Monocular Computer Vision for Driver Assistance

Reinhard Klette

Mahdi Rezaei, Simon Hermann,
Ruyi Jiang, and Bok-Suk Shin

The .*enpeda*.. Project
The University of Auckland - New Zealand

Vision-based Driver Assistance

What is Vision-based Driver Assistance (VbDA)?

Why Vision-based Driver Assistance?

Proofs of Existence

Performance Requirements and Evaluation

Adaptive Solutions

In this talk:

What if monocular vision only?

How far can we go with **monocular camera data only**
(i.e. no stereo vision, no radar, no LIDAR, etc.)

F. Ren et al., ArtsIT, 2009

Motivation

Can be an add-on for existing cars

Appropriate for mobile devices (with limited base line)

No need to buy a car where cameras are already integrated

Solutions Adaptive to Given Situations

A **situation** (or **scenario**) is a combination of events defining a class of video sequences of about 3-5 sec length

Still an open problem:

How to recognize a situation *in real time during recording* such that algorithms (to be used for analyzing the data) can adapt accordingly? There is no “best algorithm”, just “best under particular conditions”.

Situation: "Default driving conditions"

"Normal day-time highway" conditions

Situation: “Close objects”

Stopping and close objects

Situation: "Inner-city at night"

Night time with dense traffic and multiple lights

Situation: "Brightness changes"

Changing angle to the sun and clouds

Situation: “Illumination artifacts”

Local light changes in recorded images

1: VbDA for Driver Perfection

Driver Monitoring (Sleepiness ...)

Providing Night Vision

Providing Defogged Views

Weather-defined Road Conditions

Virtual Windshield as Interface

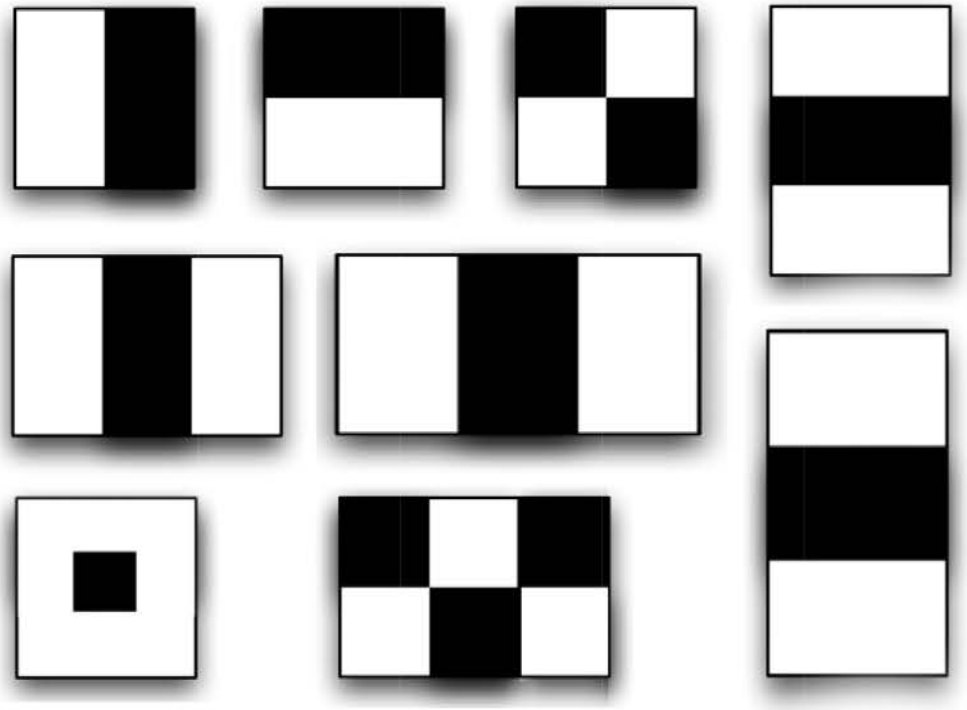
Helping Visually Impaired Drivers

Difficult if only monocular vision

Light Control (for Oncoming Traffic)

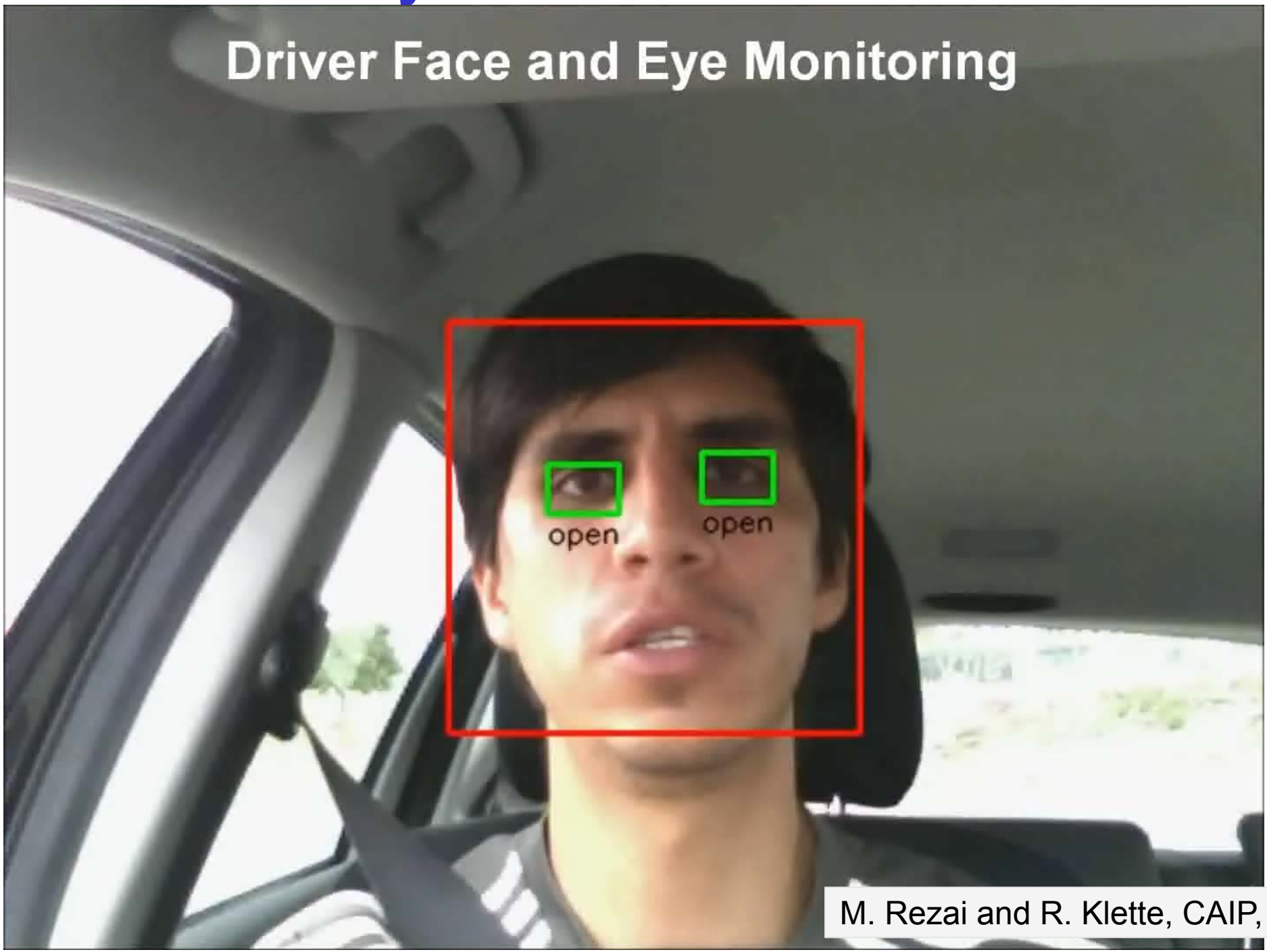
Monocular Face Detection & and Eye Status

Haar-like features and training of strong classifiers



P. Viola and M. Jones, IJCV, 2001

2D Face and Eye Detection

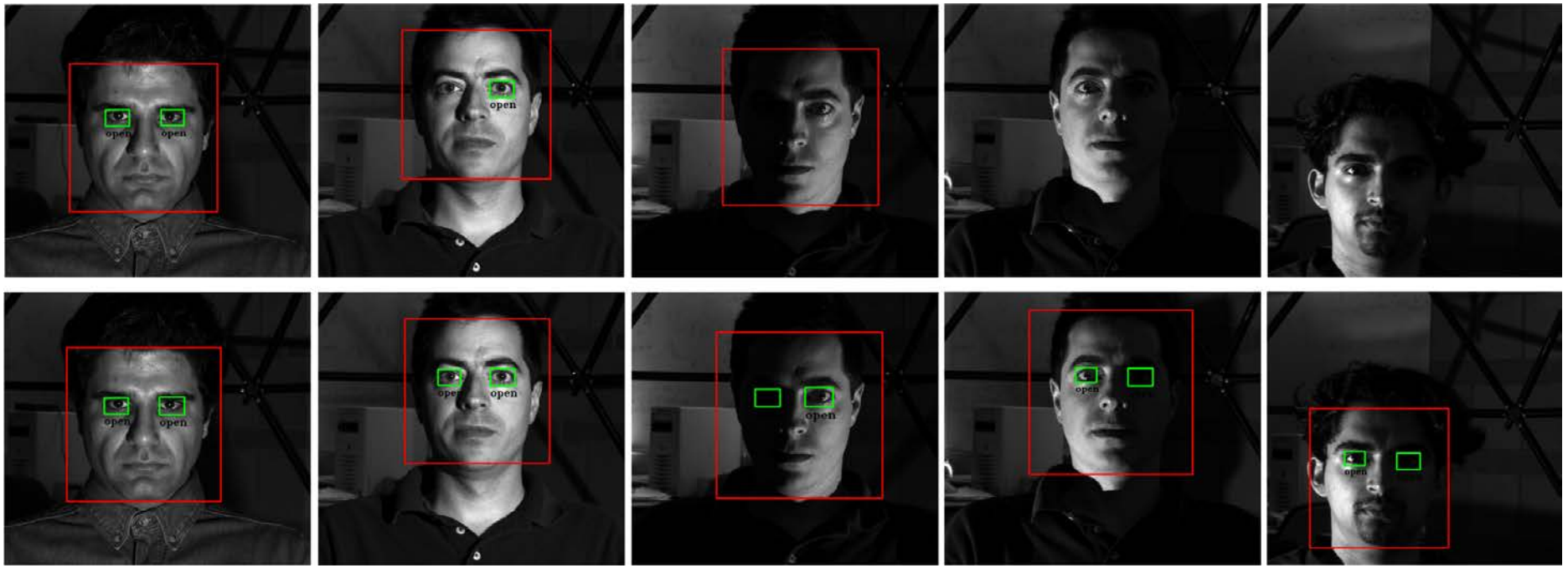


M. Rezai and R. Klette, CAIP, 2011

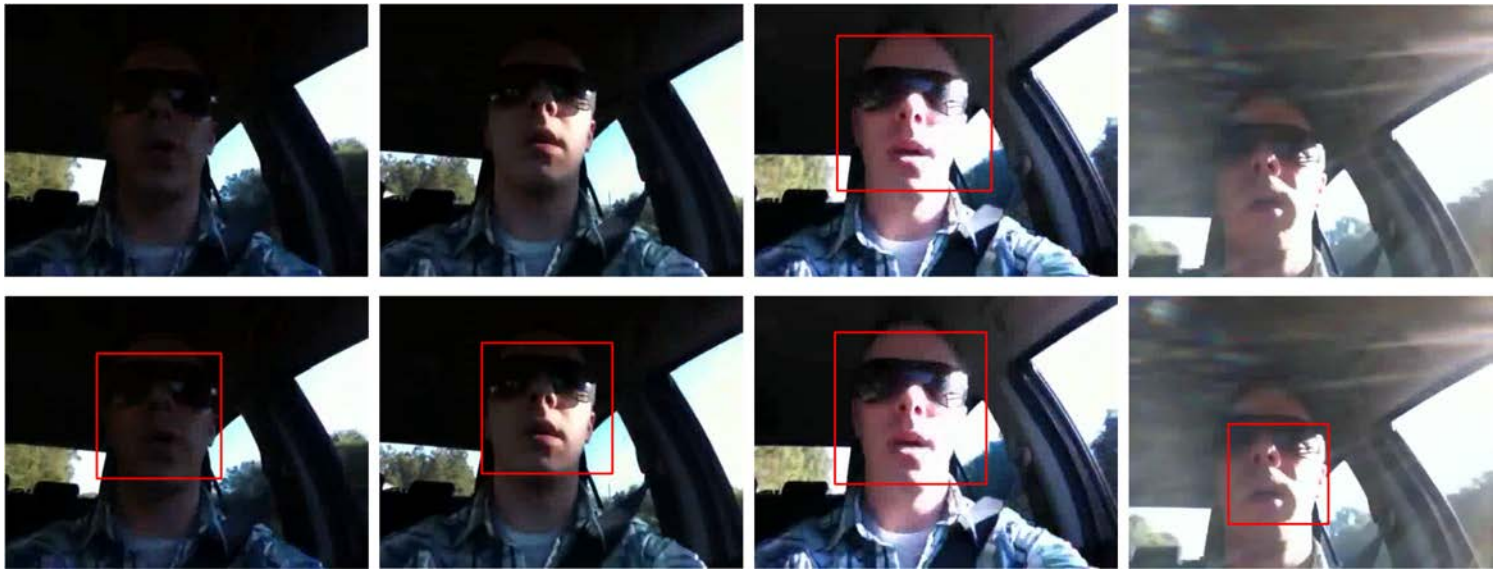
Hon. Dr. **Wan Gang**, Minister of Science and Technology of PRC, when visiting the University of Auckland in 2011:

" Every truck or long-distance bus in China should be equipped with such a driver monitoring system."

