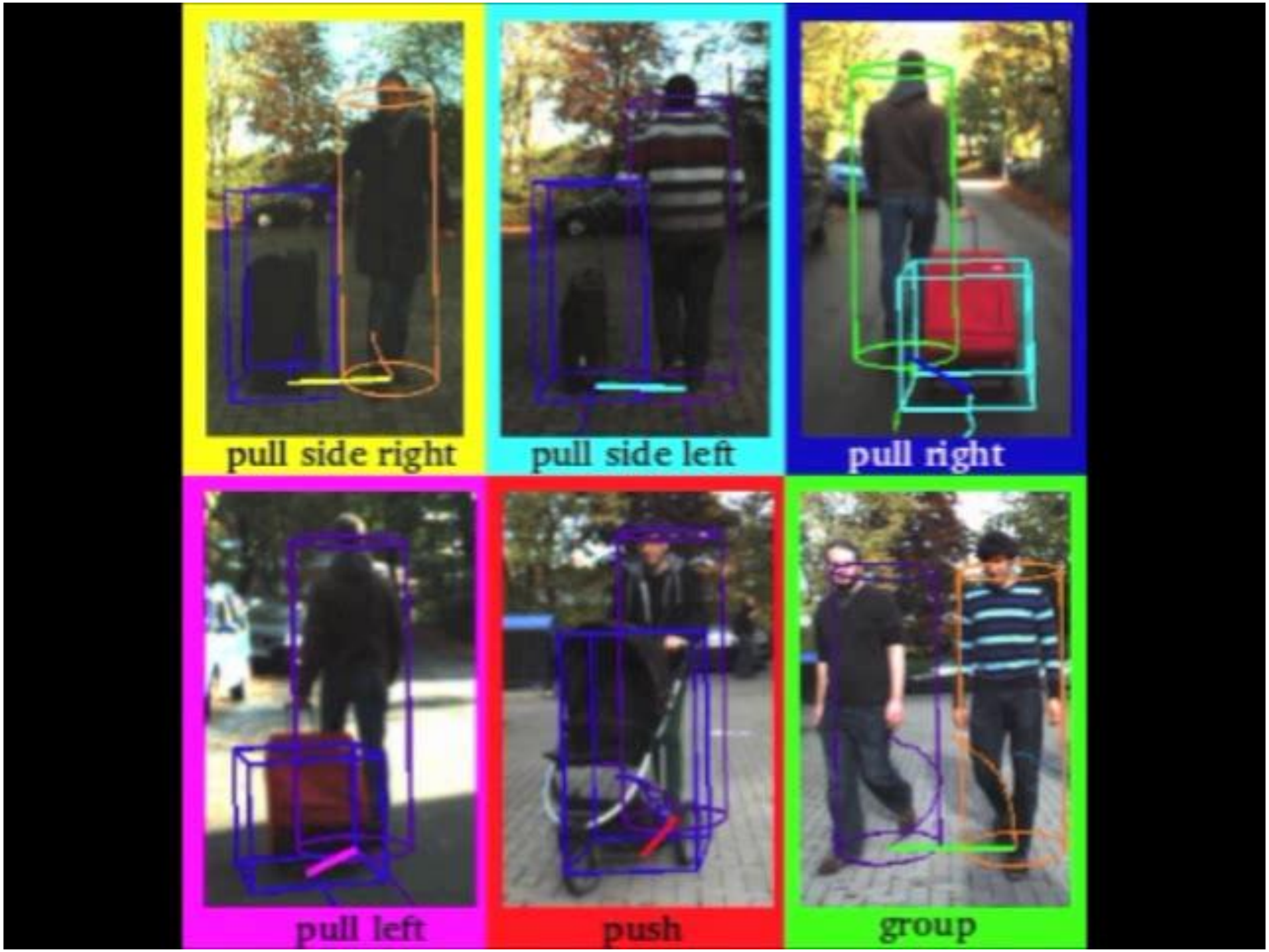
[O. Hosseini Jafari, D. Mitzel, B. Leibe, ICRA'14]

Mobile Tracking in Densely Populated Settings

(Tracking based on stereo depth only, no detector verification)

[D. Mitzel, B. Leibe, ECCV'12]

