Forschungsinitiative
Ko-FAS

Does He Run into Danger?

Läuft er in Gefahr?

Sebastian Köhler, Konrad Doll, Ulrich Brunsmann
University of Applied Sciences Aschaffenburg

Supported by:
Federal Ministry of Economics and Technology
on the basis of a decision by the German Bundestag

Outline

1. Motivation
2. Our Approach
   - Overview
   - Edge Based Motion History Image (MHI)
   - MCHOG Descriptor
   - Classification
3. Results
   - Dataset
   - Quality
   - Response Time
4. Conclusion/Outlook

http://derstandard.at/3169382?seite=3 [7/29/2013]

1. Motivation

FORSCHUNGSINITIATIVE KO-FAS

- Problem scope
  - A pedestrian‘s intention to cross the street can immediately lead to a dangerous situation

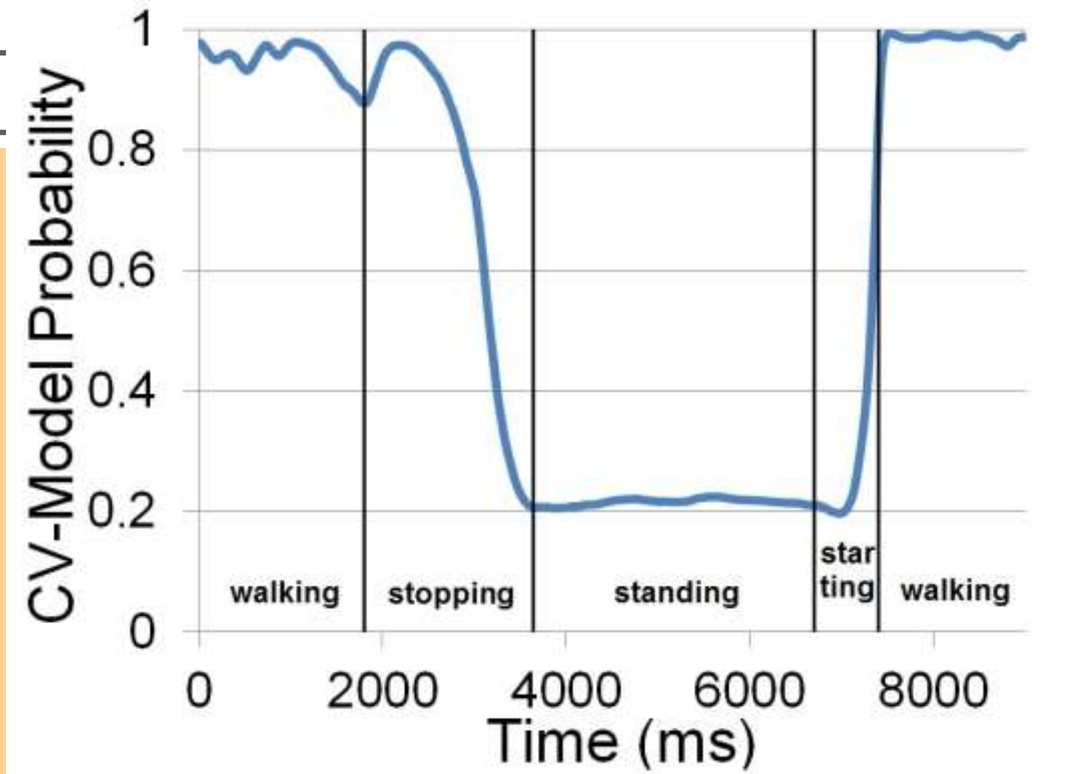
- State of the art (research)
  - Trajectory-based (Kalman filter)
  - Orientation-based change of movement
  - Probabilistic path prediction of learned motion features gathered from dense optical flow[1]
  - Our appraoch: MCHOG[2]

[1] C. Keller, C. Hermes, and D. Gavrila, “Will the pedestrian cross? probabilistic path prediction based on learned motion features,” in Pattern Recognition. Springer, 2011, vol. 6835, pp. 386–395.