Challenging Lighting Conditions

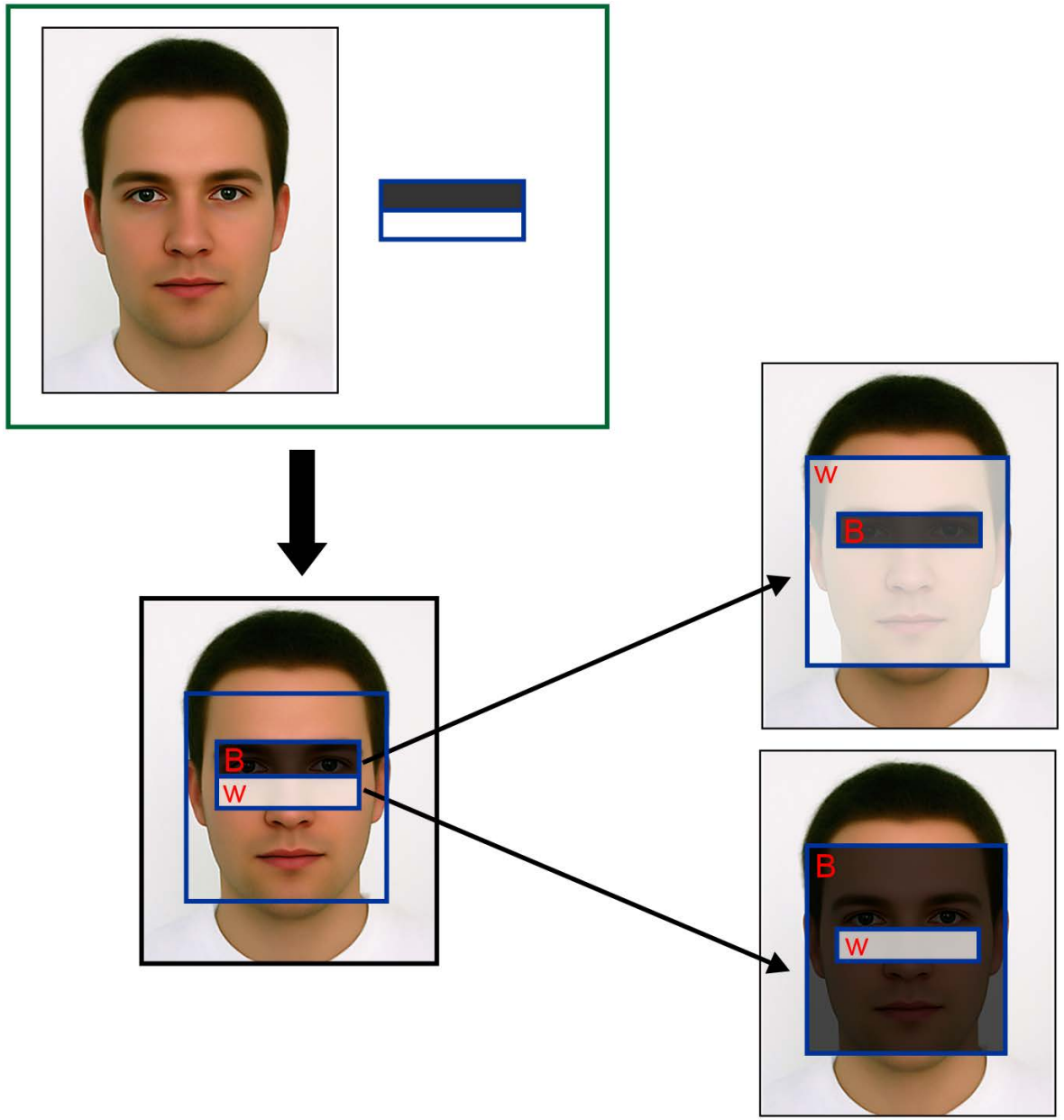


Standard VJ *versus* asymmetric adaptive global Haar features

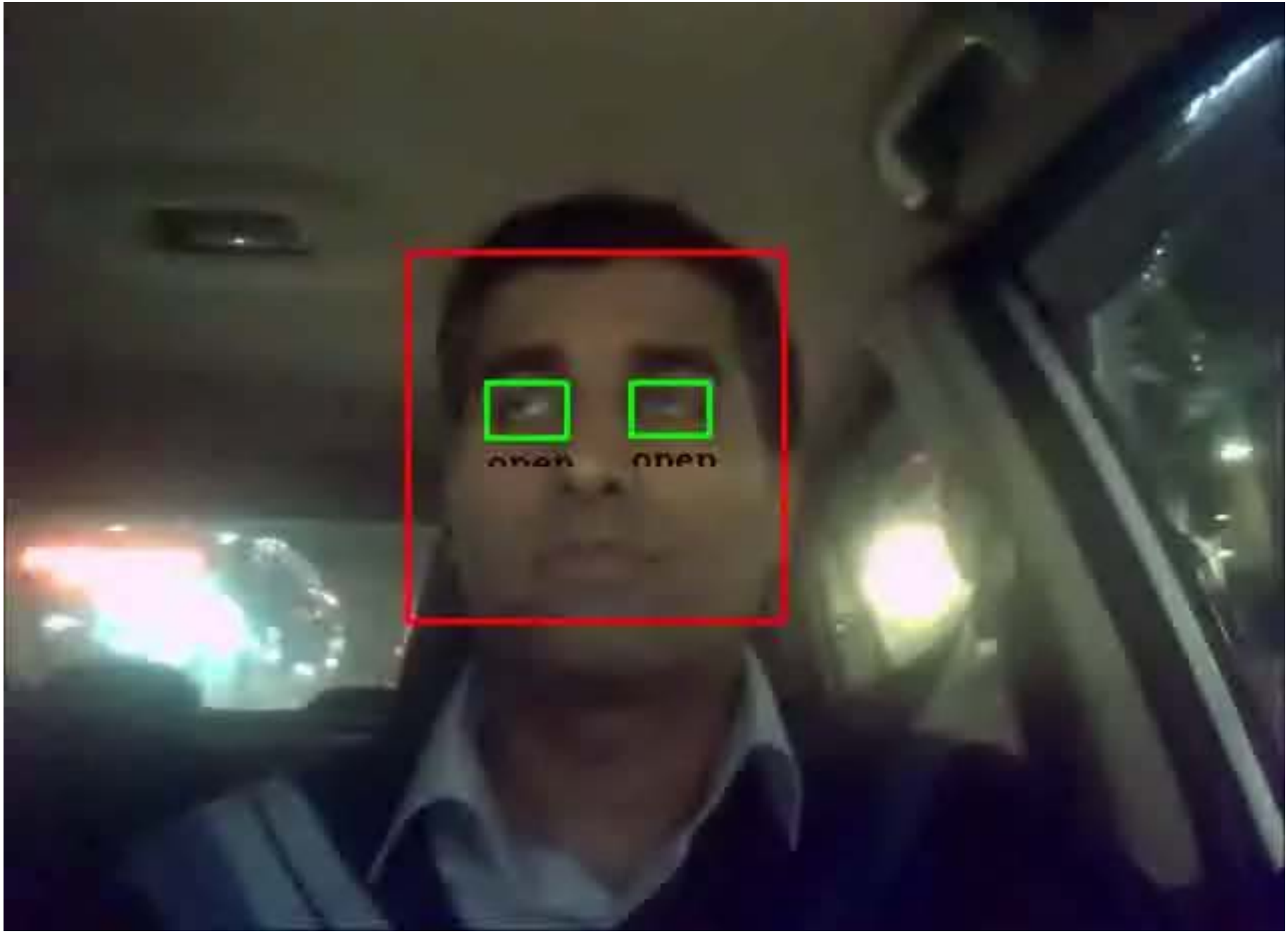


M. Rezaei and R. Klette, Workshop ACCV 2012

Local and Global Haar-Like Features



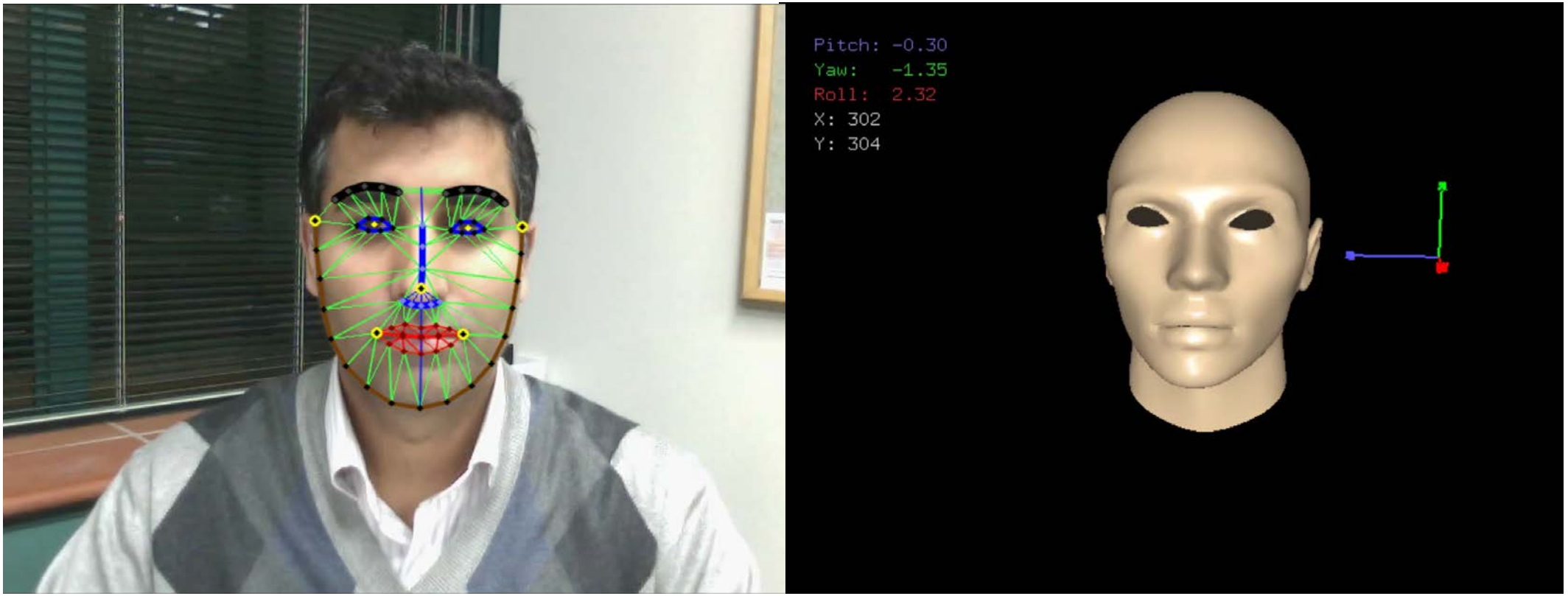
M. Rezaei and R. Klette, PSIVT, 2013



Combined with the use of a Kalman filter for face tracking

M. Rezaei and R. Klette, Workshop ACCV 2012

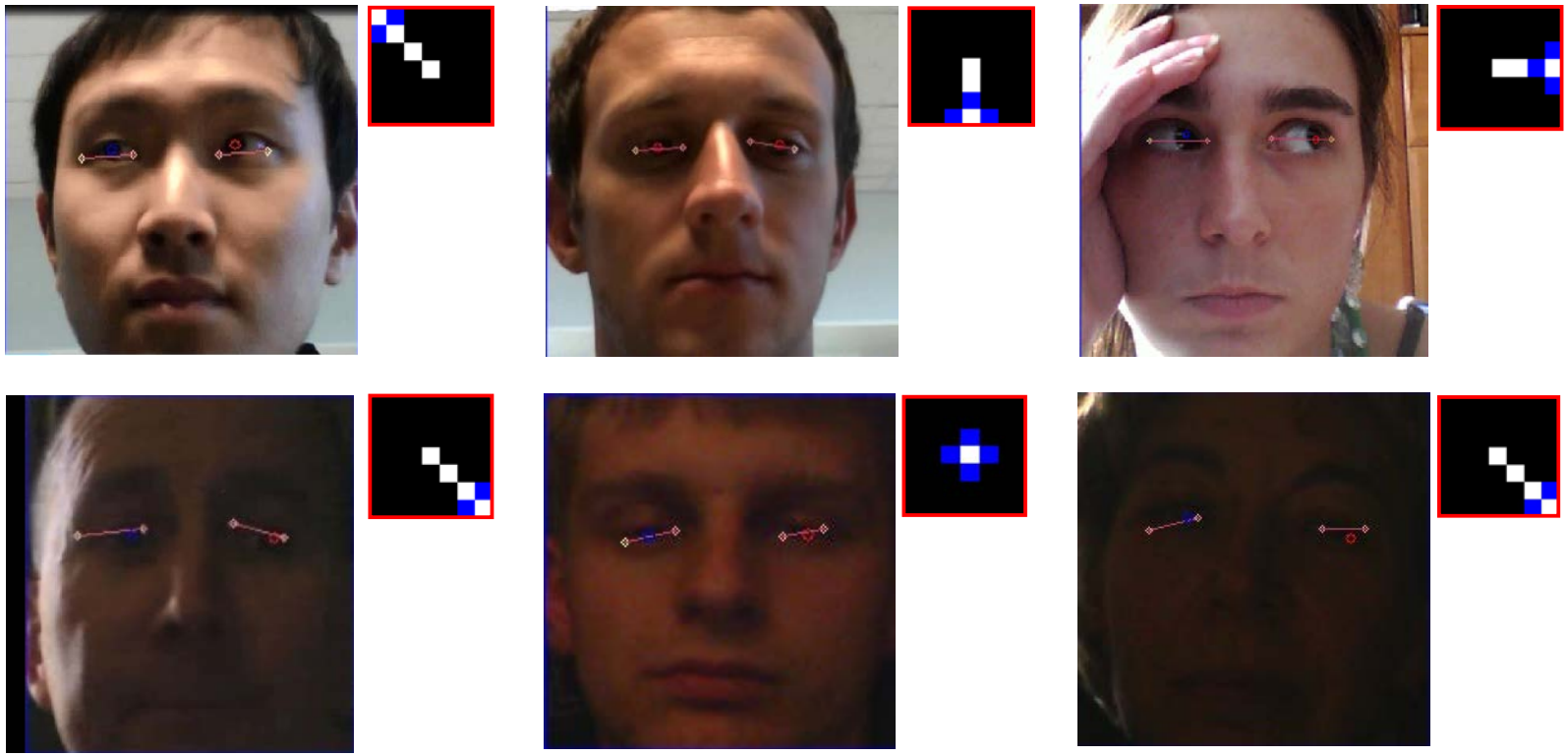
Monocular 3D Head Pose and Yawning Detection



Detection of facial features & mapping on generic face model using an asymmetric appearance model

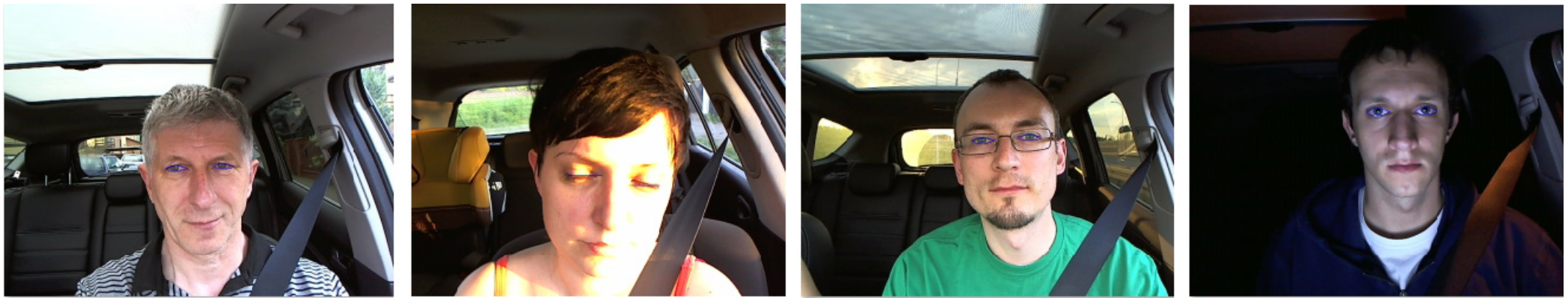
M. Rezaei, 2013

Monocular Eye Gaze and Eye Status Detection



J. Wiśniewska, M. Daniluk, et al., ICCVG, 2014

Test video data on EISATS web site



2: VbDA for Basic Navigation Support

Blind Spot Sur

Parking Assista

Distance Estim

Obstacle Dete

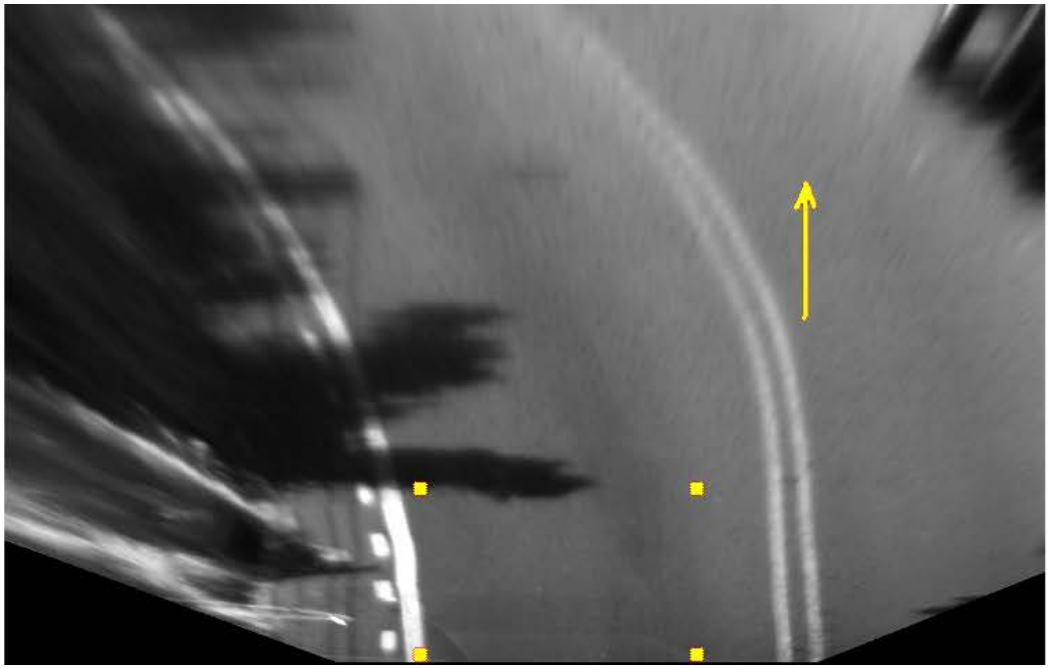
Speed Adaptat

Queuing of Mu

Difficult if only monocular vision

s

Homography into a Bird's-eye View



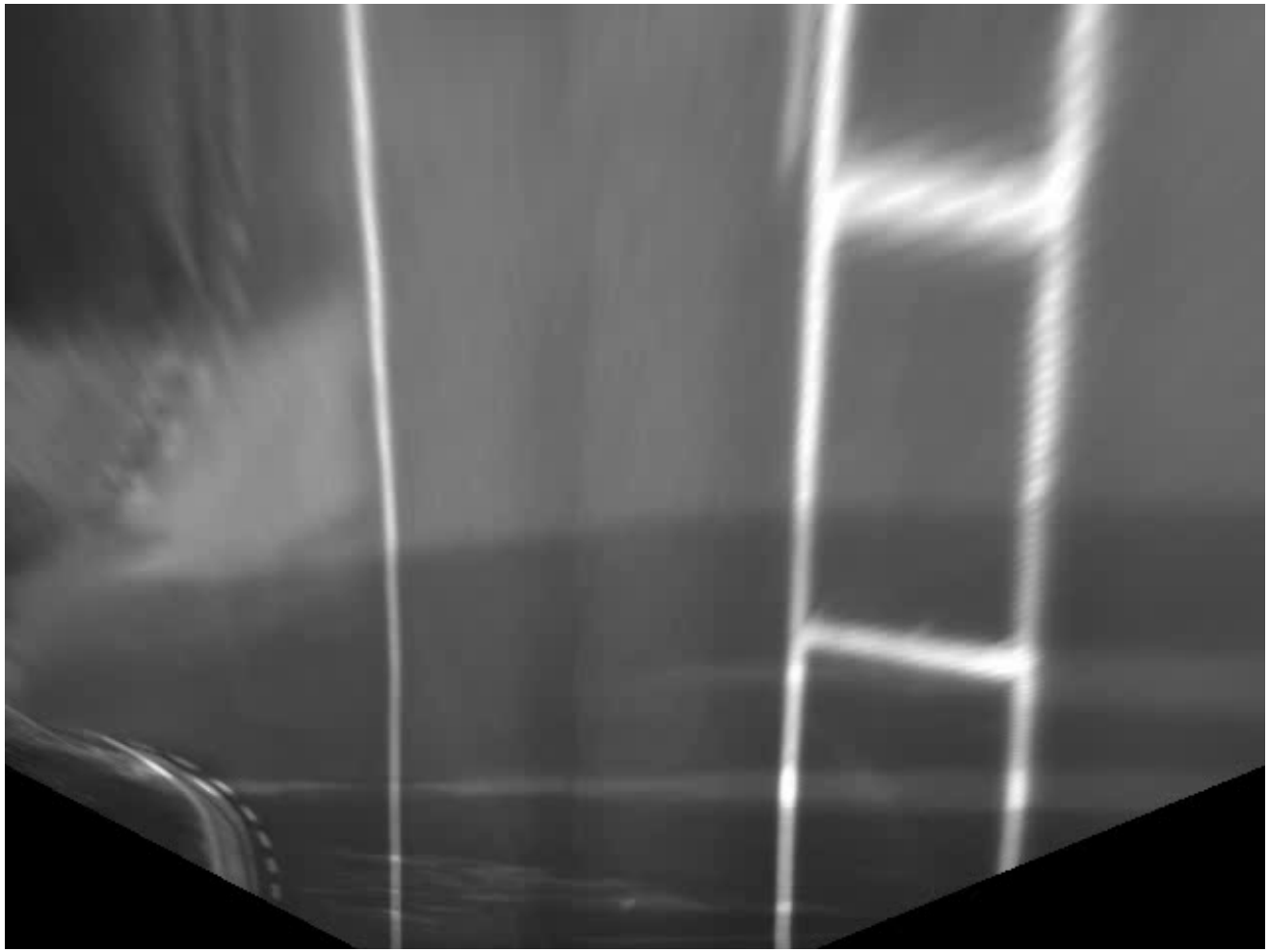
Warp-perspective mapping (simple; see above)

OR

Inverse perspective mapping (using camera calibration data)

