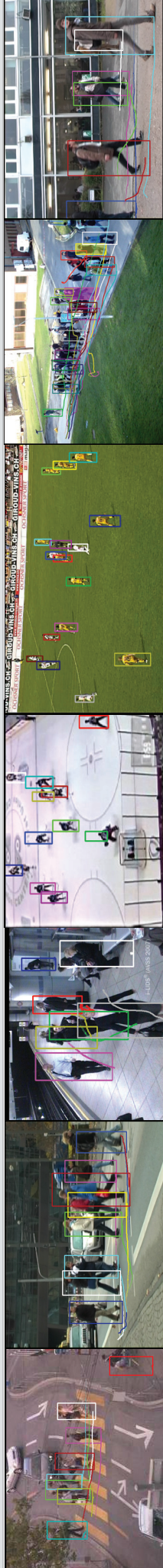# Evaluation of Agent Motion in Video: Online Tracking-by-Detection

Michael D. Breitenstein   Bastian Leibe   Luc Van Gool

ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

HERMES Human Expressive Representations of Motion and their Evaluation in Sequences

CogSys Cognitive Systems

Motivation

- New cameras are installed daily everywhere (webcams, mobile phones, cars, surveillance cameras, ...)
- **Goal:** Analyze the behaviour of moving persons
- **Key task:** Detect, track and distinguish persons
  - Arbitrary video input (no scene knowledge)
  - Single camera, potentially moving
  - Online (no delay or offline processing)
- ⇒ **Challenge:** Single-frame detection not reliable!

Applications

- Surveillance: monitoring of public places, alerting
- Broadcasting: 3DTV, sports analysis and statistics
- Traffic safety: emergency brake assist, attention assist

Approach: Detector Confidence Particle Filter

- Robustly integrate complementary information:
  motion + object-class + instance-specific knowledge
- Exploit state-of-the-art object detectors
  - Train person-specific classifiers online for each target to handle false positive detections
  - Exploit continuity of detector confidence to handle missing detections

Input: unreliable detections
and detector confidence
Output: trajectories
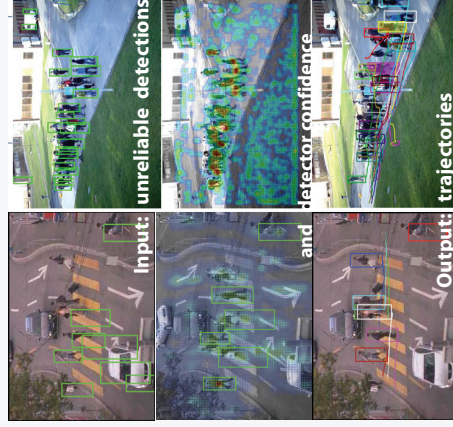
Sequential Monte Carlo Estimation

- Estimate multi-modal distribution of target in the image (maintain multiple hypotheses)
- One particle filter per target
  - Initialization based on detections
  - Constant velocity motion model
  - Novel observation model
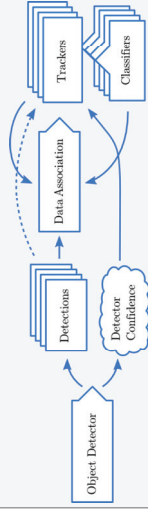
Observation Model with 3 Terms

$$w_{tr,p} = p(y_t | x_t^{(i)}) =$$
$$\underbrace{\beta \cdot \mathcal{I}(tr) \cdot p_N(p - d^*)}_{detection} + \underbrace{\gamma \cdot d_c(p) \cdot p_o(tr)}_{det.\ confidence\ density} + \underbrace{\eta \cdot c_{tr}(p)}_{classifier}$$

1. Detection Term: ⇒ select which detection d^* to use

- Greedy data association:
  - Associate max. 1 detection d with max. 1 tracker tr
- Matching Score:

$$s(tr, d) = g(tr, d) \cdot (c_{tr}(d) + \alpha \cdot \sum_{p \in tr}^{N} p_N(d - p))$$

- Evaluate appearance: person-specific classifier $c_{tr}(d)$
- Evaluate distance between d and tr (particles p)
- Evaluate motion direction by gating function $g(tr, d)$ (fast moving objects don't change course abruptly)

Gating function

2. Detector Confidence Term: ⇒ guide particles if no reliable detection

- Detector confidence $d_c(p)$ for particle p
  - Hough voting space density (for ISM detector *[Leibe 08]*)
  - SVM score before NMS (for HOG detector *[Dalal 05]*)

Sliding window   **Hough voting**
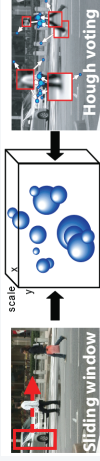
- Assess reliability:
  - Only use $d_c(p)$ if probably caused by foreground:
  ⇒ if another tracker tr' is nearby and if tr' is associated to a detection d'

Reliability function
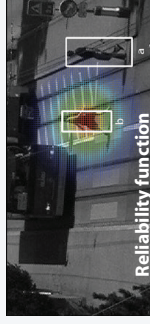
3. Classifier Confidence Term: ⇒ differentiate between targets

- Learn target-specific classifiers:
  - Evaluate patch at every particle position
- Boosted classifier *[Grabner, 06]*
  - Trained for each d^* vs. all other d
  - Use color+texture features
- Classifiers adaptive, get more and more discriminative

Key Factors

- Robustness during partial occlusion & false pos. detections:
  ⇒ Detector and classifier confidence terms
- Recovery after occlusion:
  ⇒ Multi-modality of particle filter, target-specific classifiers
- Appearance change:
  ⇒ Online trained, target-specific classifiers

Results

- Tested for a variety of application scenarios:
  Outperforms state-of-the-art methods that use multi-camera input/scene knowledge/offline processing
  ([Okuma 04] [Berclaz 06] [Leibe 07] [Wu 07] [Huang 08])
- Suitable for online applications
- Every observation model term adds robustness
  CLEAR MOT evaluation for ETH Central:

Observation Model Terms	MOTP	MOTA	FN	FP	ID Sw.
1: Det+DetConf+Class	70.0%	72.9%	26.8%	0.3%	0
2: Det+DetConf	64.0%	54.5%	28.2%	17.2%	5
3: Det+Class	65.0%	55.3%	31.3%	13.4%	0
4: Det	67.0%	40.9%	30.7%	28.0%	10

PETS'09-S2.L1

PETS'09-S2.L2

ETH Central

TUD Crossing

AVSS'07 i-Lids

UBC Hockey

ETH Soccer

Soccer dataset courtesy of LiberoVision and Teleclub

Videos: www.vision.ee.ethz.ch/~bremicha/tracking