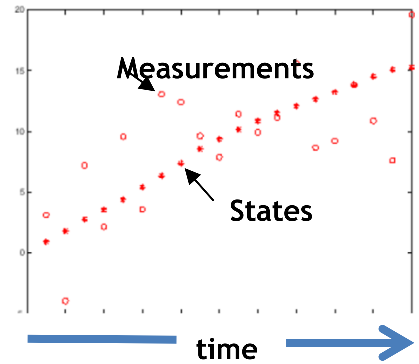
Analyzing Person-Object Interactions

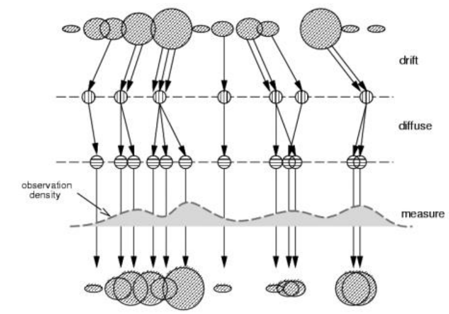


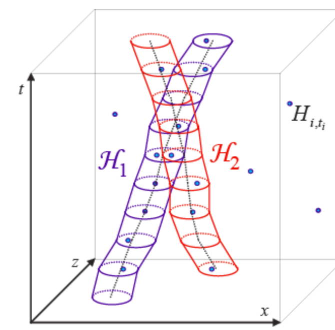
[T. Baumgartner, D. Mitzel, B. Leibe, CVPR'13]

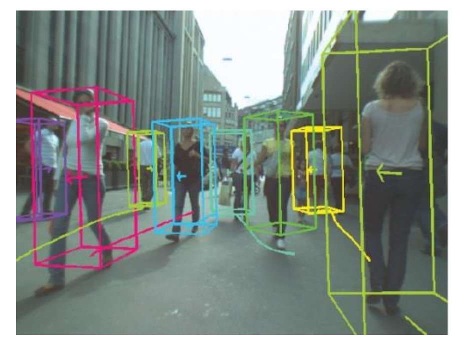
Outline of This Lecture

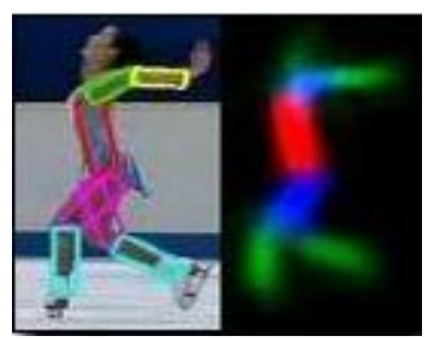
- Single-Object Tracking
- Bayesian Filtering
  - Kalman Filters, EKF
  - Particle Filters
- Multi-Object Tracking
  - Multi-hypothesis data association
  - MHT, JPDAF, MCMCDA
  - Network flow optimization
- Articulated Tracking
  - GP body pose estimation
  - Model-based tracking, AAMs
  - Pictorial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