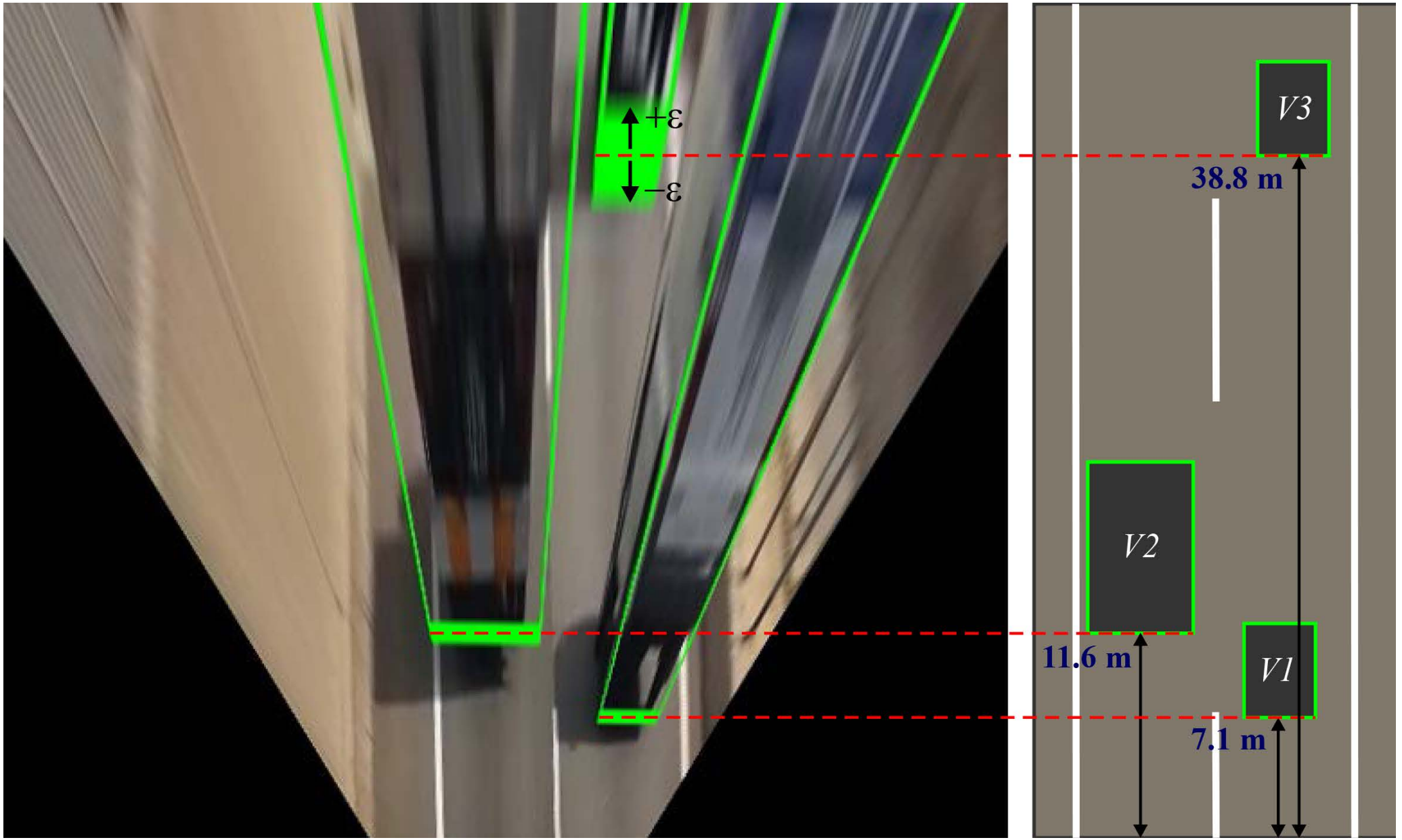
R. Jiang et al., ArtsIT, 2009

Example of Bird's-eye-View Video



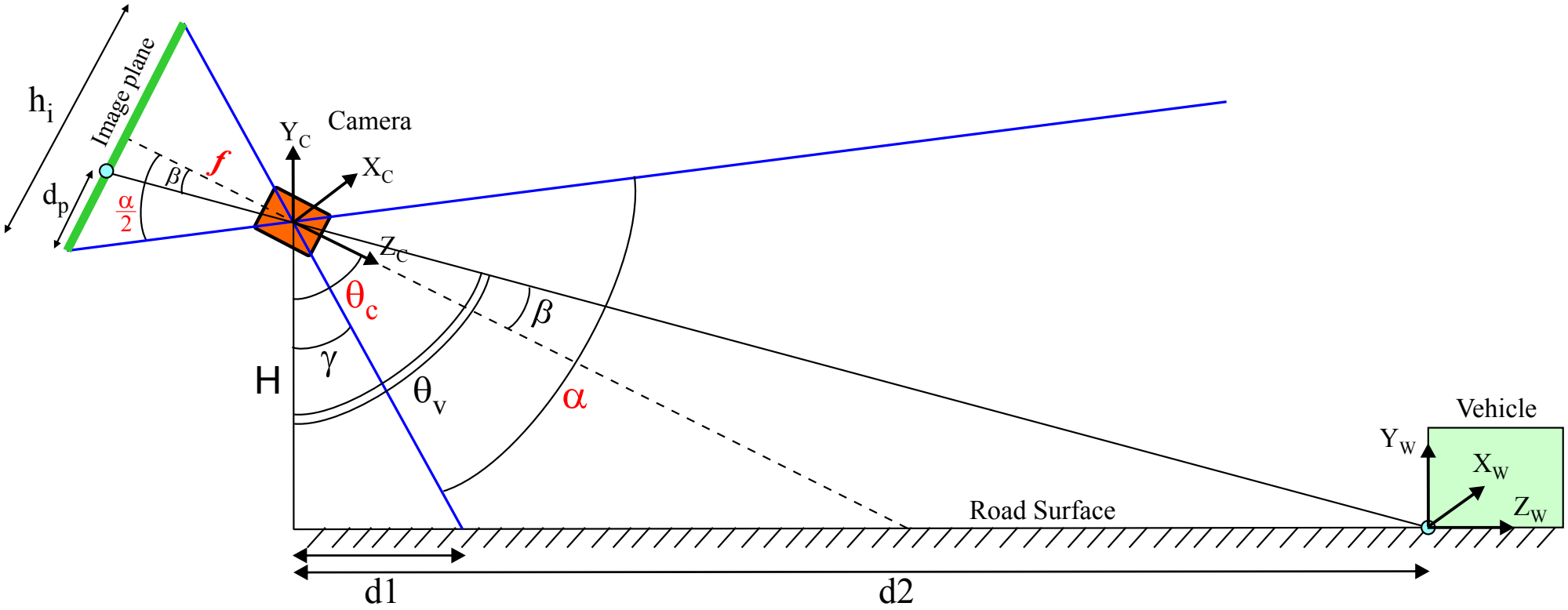
R. Jiang et al., DAGM, 2009

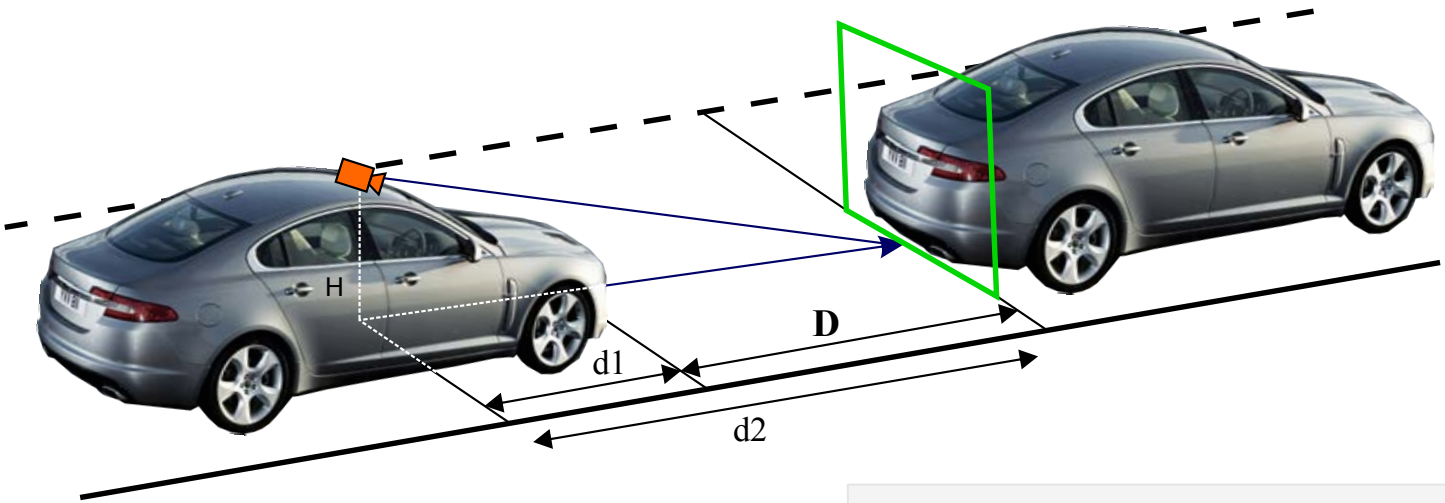
Distance Estimation in Bird's-eye View



M. Rezai and M. Terauchi, PSIVT, 2013

Estimate Distance Using Alpha and Theta





M. Rezai and M. Terauchi, PSIVT, 2013

3: VbDA for Understanding Dynamic Objects

Motion Field Calculation (Optical Flow, ...)

Visual Odometry

Vehicle Detection ai

Pedestrian Detectio

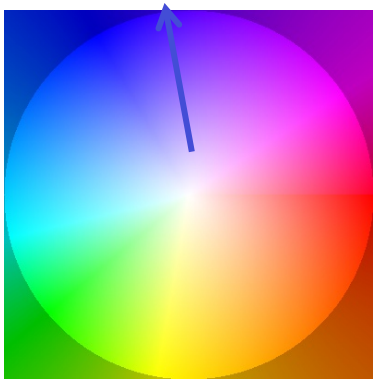
Difficult if only
monocular vision

Motion Analysis

is a 2D (in image plane) correspondence problem

at 25 Hz

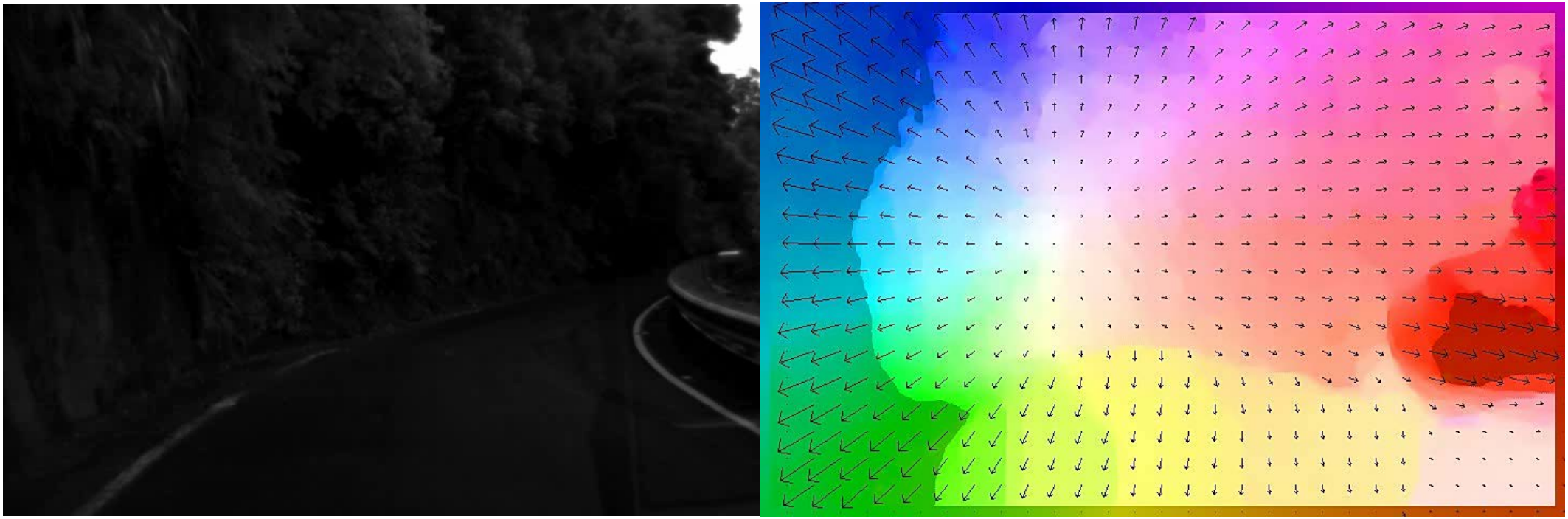
Color key



(u,v)

optic flow – aims at subpixel accuracy

Situation: Illumination artifacts

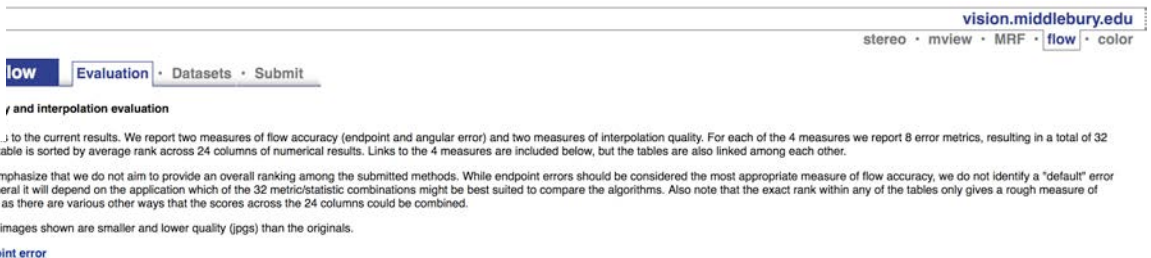


TV-L$_1$ optic flow (10 bit data)

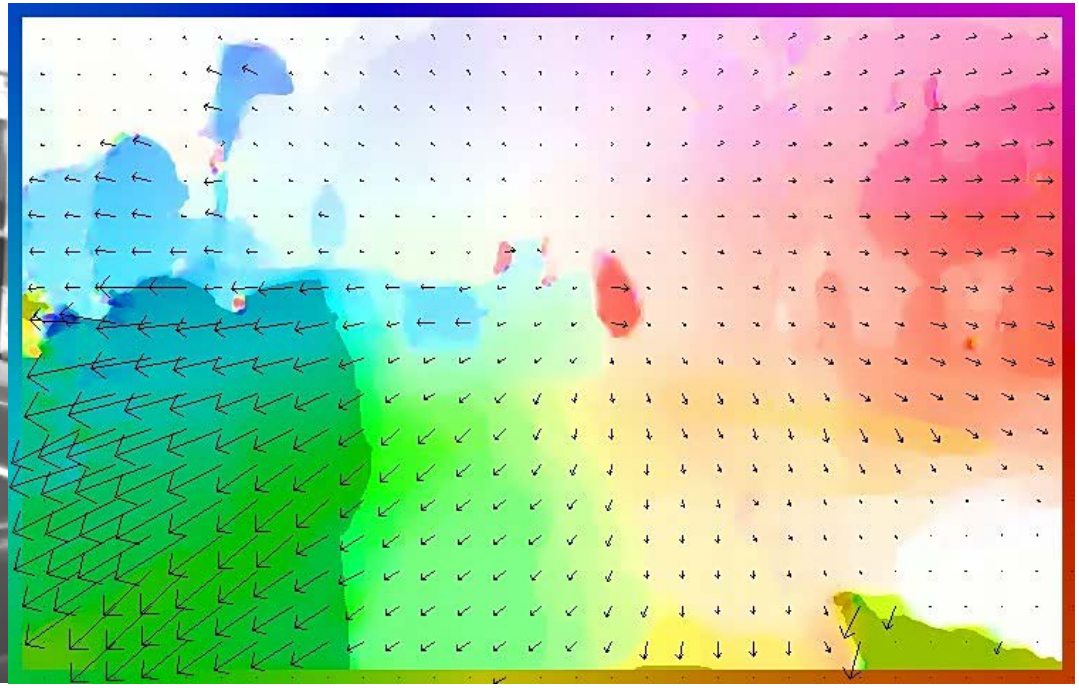
T. Brox et al., ECCV, 2004

Best-paper award at ECCV 2004

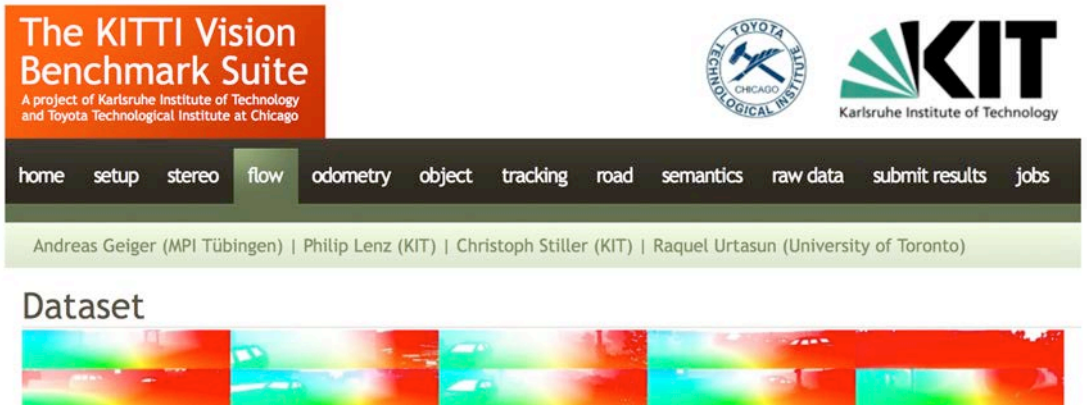
No. **63** on Middlebury,
Optical flow, EP error page
on 18 Nov. 2014



