ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

BIWI Computer Vision

Variations of a Hough-Voting Action Recognition System

Daniel Waltisberg, Angela Yao, Juergen Gall, Luc Van Gool

Motivation

- Monitoring children and elderly people

- Intelligent surveillance systems

Overview

1. Hough-Voting Framework for Action Recognition
2. Aerial View Challenge
3. Interaction Challenge
4. Summary

Hough-Voting Framework for Action Recognition

- Hough Transform used to detect lines
  → Generalized Hough Transform to detect arbitrary shapes
  → Extension for action recognition

- Straight-forward extension is too high-dimensional to handle: $3t + 3$

- Split into two stages
  Localize: track the person
  Classify: label the action

Localization

detection hypotheses

particle filtering

normalize tracks

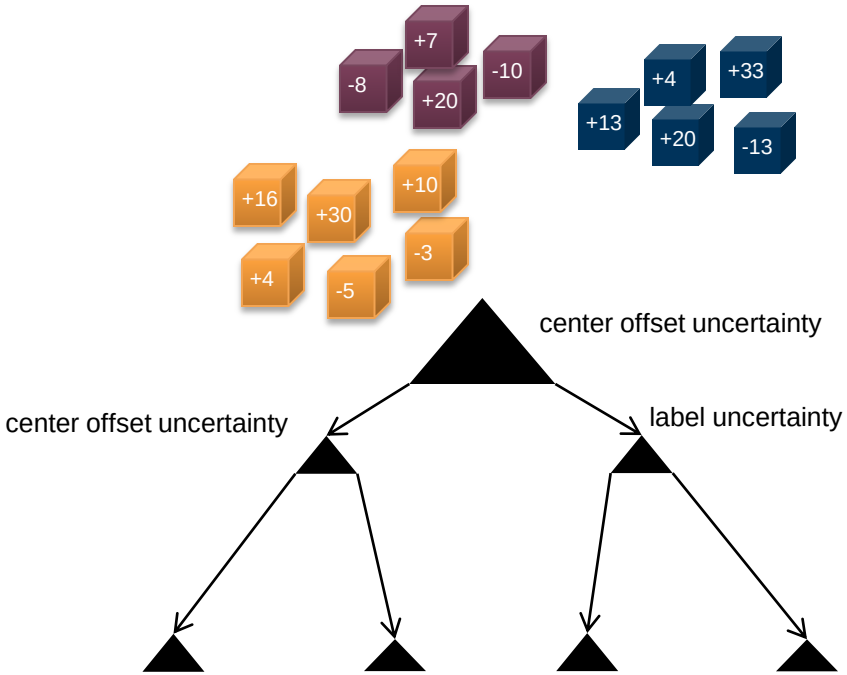
Classification - Training

- Extract randomly sampled 3D patches (e.g 25x25x5)
- Build random tree from the top
  - Create pool of binary tests
  - Select binary test which splits with minimal class uncertainty

$$U_1(A) = -|A| \cdot \sum_c p_c \ln(p_c)$$

or center offset uncertainty

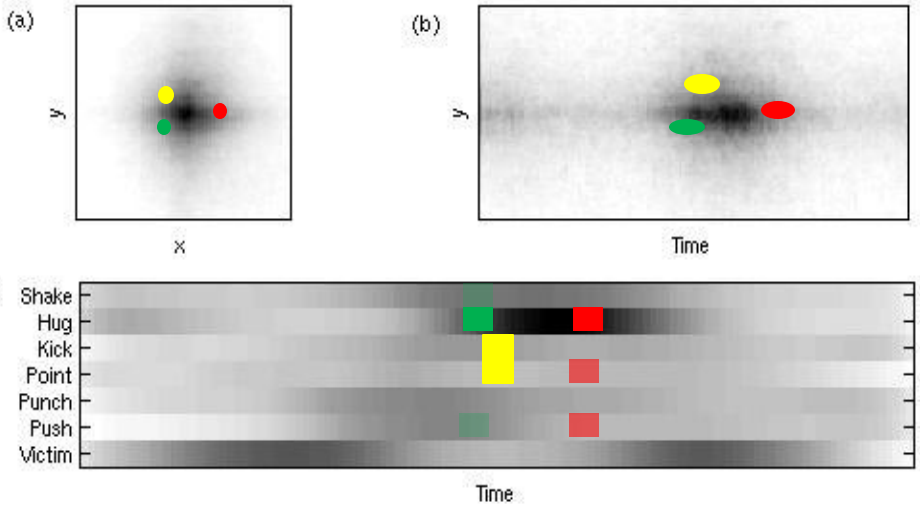
$$U_2(A) = \sum_i ||\boldsymbol{d}_i - \overline{\boldsymbol{d}_A}||^2$$