[2] S. Köhler, M. Goldhammer, S. Bauer, S. Zecha, K. Doll, U. Brunsmann, K. Dietmayer: “Stationary Detection of the Pedestrian’s Intention at Intersections,” *accepted for publication in:* IEEE Intelligent Transportation Systems Magazine ITSC 2012 Special Issue, invited Paper.

2. Our Approach

Overview | Edge Based MHI | MCHOG Descriptor

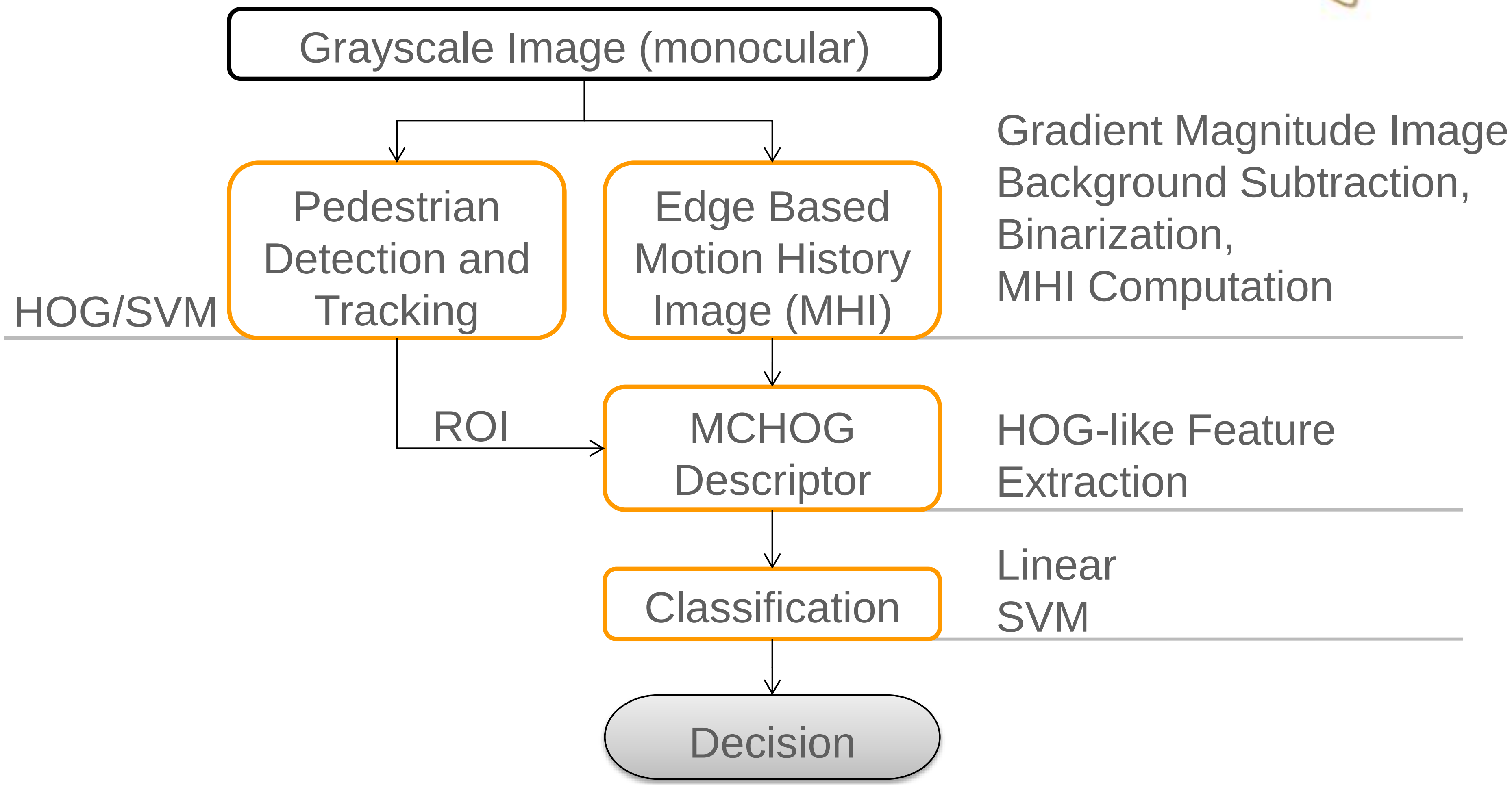


- Goal: Earlier detection of gait initiation
- Our approach: Using high resolution spatio-temporal image information