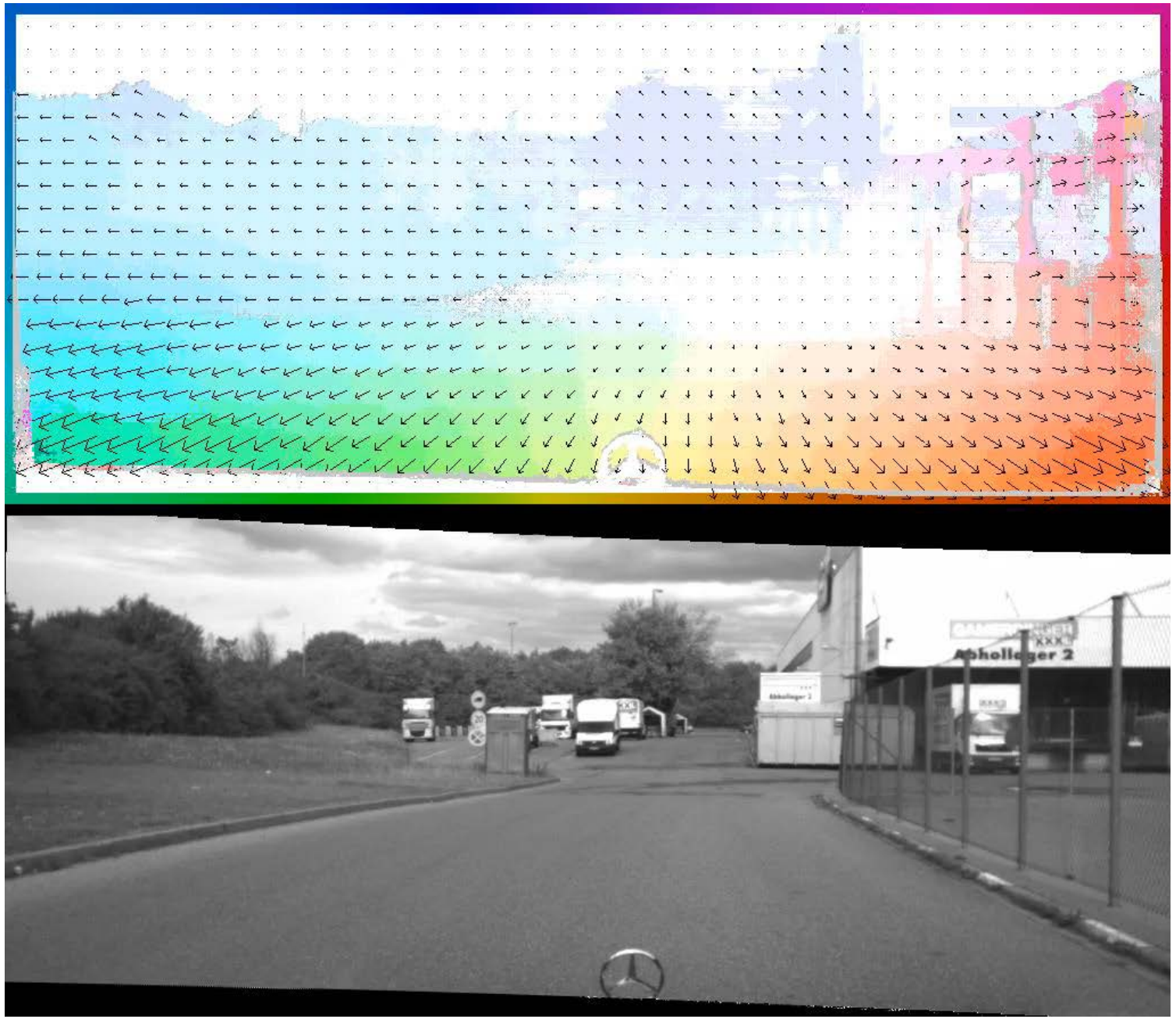
One More Example for *TV-L_1*



Not listed on KITTI, optic flow



Discrete Calculation of Optical Flow



fSGM (semi-global matching for flow)

S. Hermann & R. Klette, ACCV, 2012

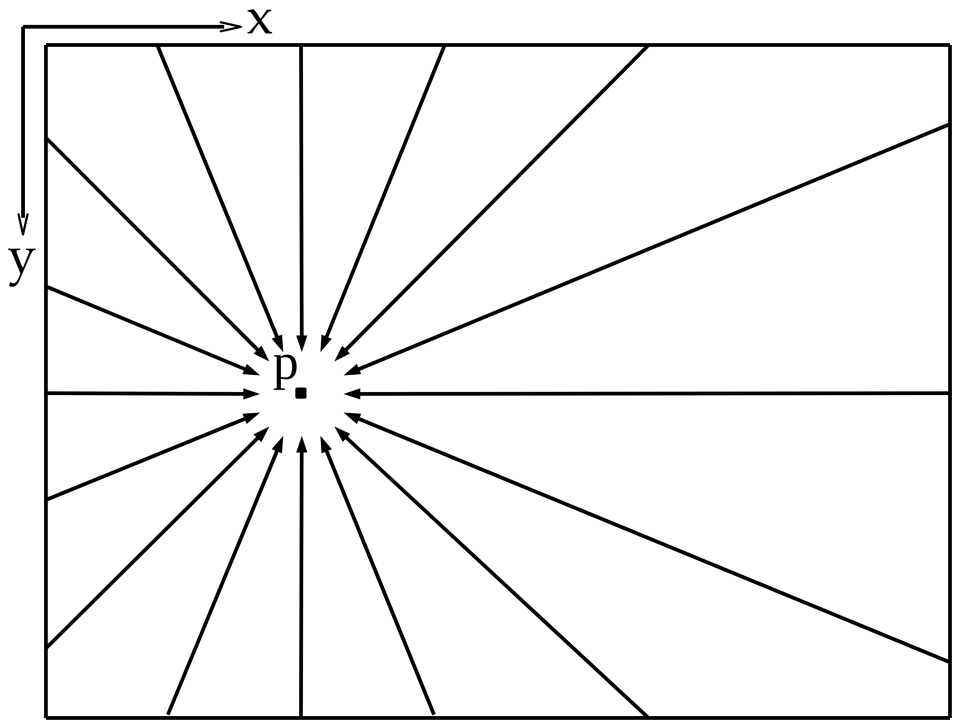
What is *fSGM* ?

fSGM is a discrete optic flow estimator

It uses dynamic programming in combination with the SGM integration strategy (SGM = semi-global matching)

fSGM can handle large pixel displacements

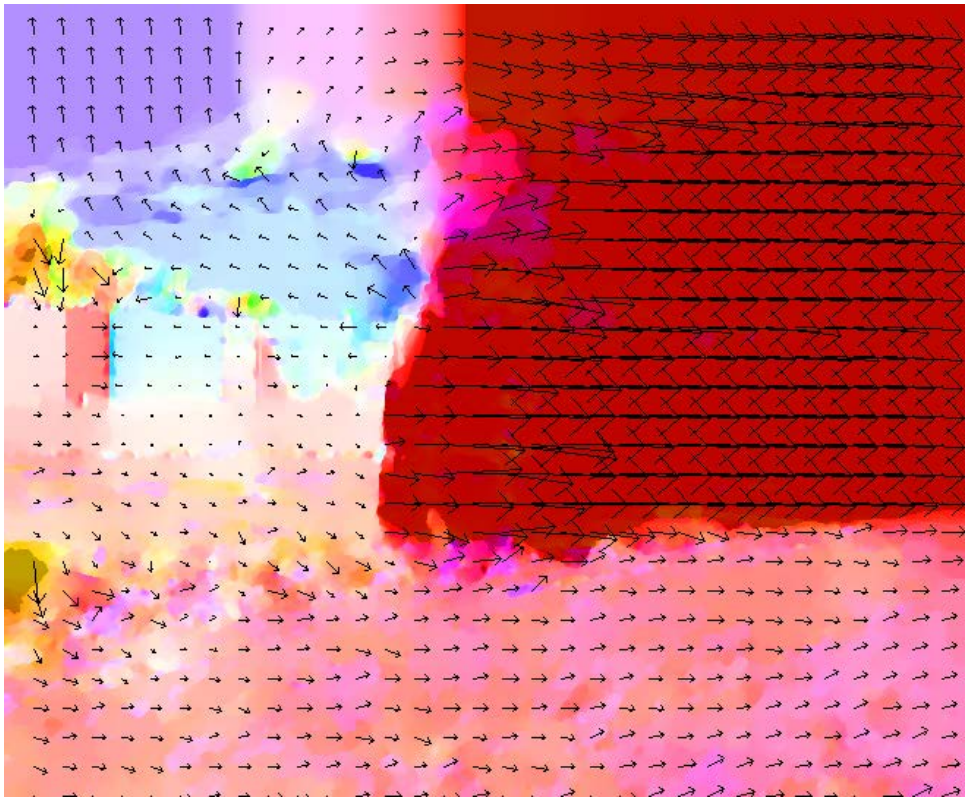
fSGM is a discrete method and thus an alternative to variational methods (which are continuous)



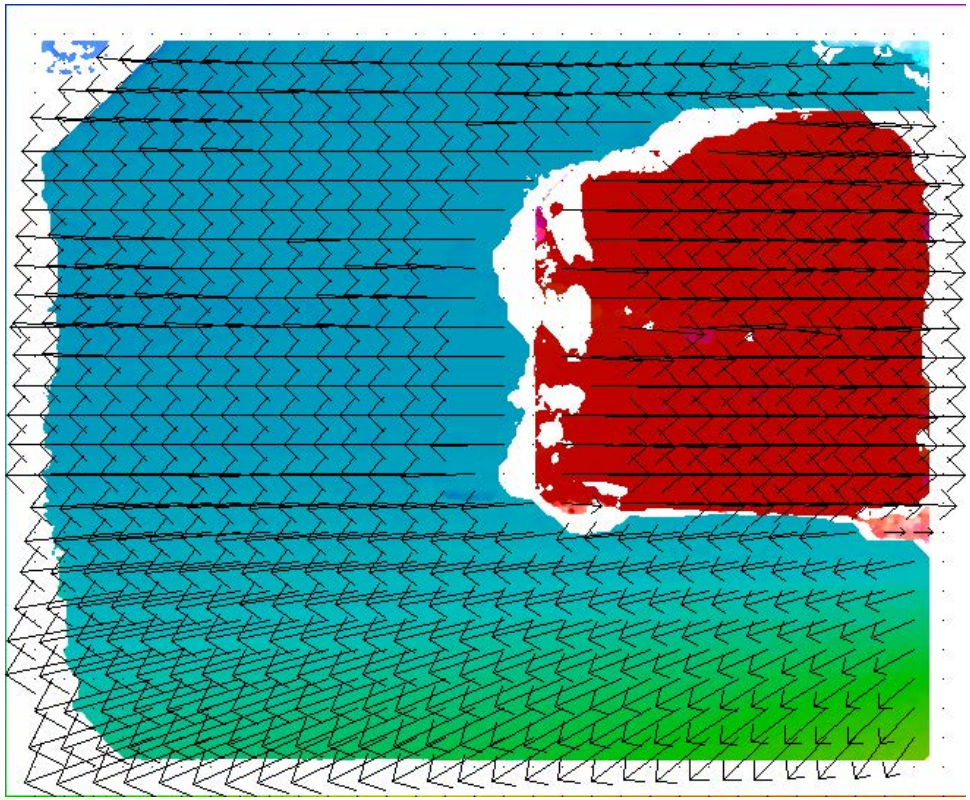
S. Hermann & R. Klette, ACCV, 2012

Variational Method vs. *fSGM*

TV-L_1



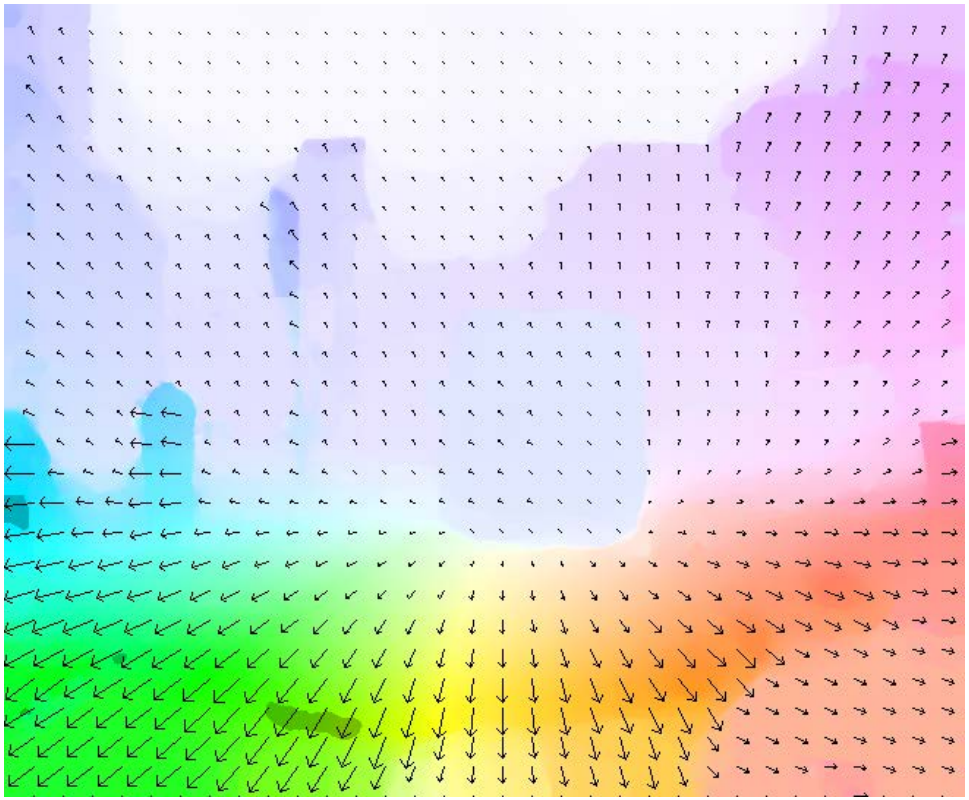
fSGM



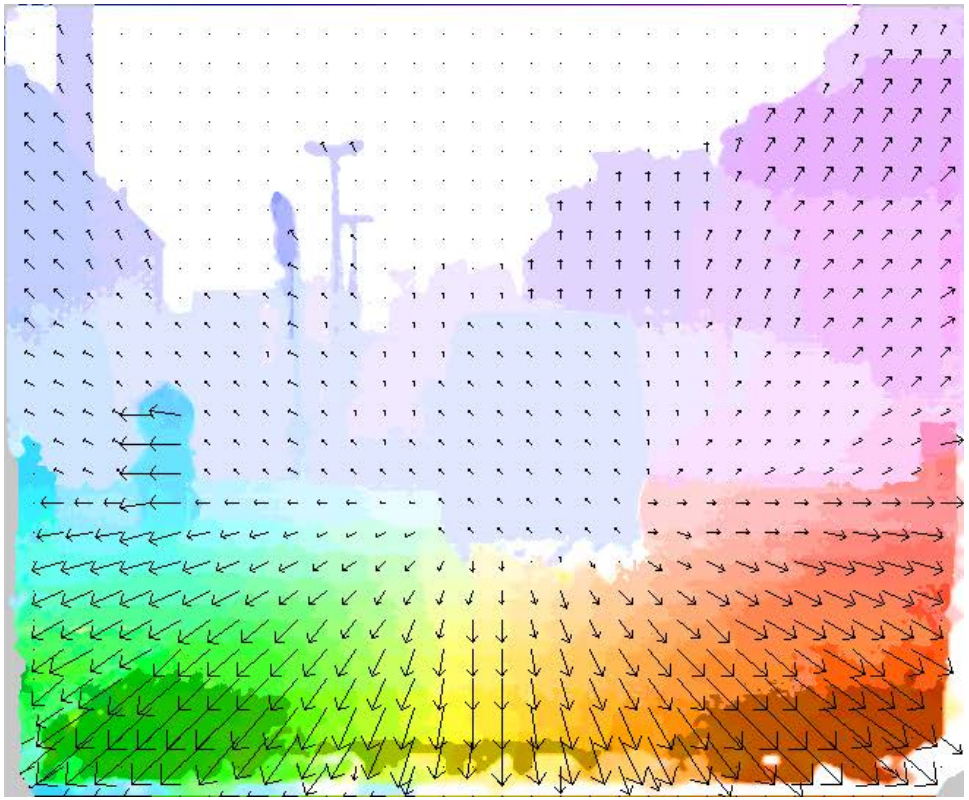
S. Hermann, 2012

Variational Method vs. *fSGM*

TV-L_1



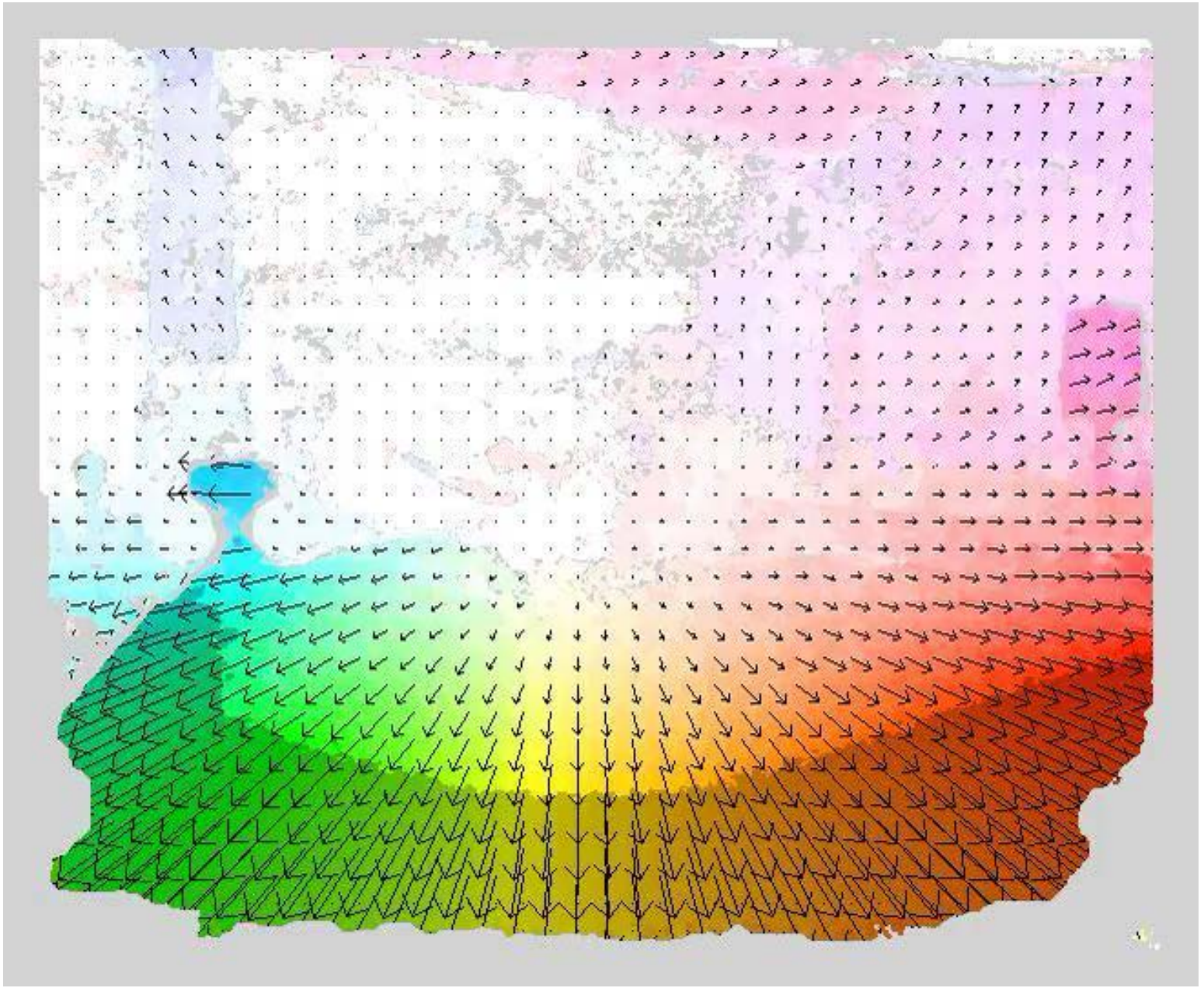
fSGM



S. Hermann, 2012

Robust Vision Challenge

in Association with the 2012 ECCV Workshop on Unsolved Problems in Optical Flow and Stereo Estimation