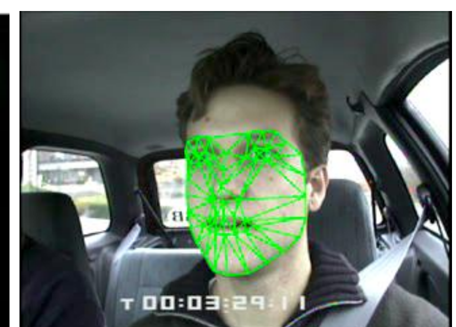


Image sources: Andreas Ess, Deva Ramanan, Ian Matthews

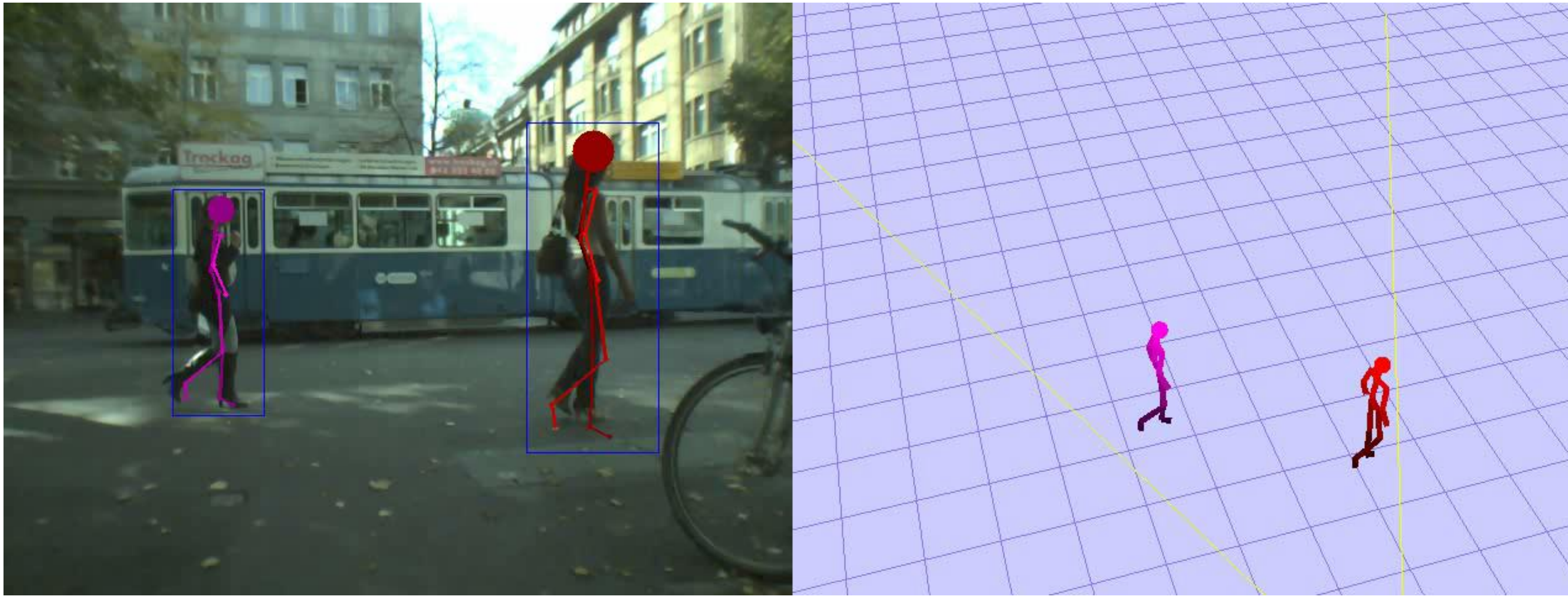
Articulated Person Tracking

- **Tracking and interpreting detailed body motion.**

D. Ramanan, D.A. Forsyth, A. Zisserman. [Tracking People by Learning their Appearance](), PAMI 2007.

[D. Ramanan, D. Forsyth, PAMI'07]