Communication

2. Our Approach

Overview | Edge Based MHI | MCHOG Descriptor



FORSCHUNGSINITIATIVE KO-FAS

1. Magnitude Image Computation

$$\|I_{mag}\| = \sqrt{{I_{edgeX}}^2 + {I_{edgeY}}^2}$$

$$I_{edgeX} = I_{gray} * K_x \qquad K_x = \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}$$

$$I_{edgeY} = I_{gray} * K_y \qquad K_y = {K_x}^T$$

2. Background Subtraction and Binarization

Mixture of Gaussians

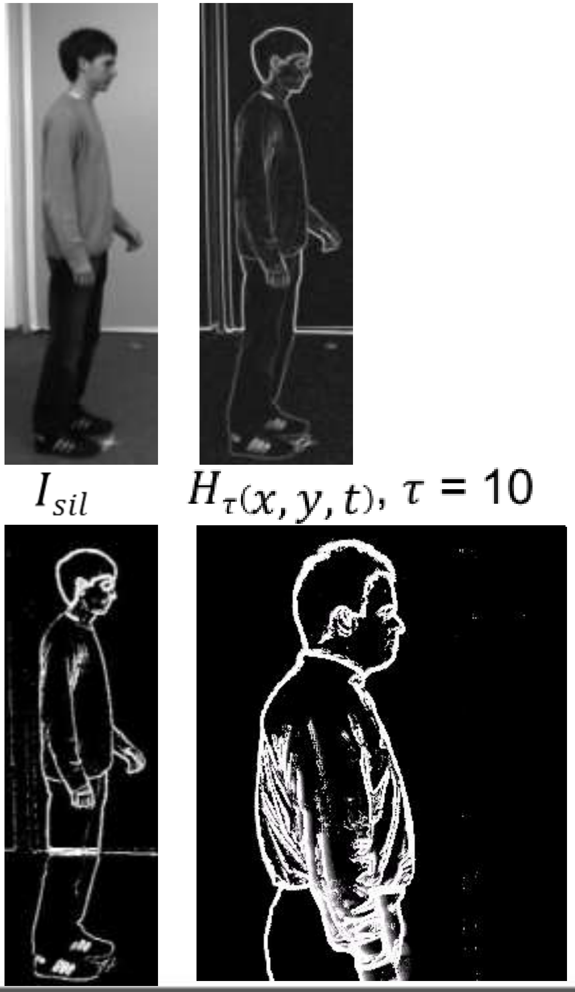
3. Motion History Image

$$H_\tau(x,y,t) = \begin{cases} \tau & if\ \Psi(I_{sil}(x,y,t)) \neq 0 \\ \max(0, H_\tau(x,y,t-1)-1) & otherwise \end{cases}$$