Jury

Simon Baker
(Microsoft Research)

Goksel Dedeoglu
(Texas Instruments)

Jan Effertz
(Volkswagen Research)

Oliver Erdler
(Sony)

Wolfgang Niehsen
(Robert Bosch GmbH)

Phil Parsonage
(The Foundry)

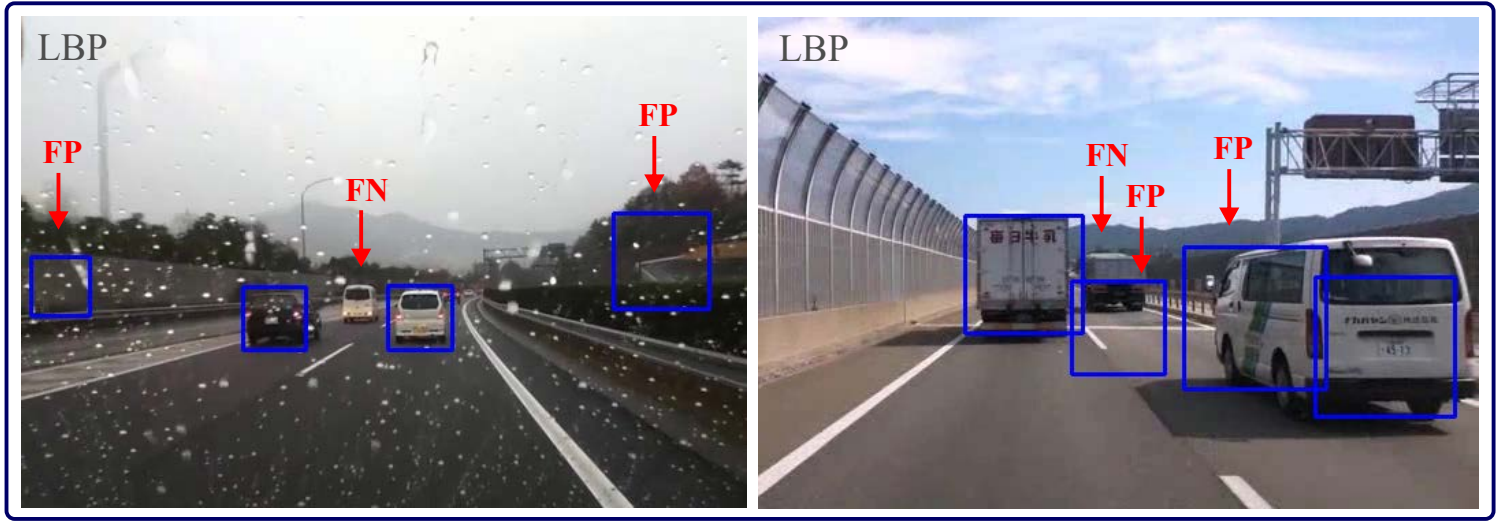
Stephan Simon
(Robert Bosch GmbH)

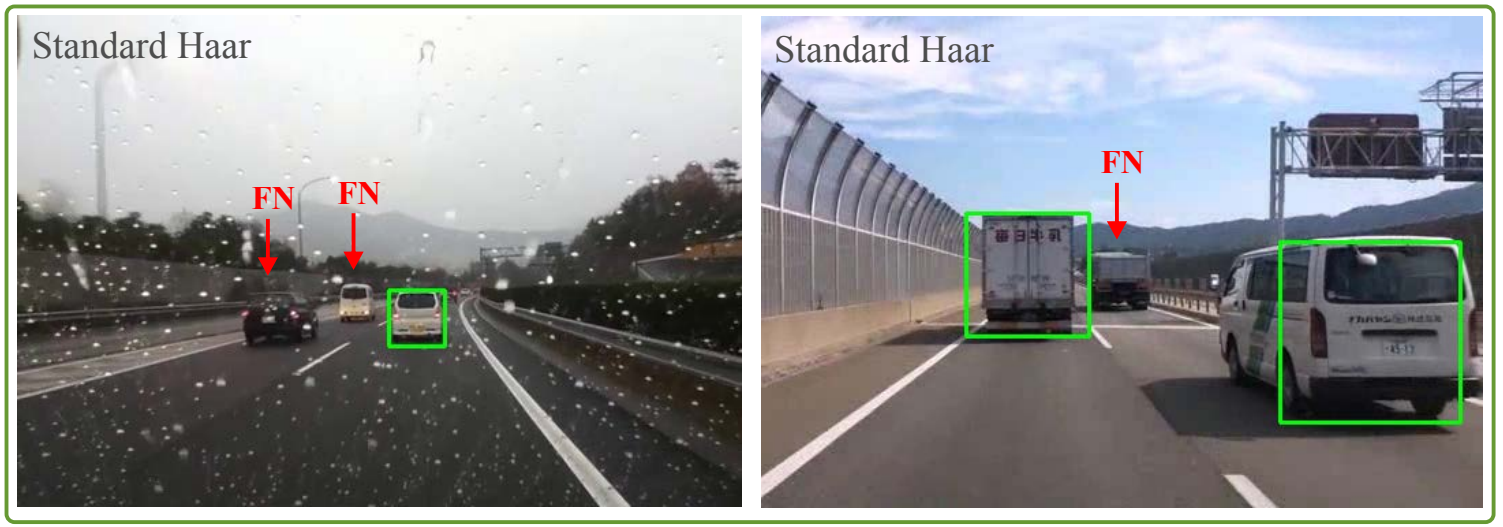
Christian Unger
(BMW Group)

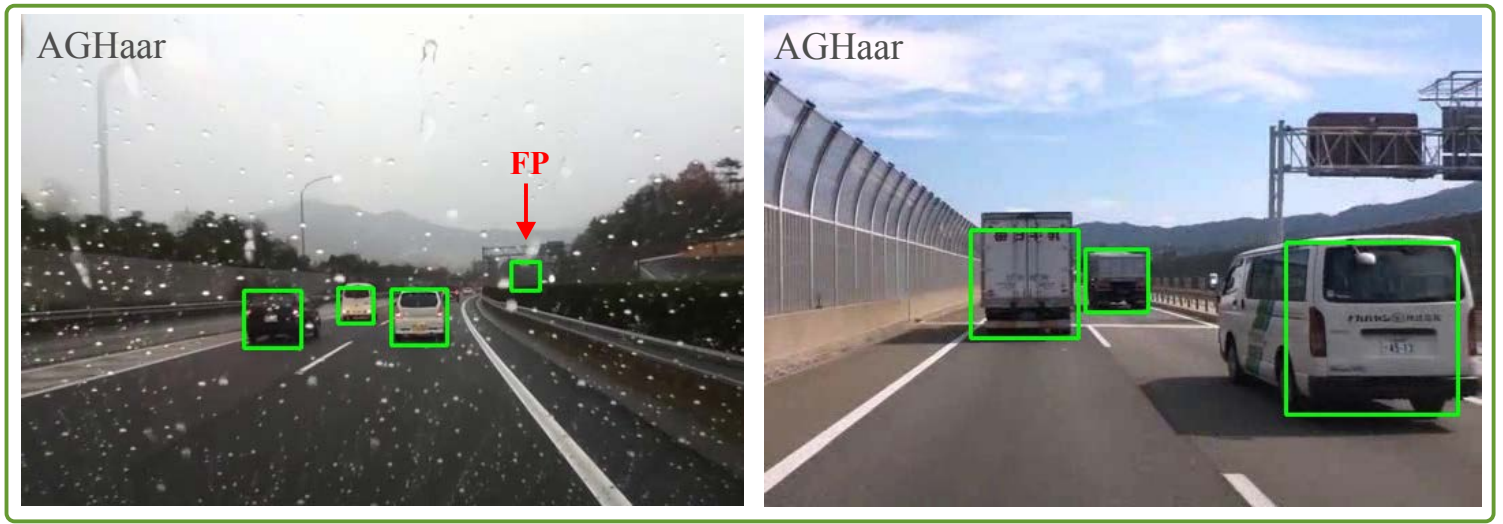
fSGM is on 18 Nov 2014 listed as no. **28** on KITTI, optical flow

Vehicle Detection

(also for challenging situations)