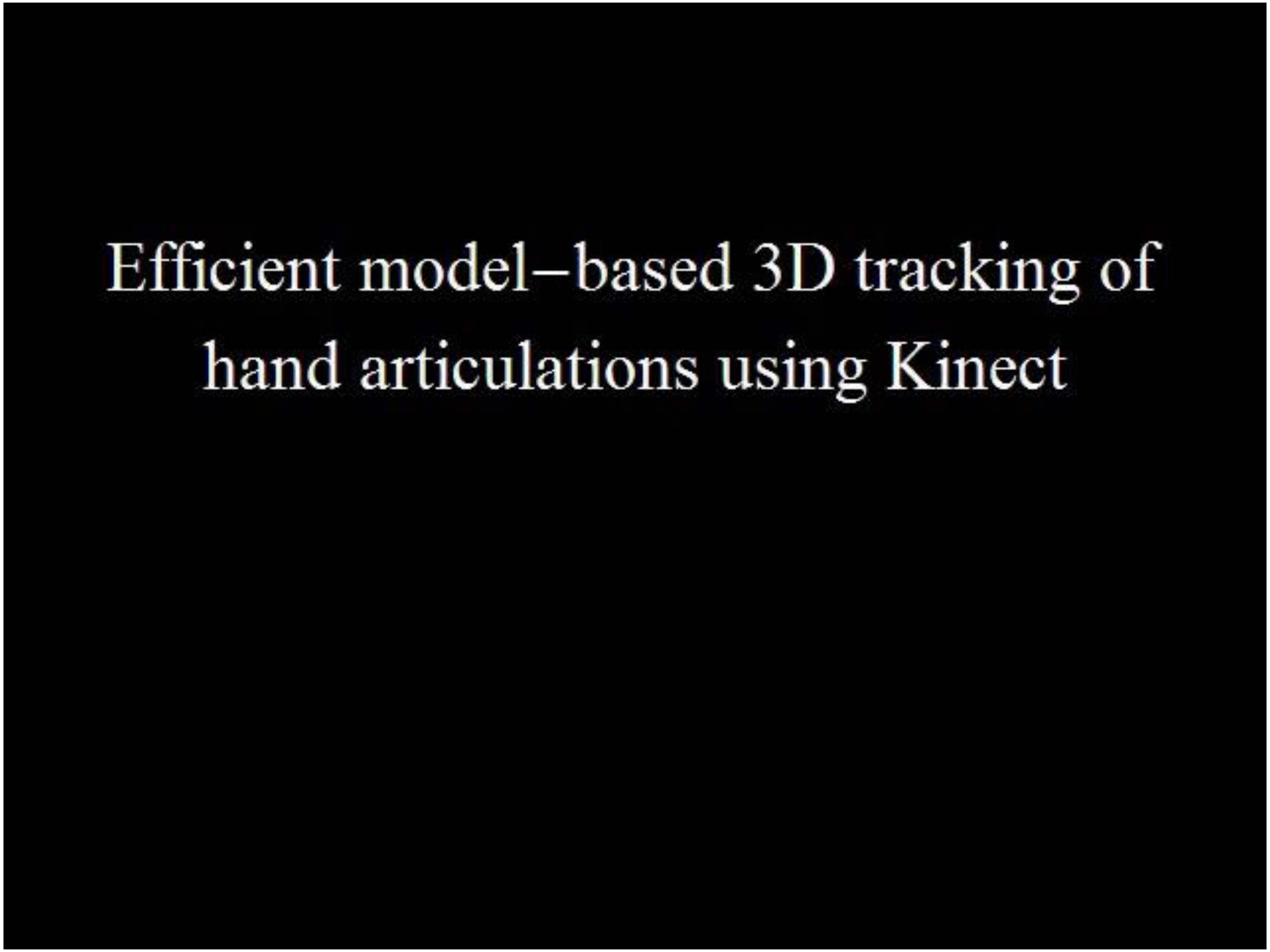
Combination: Articulated Multi-Person Tracking



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

[Gammeter, Ess, Jaeggli, Schindler, Leibe, Van Gool, ECCV'08]

Articulated Hand Tracking



[J. Oikonomidis, A. Argyros, BMVC’11]

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

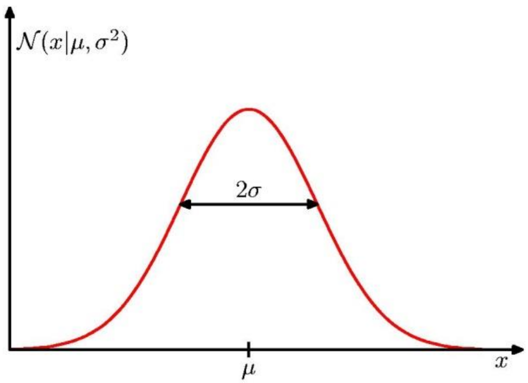
Applications: Facial Animation Transfer

Commercialized by faceshift
https://www.youtube.com/watch?v=b5m40eF3C9E

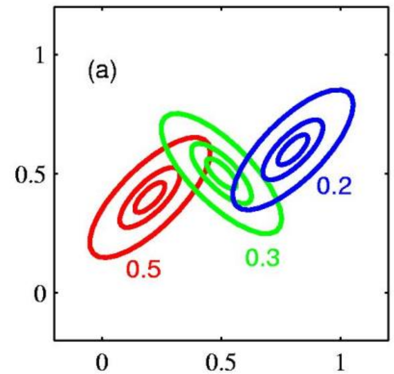
Wrap-Up

- You've seen many exciting applications
  - In this lecture, we'll look at how all of this works.
  - We will build upon basics from the CV I and ML lectures.
  - *You can attend the class without having heard those, but please use the available online material for self-study.*

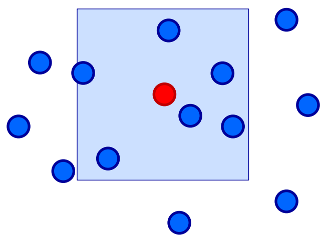
- Next lecture: Background modeling
  - Please repeat the following topics from the ML lecture:



Gaussians & ML estimation



Mixtures of Gaussians & EM



Kernel density estimation

Questions?