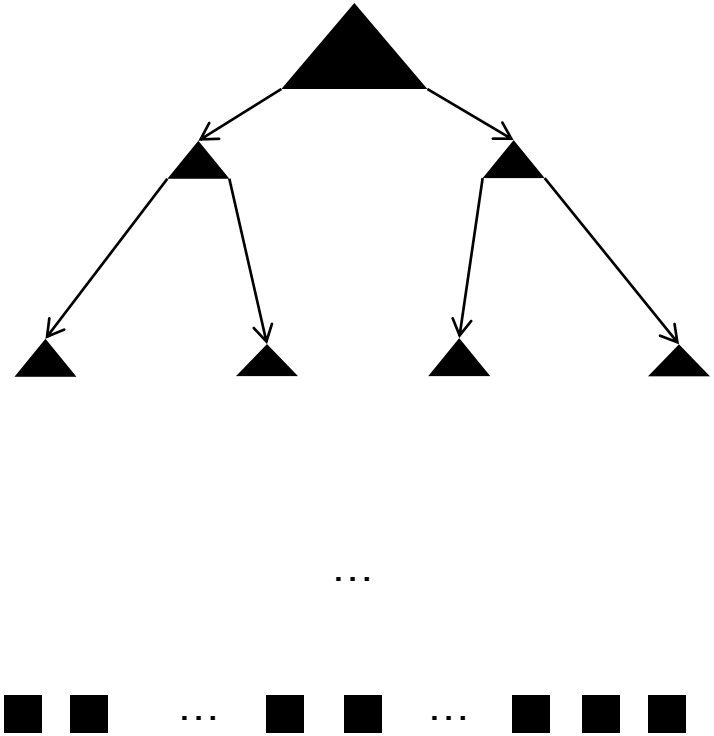


shared features across classes

Classification - Testing

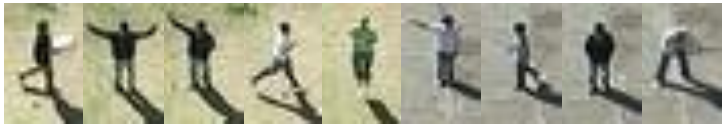
- Extract densely sampled 3D patches
- Pass patches through tree
- Leaves are used to cast votes in the 4D space (action label, center in time and space)





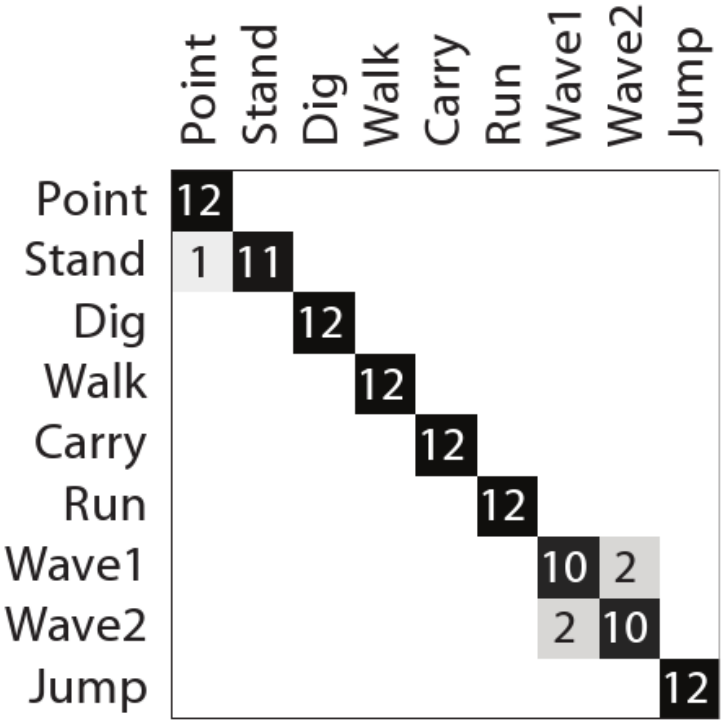
Aerial View Challenge

- Foreground masks were used to extract the tracks of the people
- Hough Voting Framework was applied directly
- The following six feature channels were used:
  - greyscale intensity
  - absolute value of x, y and time derivatives
  - absolute value of optical flow in x and y