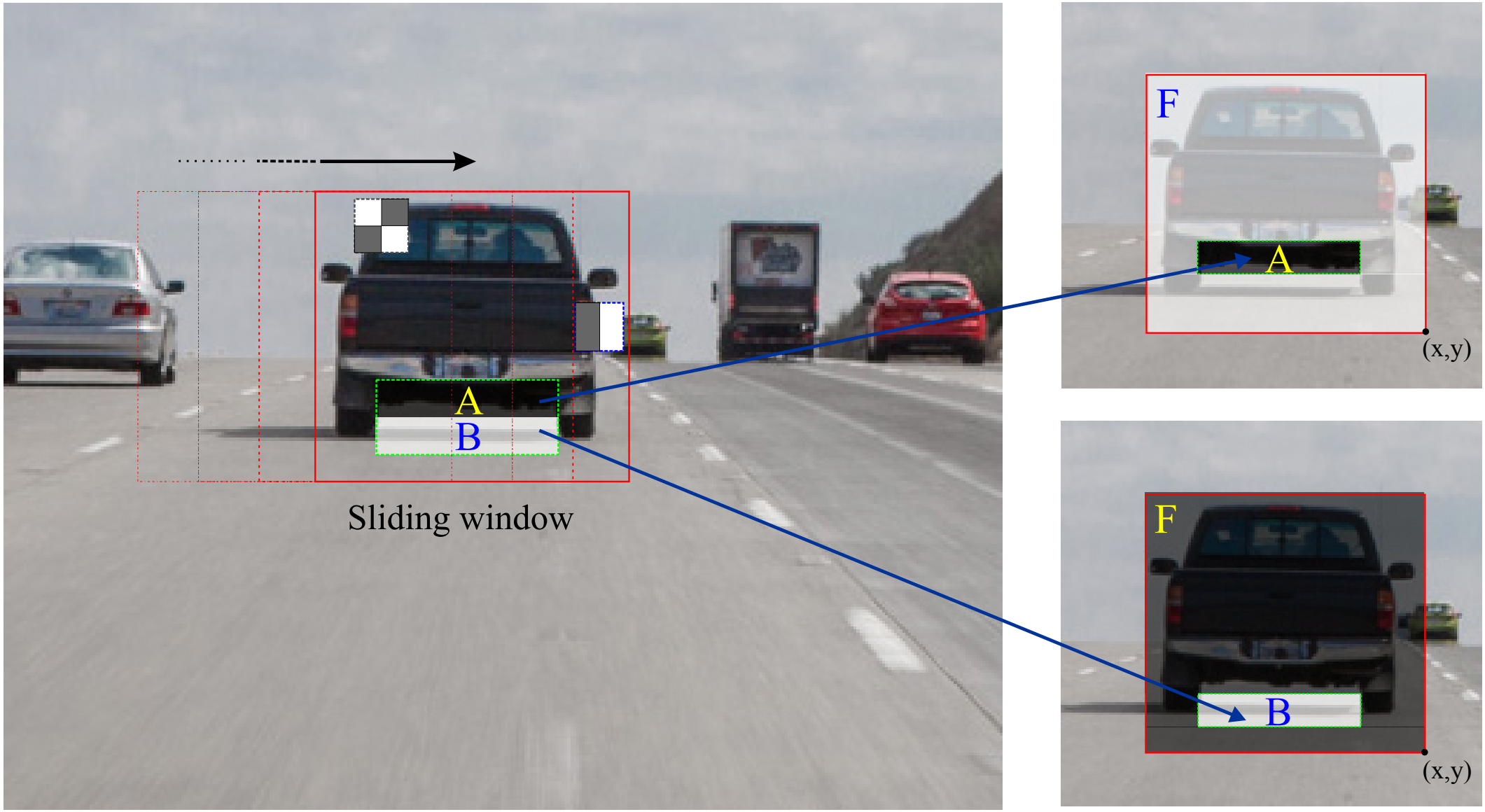






M. Rezaei,
M. Terauchi, and
R. Klette,
submitted, 2014

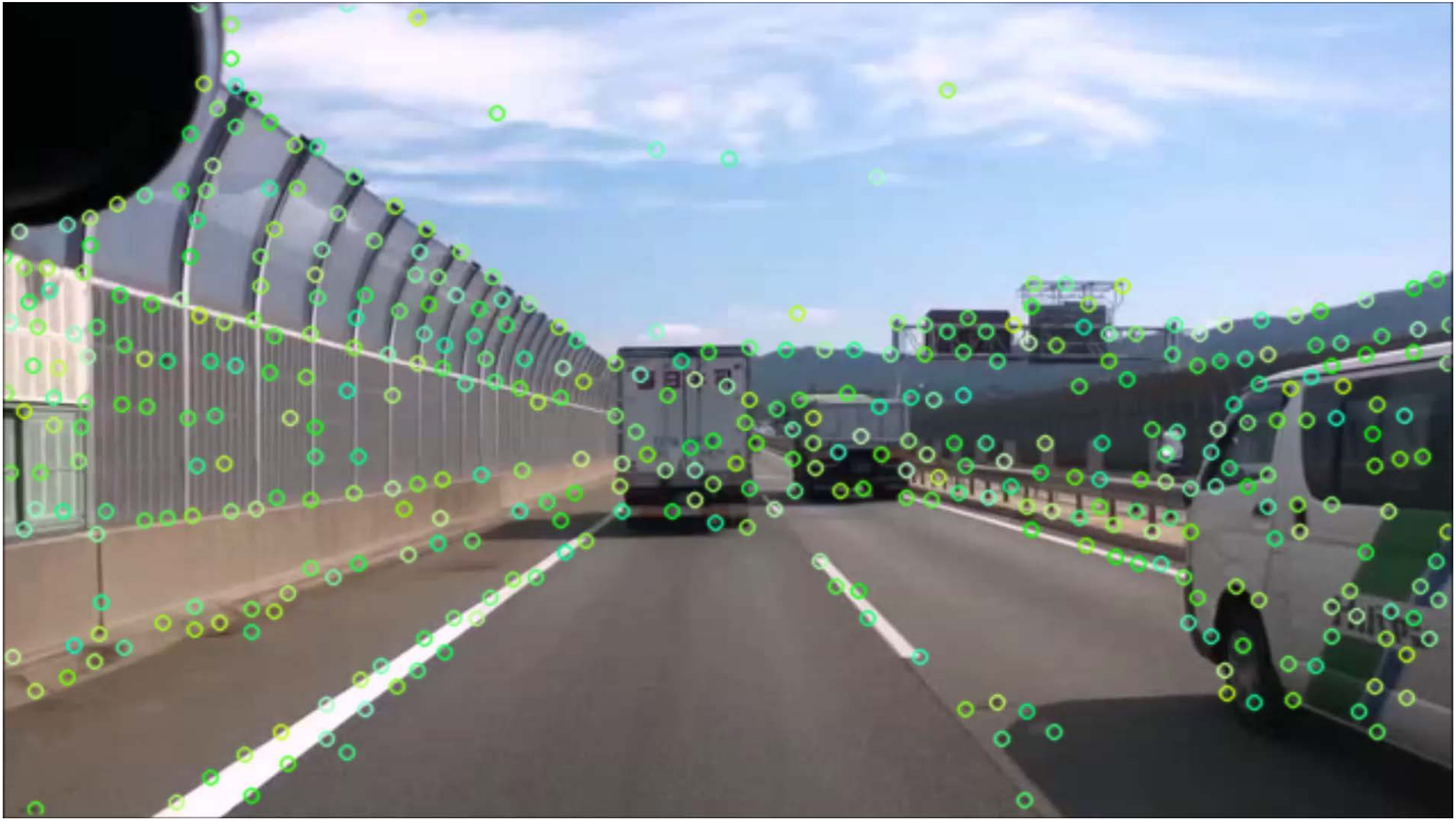
Local and Global Haar-Like Features



AGHaar: global and also adaptive and asymmetric

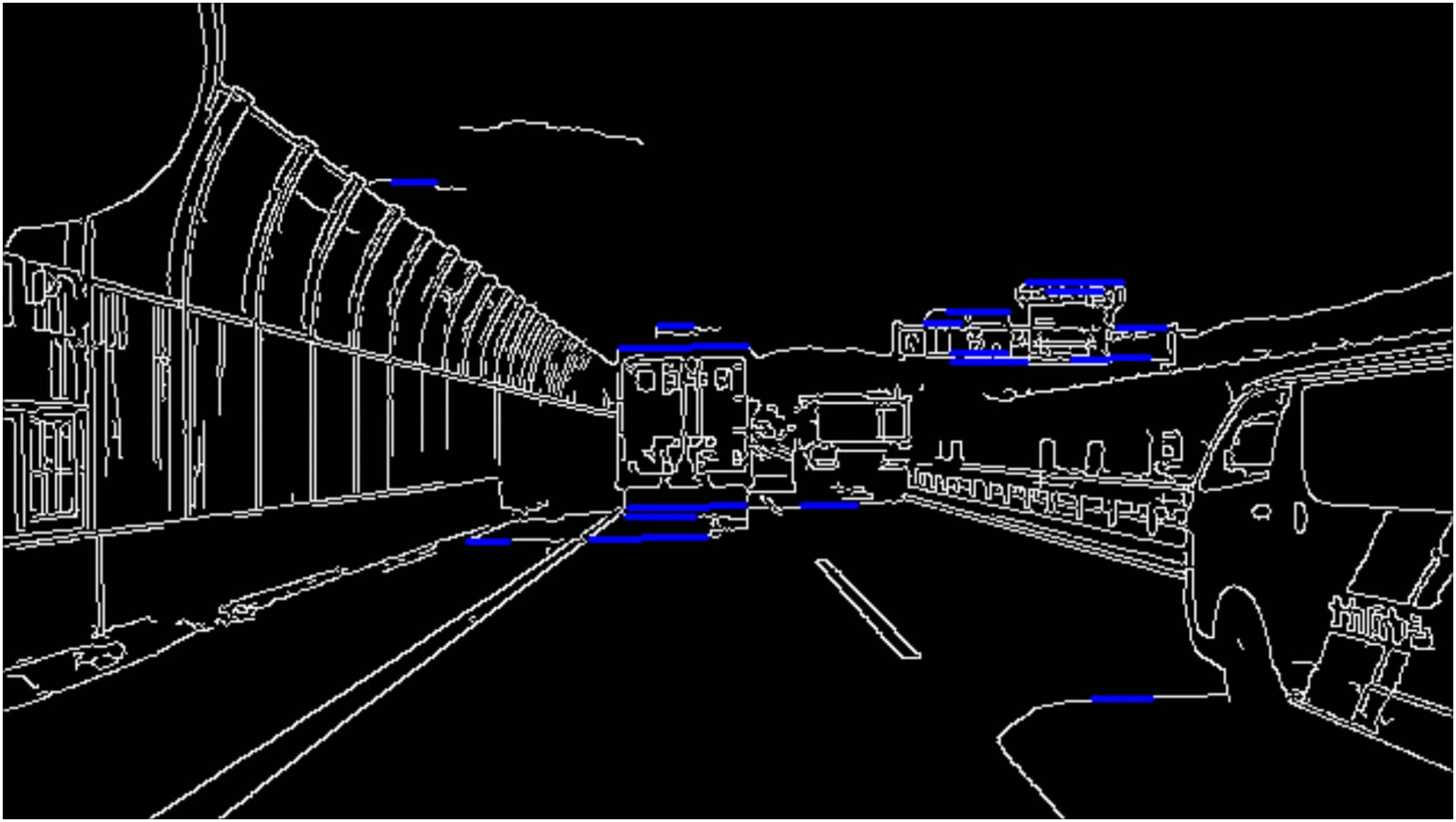
M. Rezaei and M. Terauchi, PSIVT, 2013

Density of Corners in AGHaar Regions



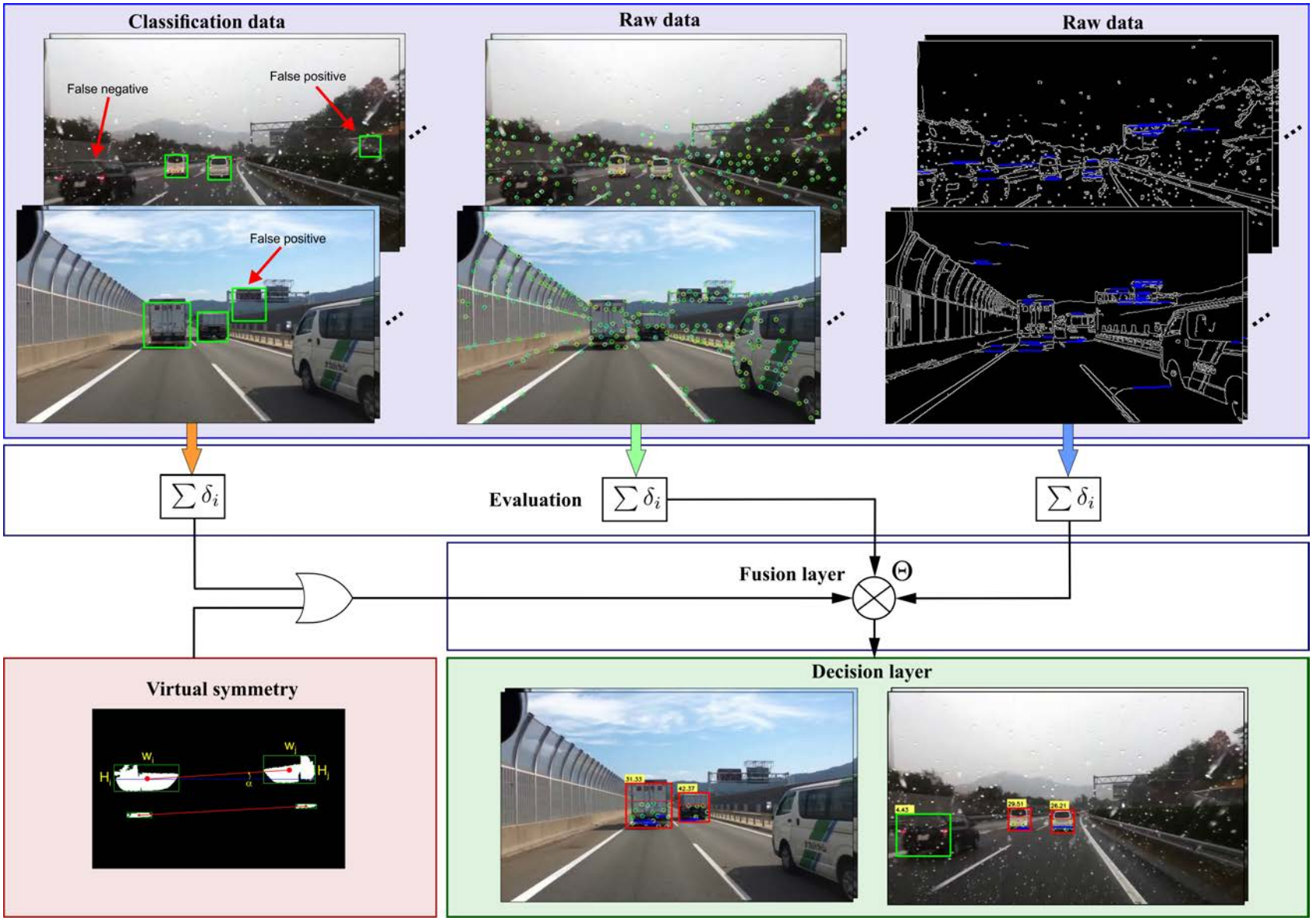
Harris corner detector

Horizontal Line Segments in AGHaar Regions



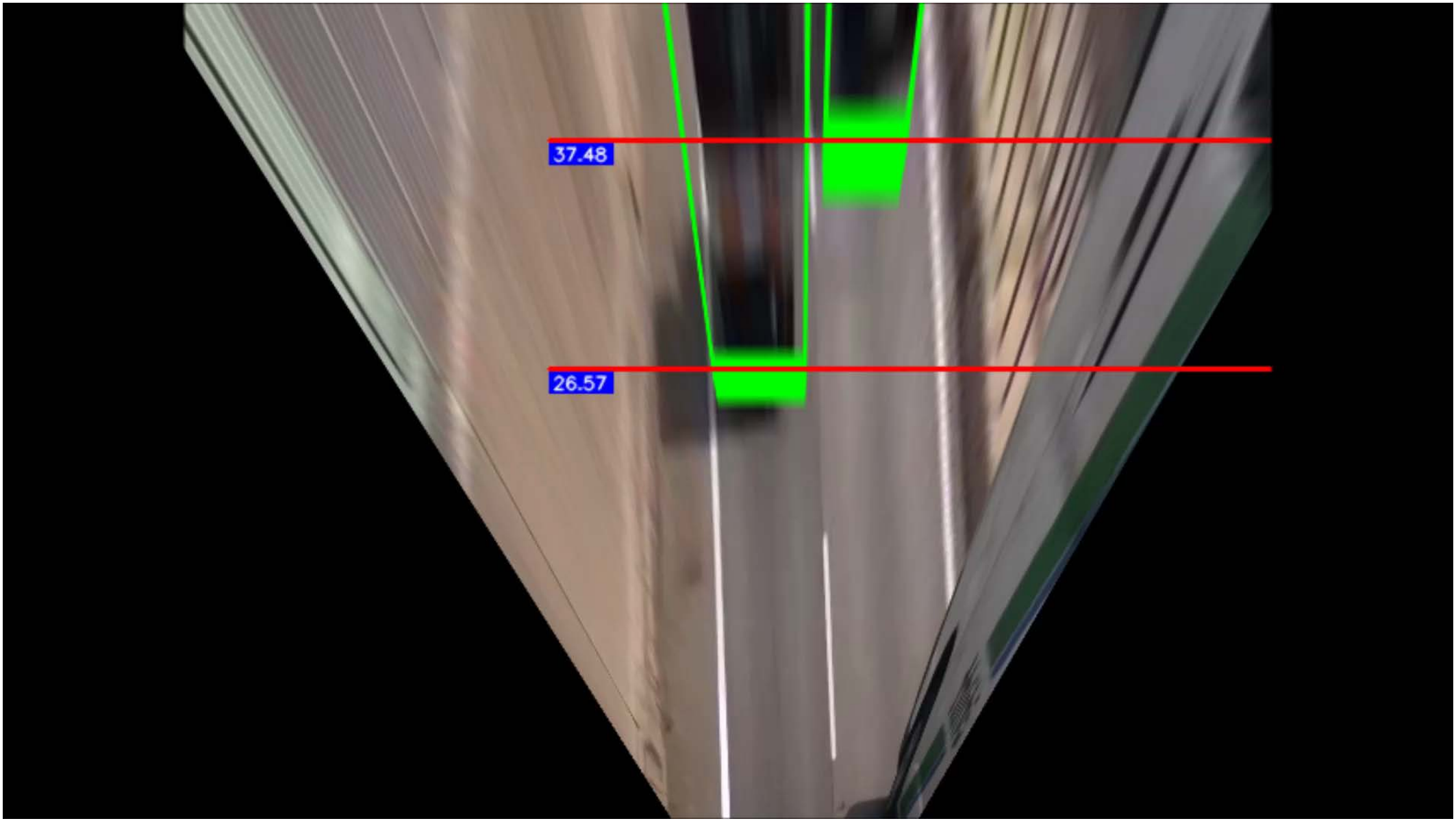
Progressive probabilistic Hough transform

Combine Hypotheses Using Dempster-Shafer

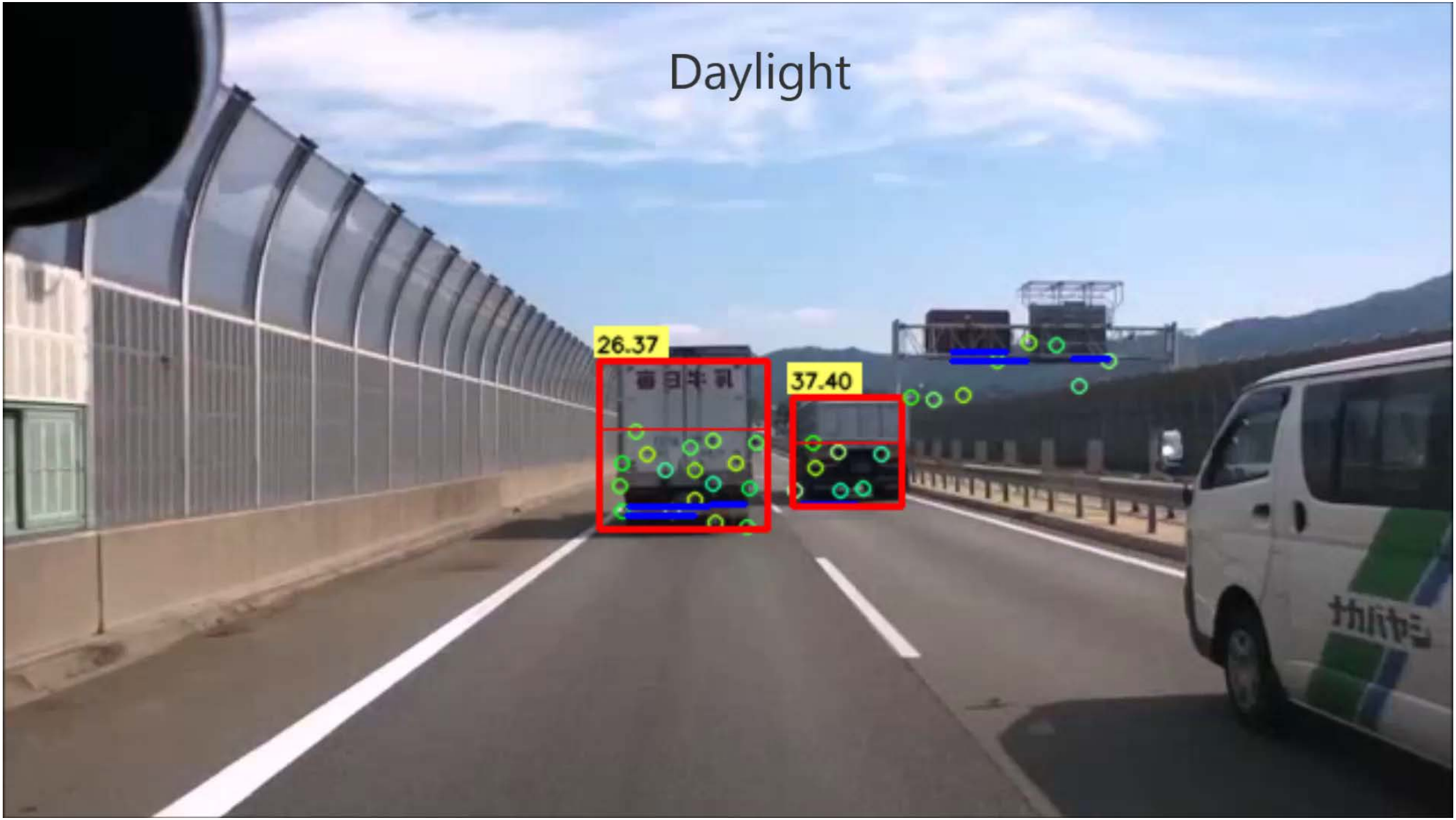


M. Rezaei, M. Terauchi, and R. Klette, submitted, 2014

Distance Estimation in Bird’s-eye View



Final Result



4: VbDA for Scene Analysis

Lane Analysis

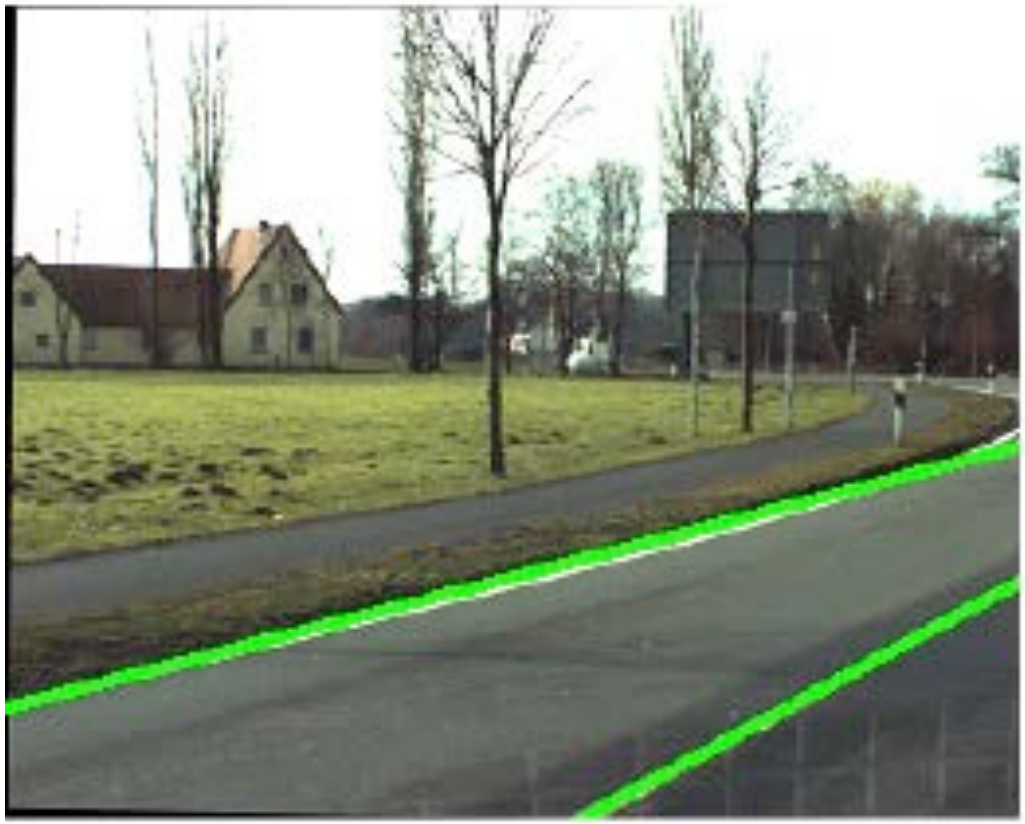
Wrong Lane Detection

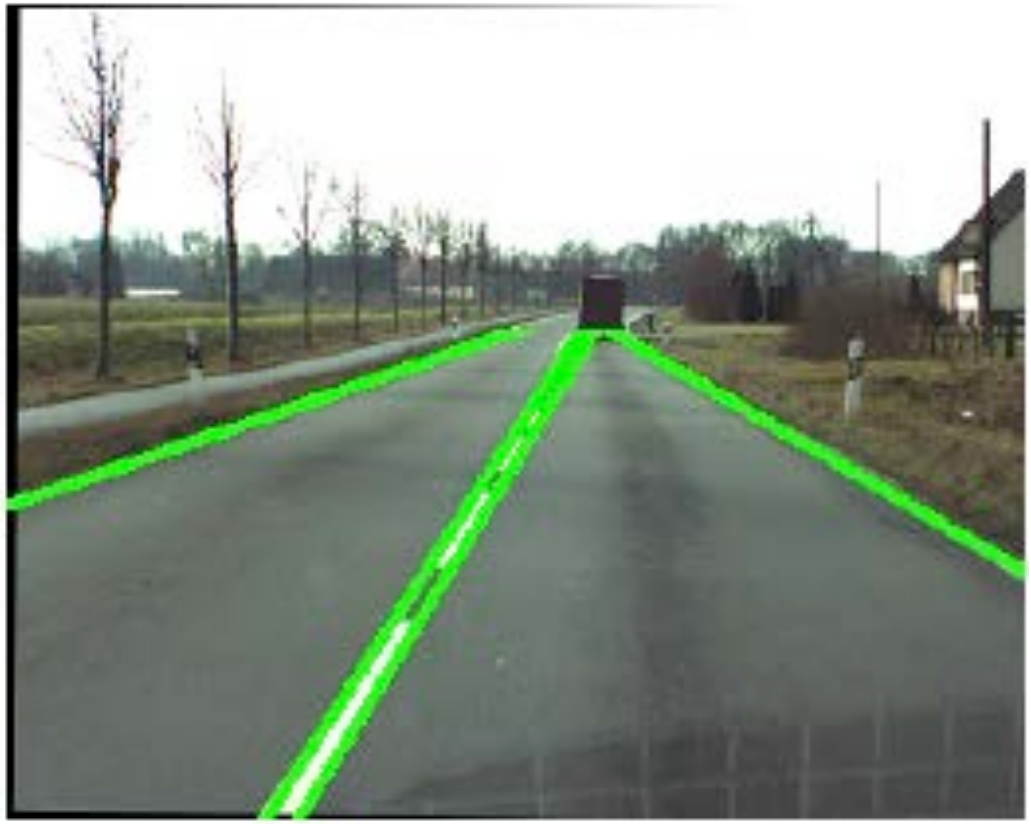
Road Signs (Traffic Signs, On-Road Writing)