τ: Decay value

$\Psi(Isil(x,y,t))$: Binarized image sequence

I_{gray} $\|I_{mag}\|$



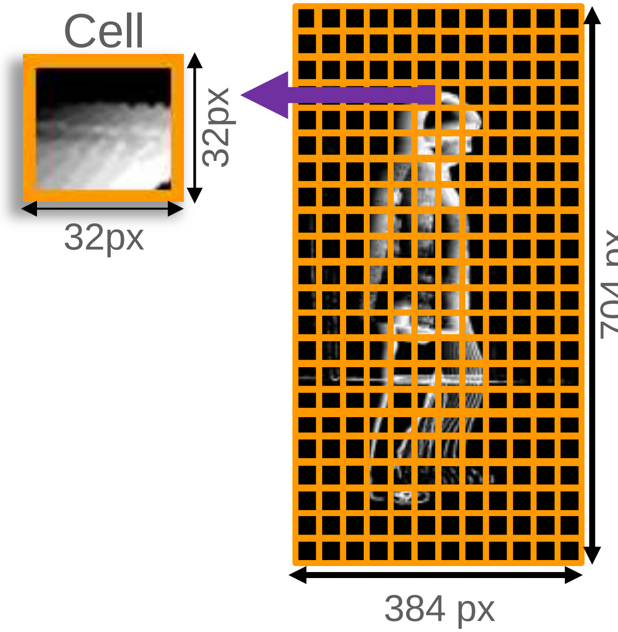
I_{sil} $H_\tau(x,y,t)$, $\tau = 10$

Motion Contour Histograms of Oriented Gradients

1. Compute magnitude and orientation from 1st order derivates in x- and y-direction

Motion **C**ontour **H**istograms of **O**riented **G**radients

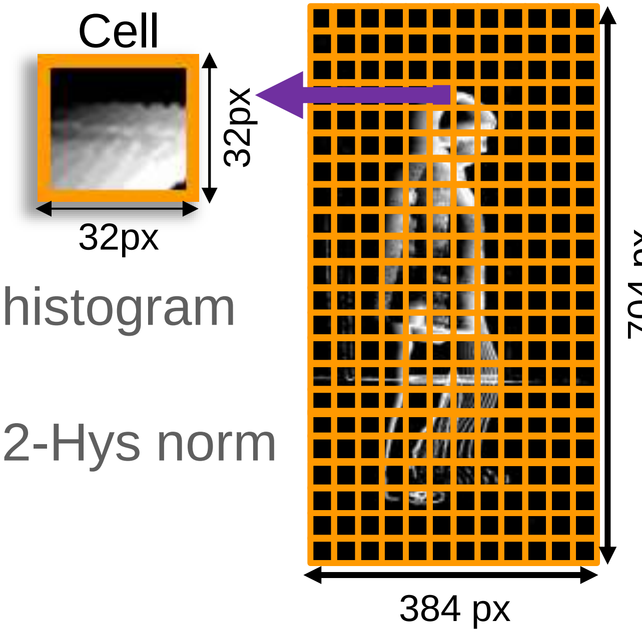
1. Compute magnitude and orientation from 1st order derivates in x- and y-direction
2. Divide detection window in cells



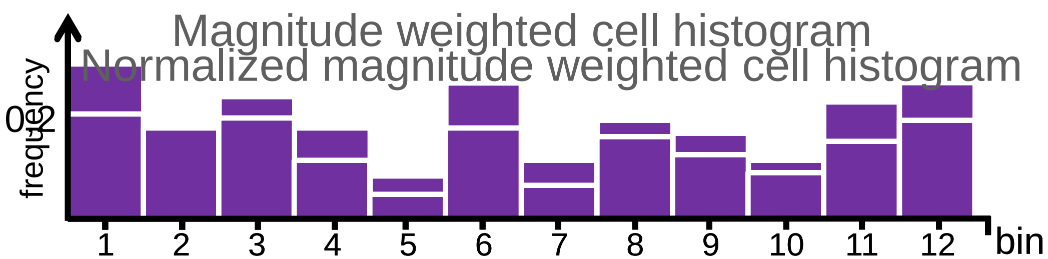
FORSCHUNGSINITIATIVE KO-FAS

Motion Contour Histograms of Oriented Gradients

1. Compute magnitude and orientation from 1st order derivates in x- and y-direction
2. Divide detection window in cells



3. Compute a magnitude weighted orientation histogram for each cell
4. Normalize cell histograms by applying the L2-Hys norm



L2 norm:

$$v \rightarrow \frac{v}{\sqrt{\|v\|_2^2 + \varepsilon^2}}$$

5. Linear 2-class soft-margin support vector machine

2. Our Approach

Overview | Edge Based MHI | **MCHOG Descriptor**

Main differences between HOG and MCHOG procedure

1. Input

HOG	MCHOG
 Camera image	 Edge-based MHI

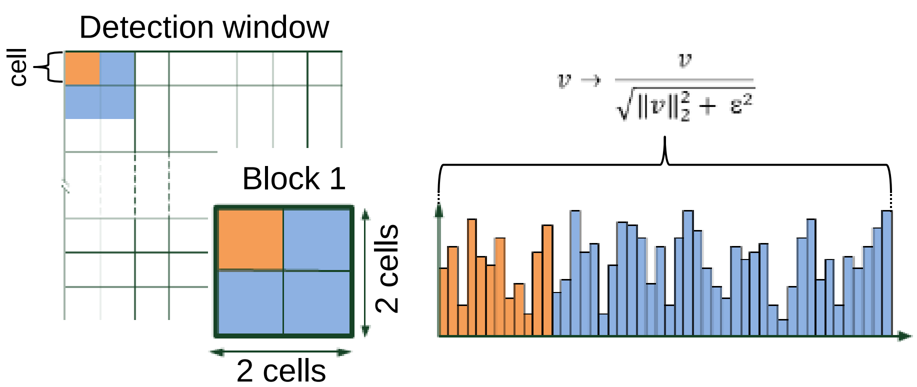
We want to capture the local magnitude and orientation of movement