Aerial View Challenge

- Performance of 95.3 %



Interaction Challenge

- Apply Hough-voting to both people

Symmetric

Shake Hands

Hug

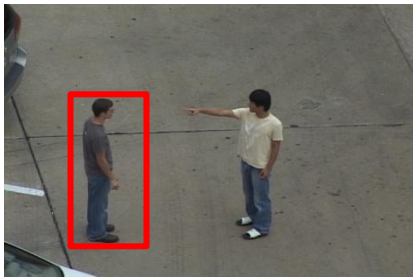
Asymmetric

Push

Kick

Victim

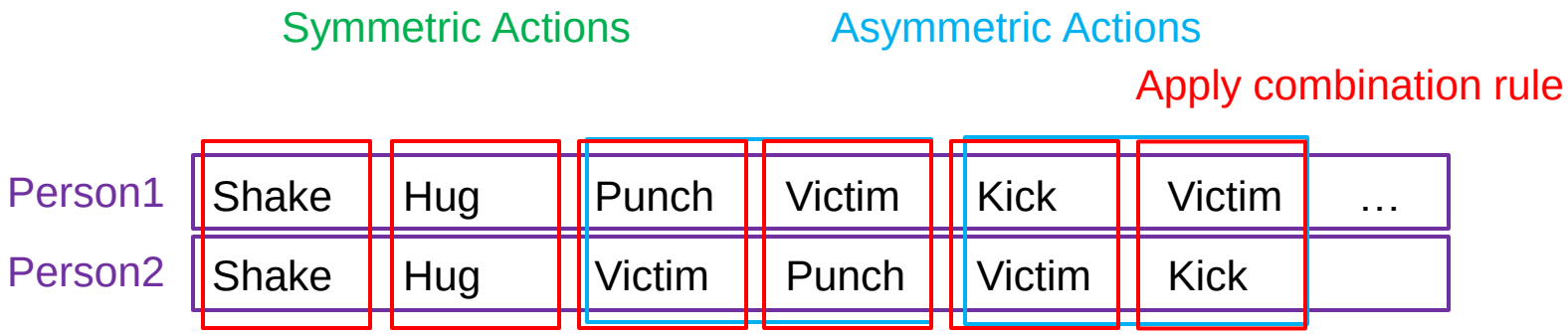
Punch



Point

Interaction Challenge

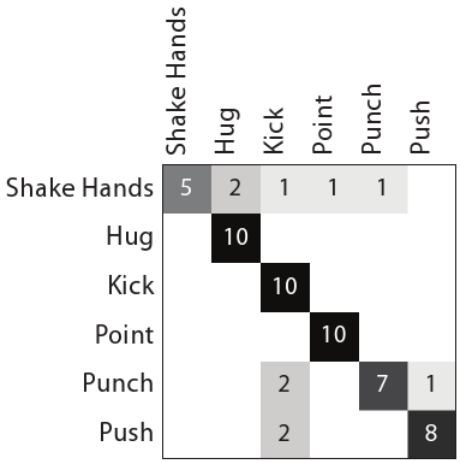
- Apply Hough-voting to both people → probabilistic measure for seven classes and each person
- Use combination rules, e.g. the min rule
- Representation with Decision Profiles



Interaction Challenge

- Similar results for min, product and sum rule

	Min	Max	Product	Sum
Set 1	0.83	0.55	0.87	0.88
Set 2	0.8	0.42	0.77	0.77



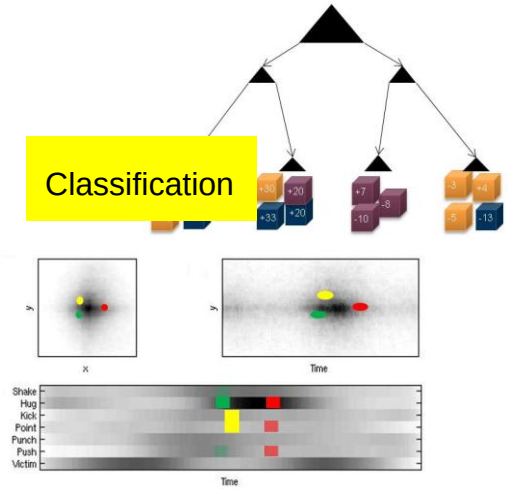
Set1 using min rule

Interaction Challenge

- Applying the Hough-voting to both people directly gives a performance of only 48.3 %
- The presented approach achieves a performance improvement more than 30 %
- Explanation: Much of the movement of the people is caught when applying the Hough-voting to both people.

Summary

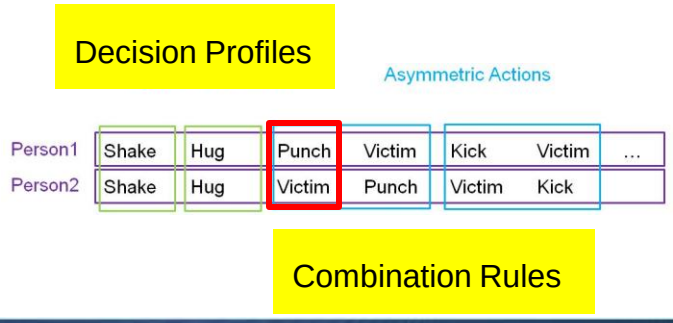
Hough-Voting framework for action recognition



Aerial View Challenge

95.3%

Interaction Challenge



Questions?