Road Envirc
- Road Mar
- 3D Roads
- Free Spac
- Curb Dete

Difficult if only monocular vision

Monocular Lane-border Detection

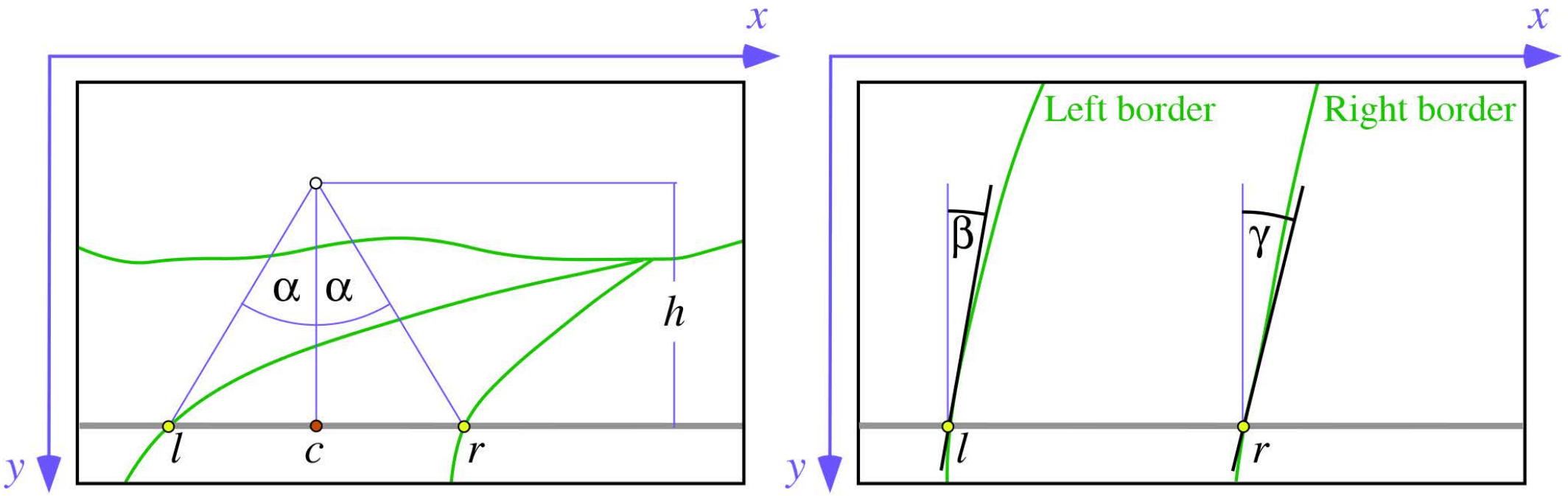




Hough transform very accurate for well-marked lanes

Z. Xu and B.-S. Shin, PSIVT, 2013

Lane Model for ``Difficult'' Scenarios



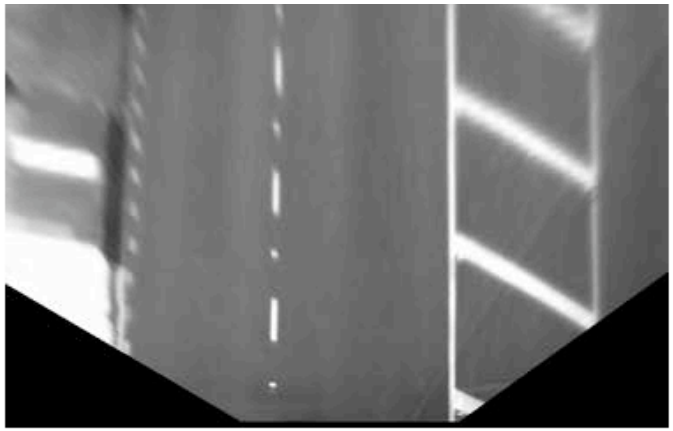
Y. Zhou, R. Xu, X. Hu, and Q. Ye, Measurement Science Technology, 2006

Particle-filter-based solution for this model, see

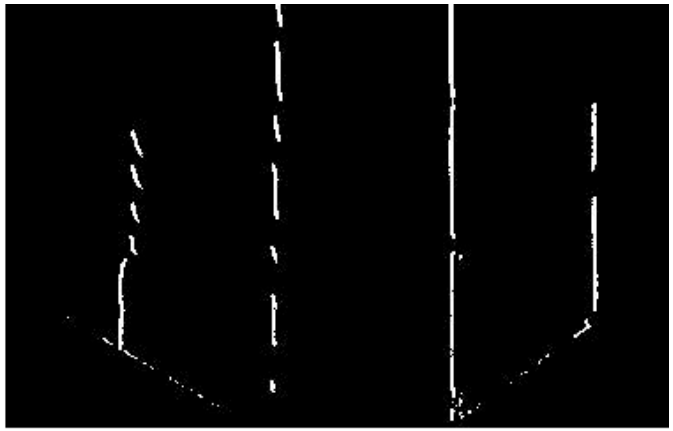
S. Sehestedt et al., *European Conf. Mobile Robots*, 2007

Workflow of Lane Detection

(a)

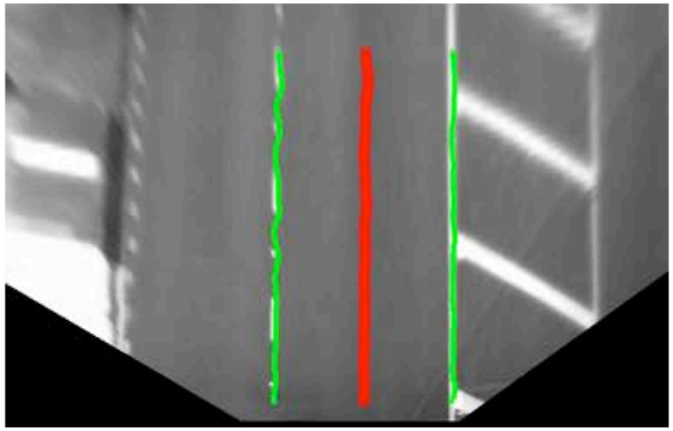


(b)

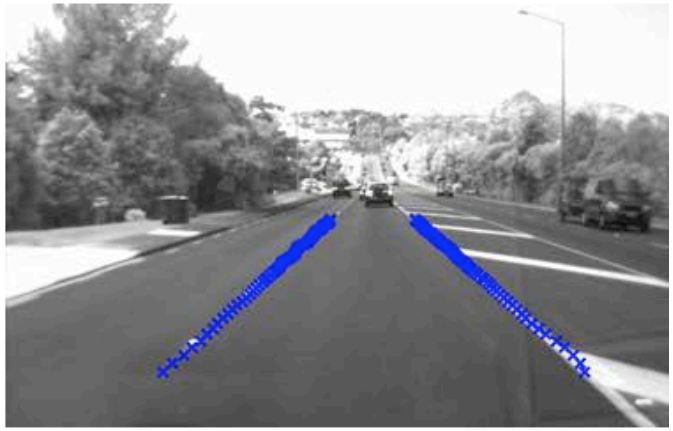


(c)

(d)



(e)

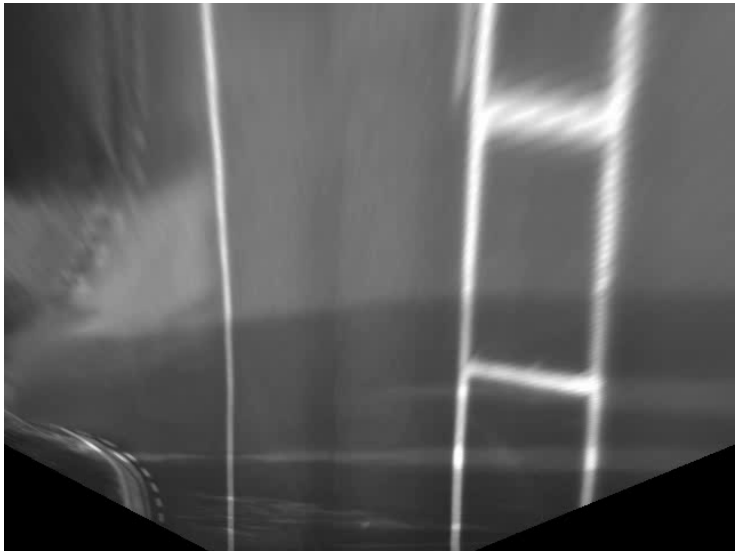


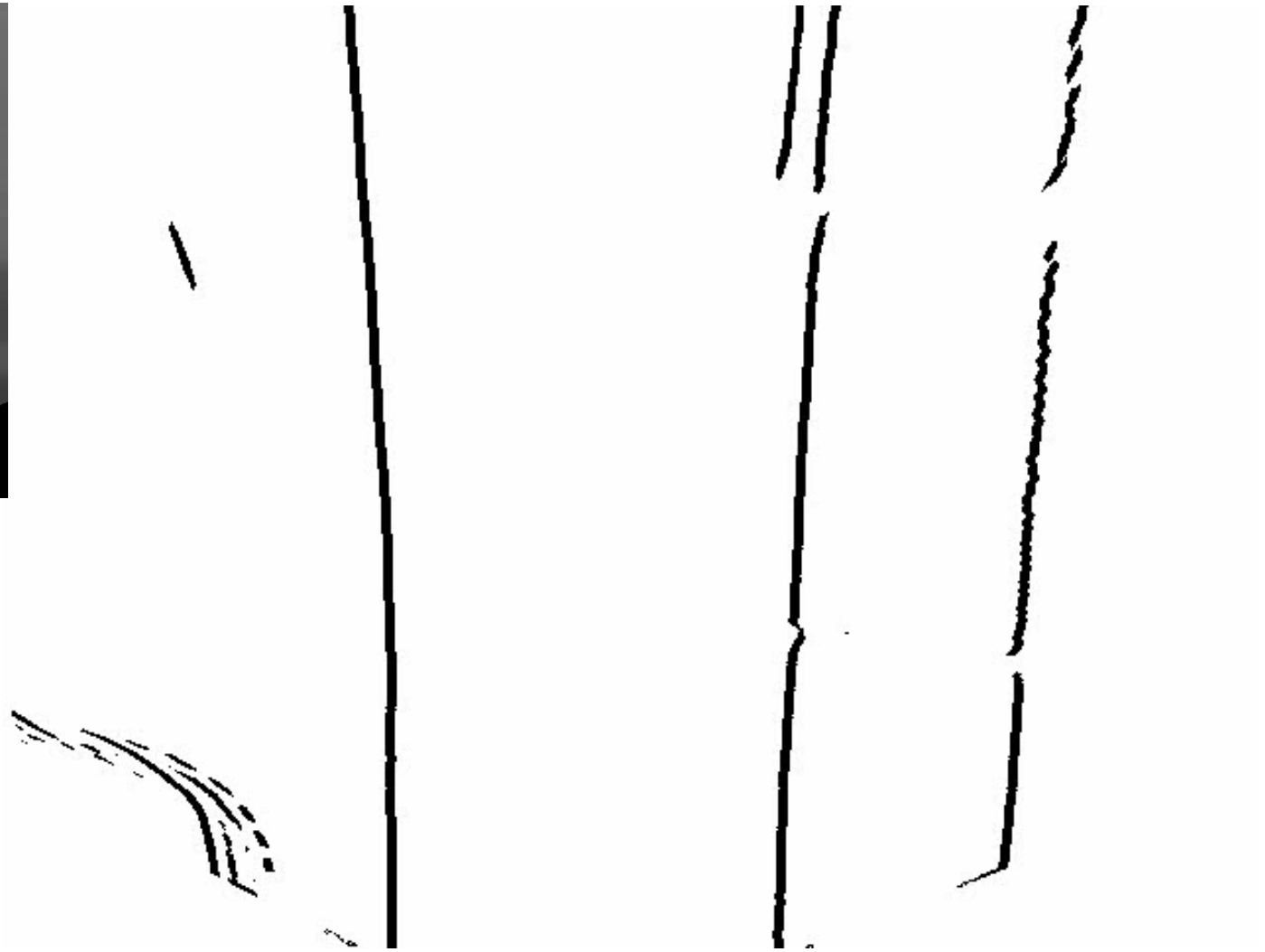
(f)

Bird's-eye view (b), vertical edges (c), EDT row component (d), and a particle filter for each processed image row

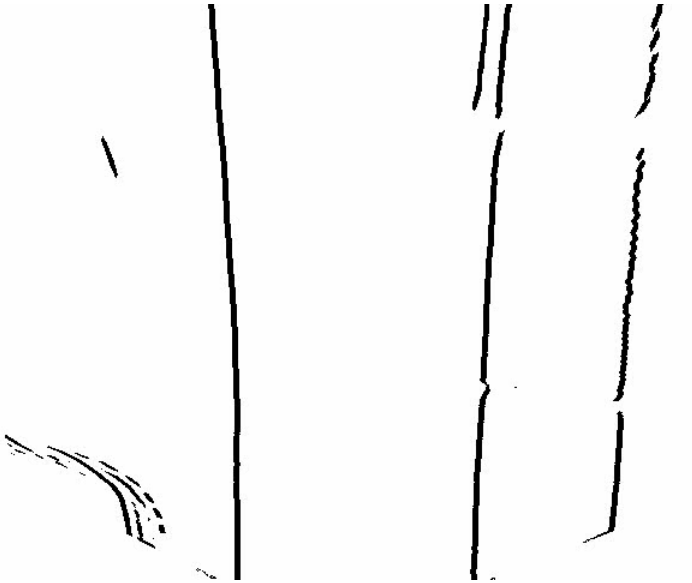
R. Jiang et al., *CAIP*, 2009

Vertical Edge Detection





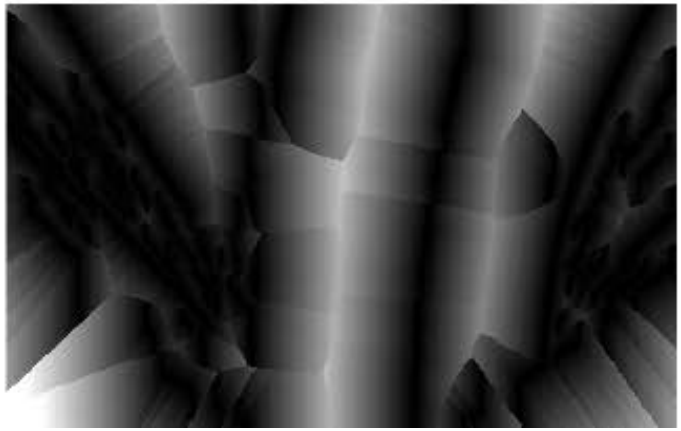
Row Component of Euclidean Distance Transform