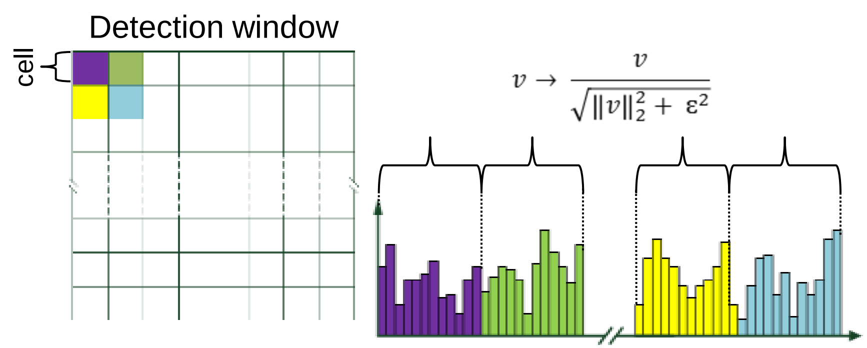
2. Our Approach

Main differences between HOG and MCHOG procedure

2. Normalization

HOG	MCHOG
Block normalization scheme	Cell normalization scheme



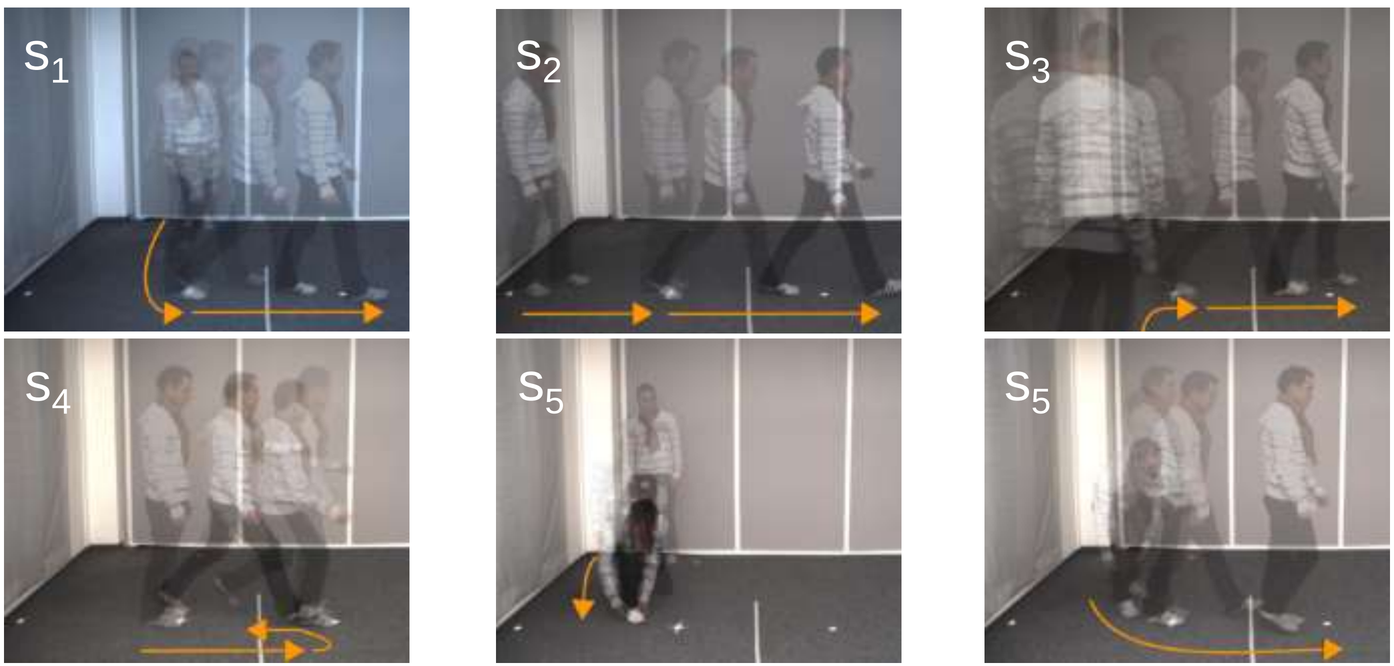


A block normalization scheme would reduce the local difference between neighboring cells whereas the very same should be captured

3. Results

Dataset | Quality | Response Time

- Data acquisition
  - 1128 x 752 px, 50 fps
  - Daylight in laboratory environment
  - 170 videos of 26 adult persons
  - Predefined scenarios
- Scenarios



- In co-operation with Continental Safety Engineering International GmbH

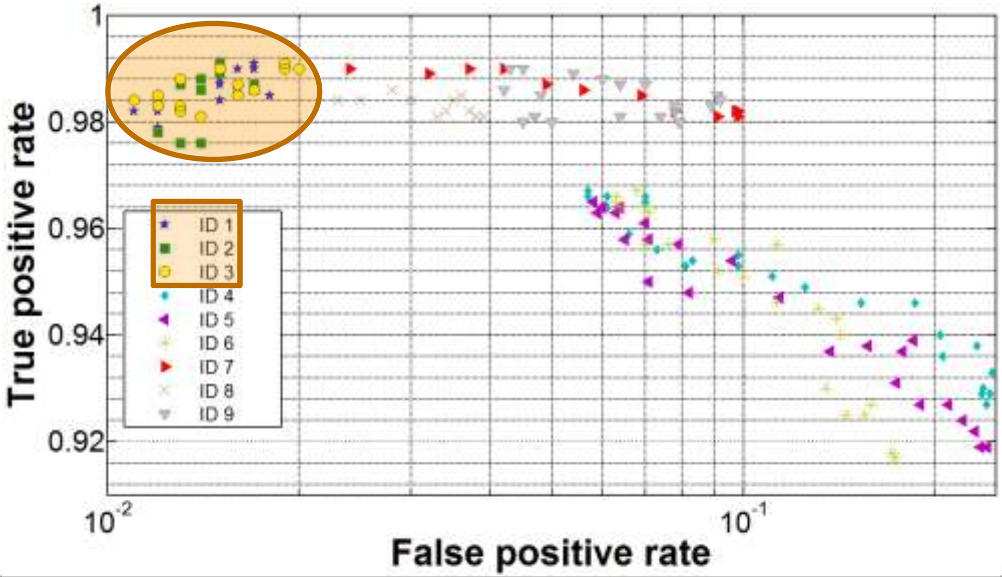
- Descriptor Settings

ID	Cell height (px)	Cell width (px)	Bins	TP rate	TN rate	FP rate	FN rate
1	32	32	9	.991	.983	.017	.009
2	**32**	**32**	**12**	.991	.985	.015	.009
3	32	32	15	.988	.987	.013	.012

Varied parameters:

- Cell size
  - Fine, coarse
  - Quadratic, rectangular
- Number of orientation bins

- Classification results on evalution set (1000 samples per class)



Inference:

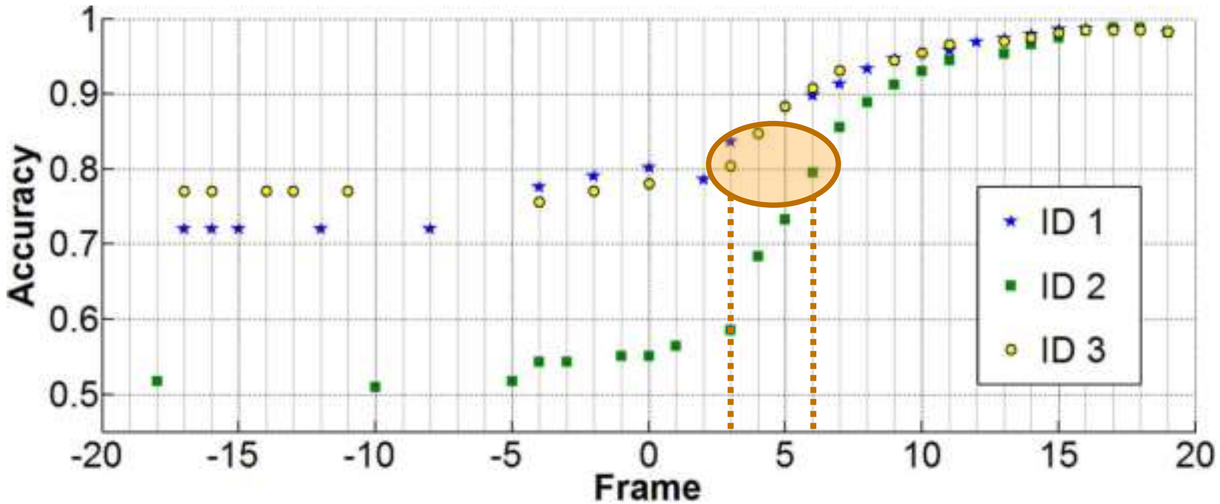
- Quadratic cells are recommended (32 x 32 px)
- Number of bins has only minimal impact (we choose 12)

3. Results

Dataset | Quality | **Response Time**

FORSCHUNGSINITIATIVE KO-FAS

- Span of time necessary for recognition



$$Accuracy = \frac{TP + TN}{P + N}$$

$Refresh\ rate:$
$50Hz = 20ms$

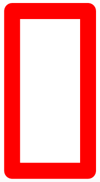
After 60-120 ms: >80% accuracy

Typical time for first step: ~ 400 ms

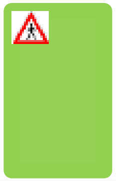
Outcome

In relation to typical gait data the result corresponds to a detection within the time span of the first step

3. Results



Red rectangle:
Pedestrian ROI, **Standing**



Superimposed green rectangle:
Pedestrian ROI, **Walking**

4. Conclusion

FORSCHUNGSINITIATIVE KO-FAS

- Motion contour image based HOG-like descriptor – MCHOG
- High accuracy levels within the first step
- Designed for monocular stationary intersection monitoring
- Evaluated at laboratory conditions
- Shown feasibilty in a real world intersection

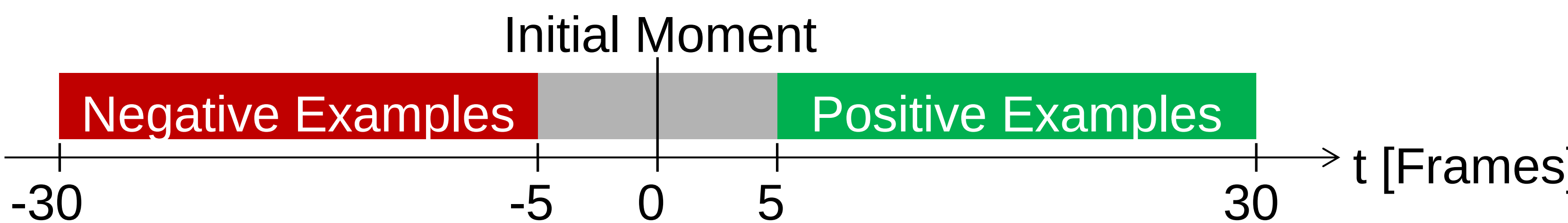
Thank you

Visit our driving demonstration (3) tomorrow

Results (Additional Slides)

Dataset | Quality | Response Time

- Labeling procedure
  - Pedestrian ROIs with HOG/SVM (with subsequent visual inspection)
  - Manual labeled initial moment (Heel-Off)
- Training and evaluation sets (images)
  - Disjunctive video set for training and evaluation
  - Using following scheme for positve and negative sample generation



- 4,663 samples total

	Negative	Positive
Training	963	1,700
Evaluation	1,000	1,000

Negative Examples Positive Examples