Situations: No Border, no Marks on the Left



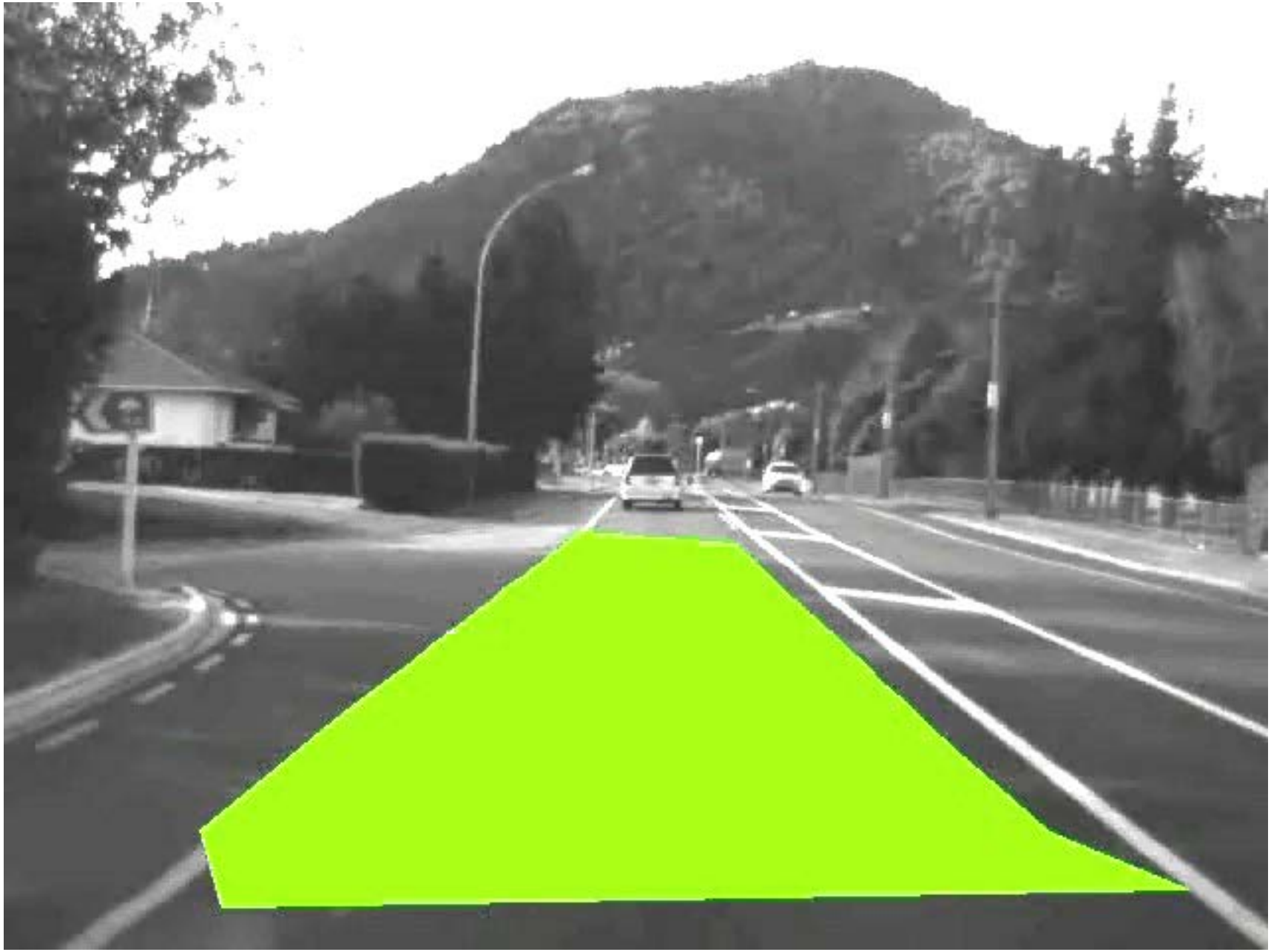




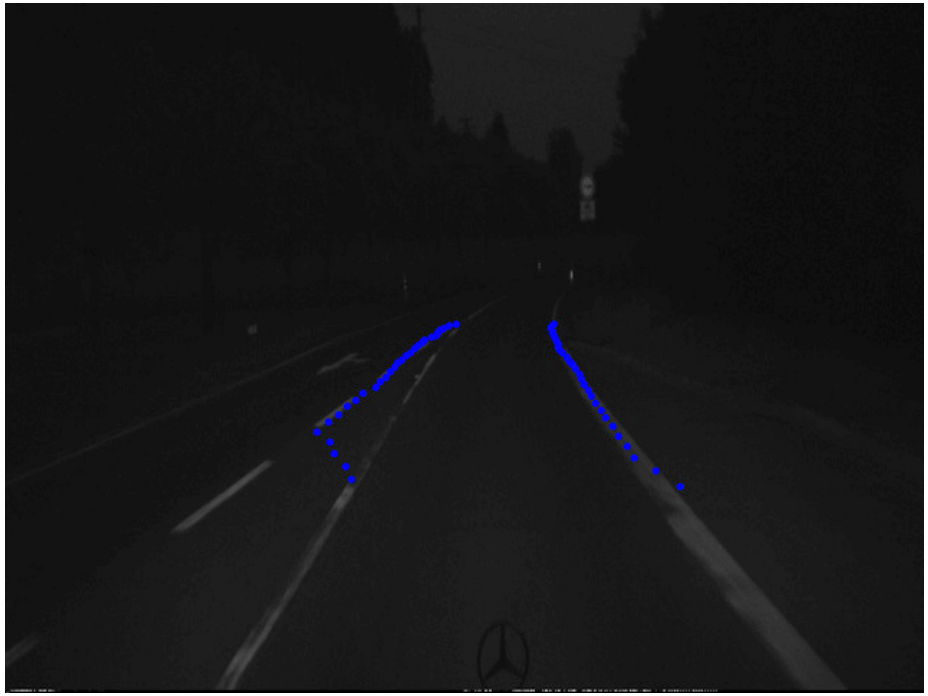

Lane Tracking (25 fps)

Predicted Space (Corridor)

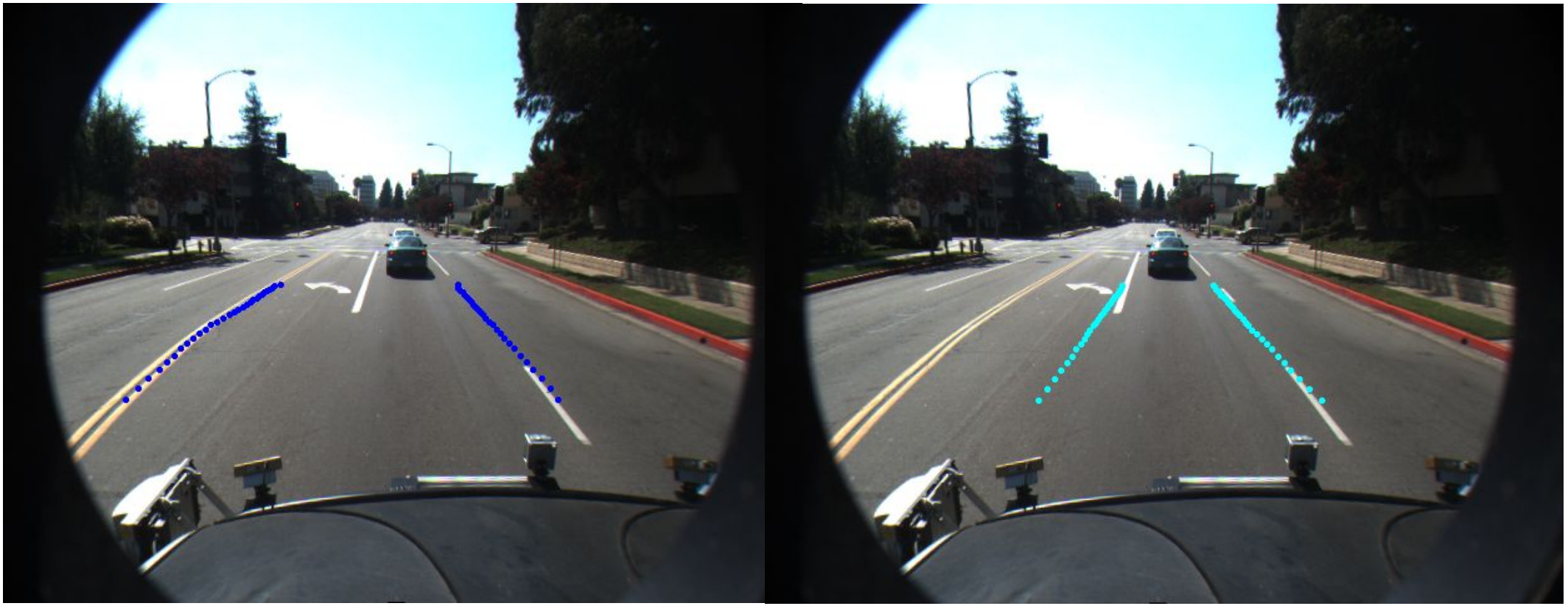
(the car will drive in the next ≈ 2-3 seconds)



Challenging Lane Border Situations



Superparticle Solution for Challenging Situations

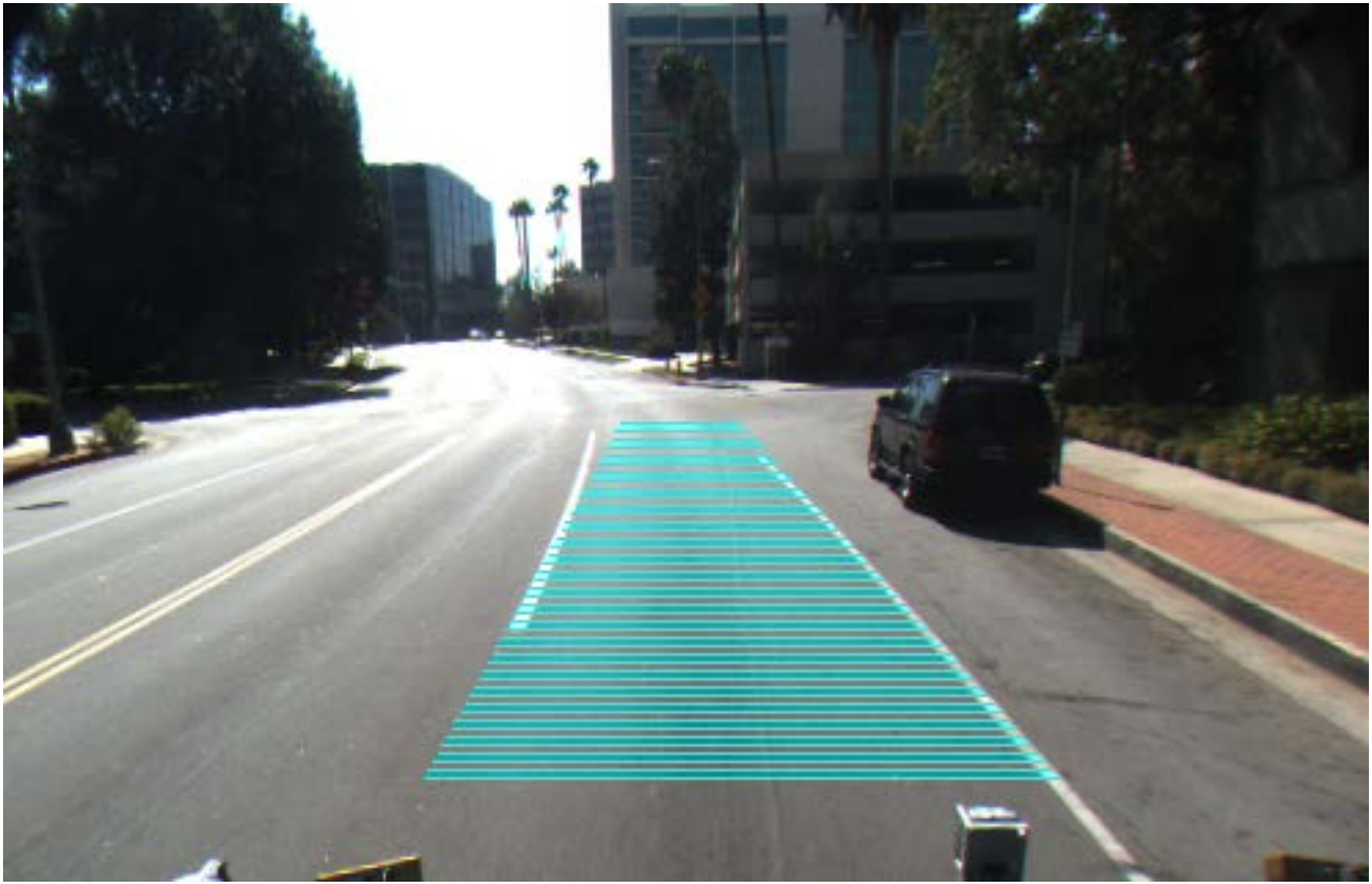


Previous approach

Superparticle

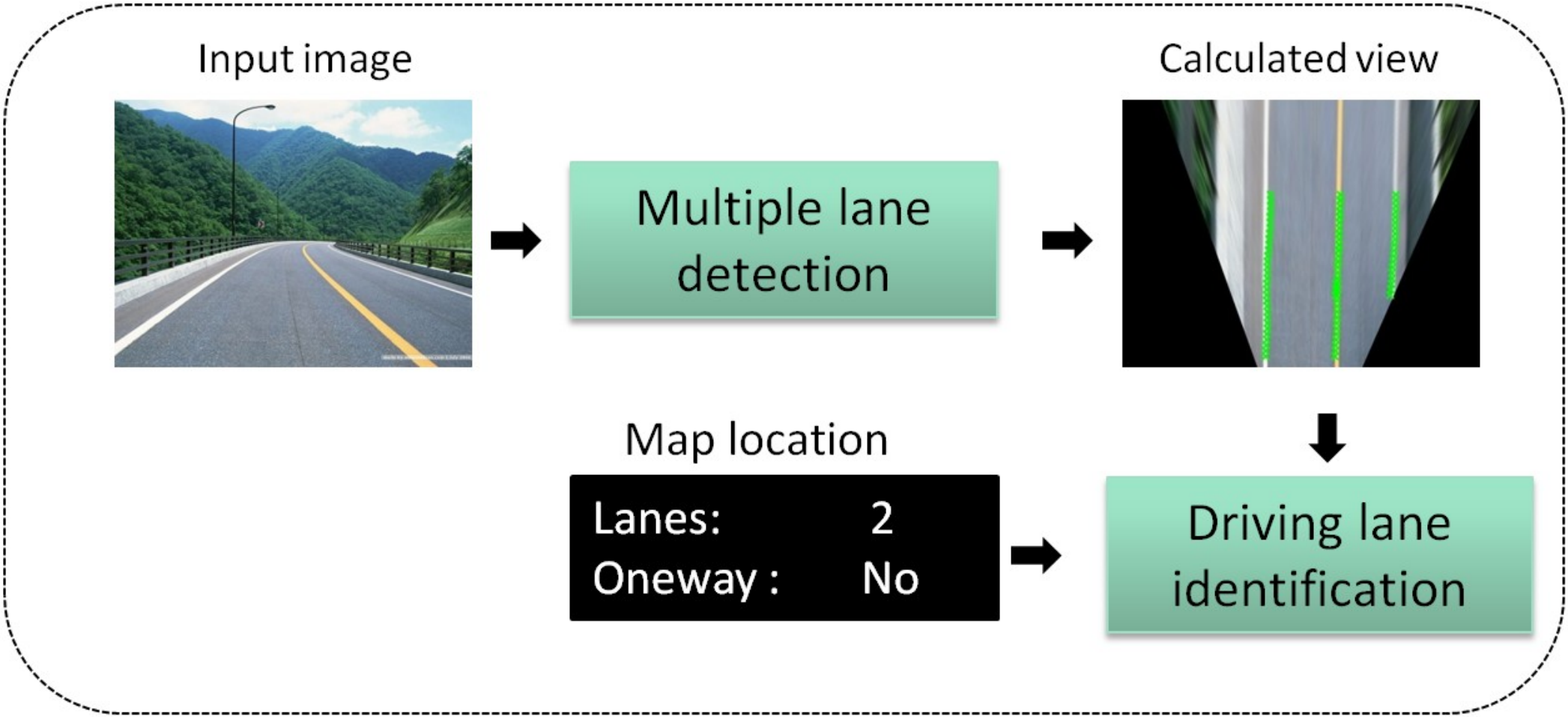
B.-S. Shin, J. Tao, and R. Klette, *Pattern Recognition,* 2014

Superparticle Solution Behaves Like Corridor

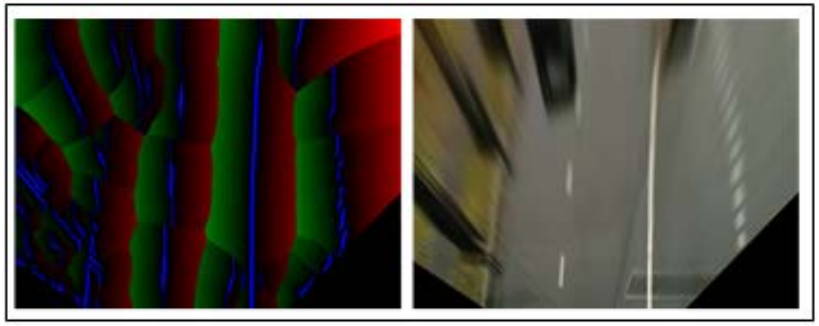


Monocular Wrong-lane Detection

Left-side or right-side driving on the road?

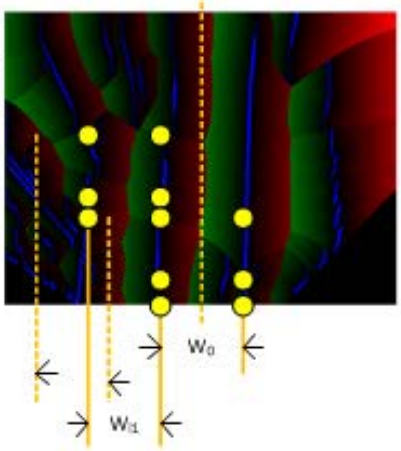


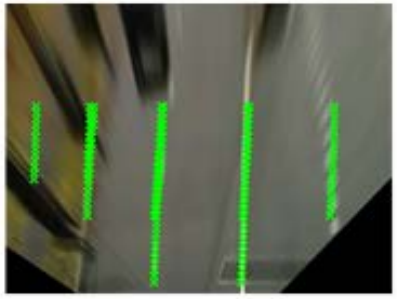
Lane Detection + Map Information



Preprocess: Inverse perspective map & Oriented distance transform

Multiple lane detection: particle filter based multiple lane detection (**Algorithm 1**)

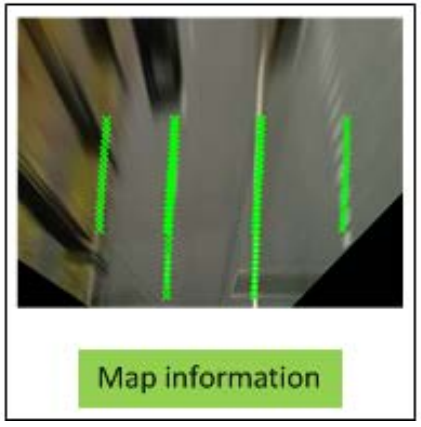




Information analysis: Conclude with detected lanes and map information (**Algorithm 3**)

Wrong lane?

Confidence measure: Measure confidence for each lane & eliminate false lane (**Algorithm 2**)



J. Tao, B.-S. Shin, and R. Klette, CAIP, 2013

5: VbDA for Driver and Road, & the Future

Driver Awa

Inter-Car C

Difficult if only monocular vision

Autonomou

...

Monocular Driver Awareness vs. Traffic

Monocular OUT-monitoring

Detect vehicles in front of ego-vehicle
Determine distance and angle to detected vehicles
at each time slot *t*

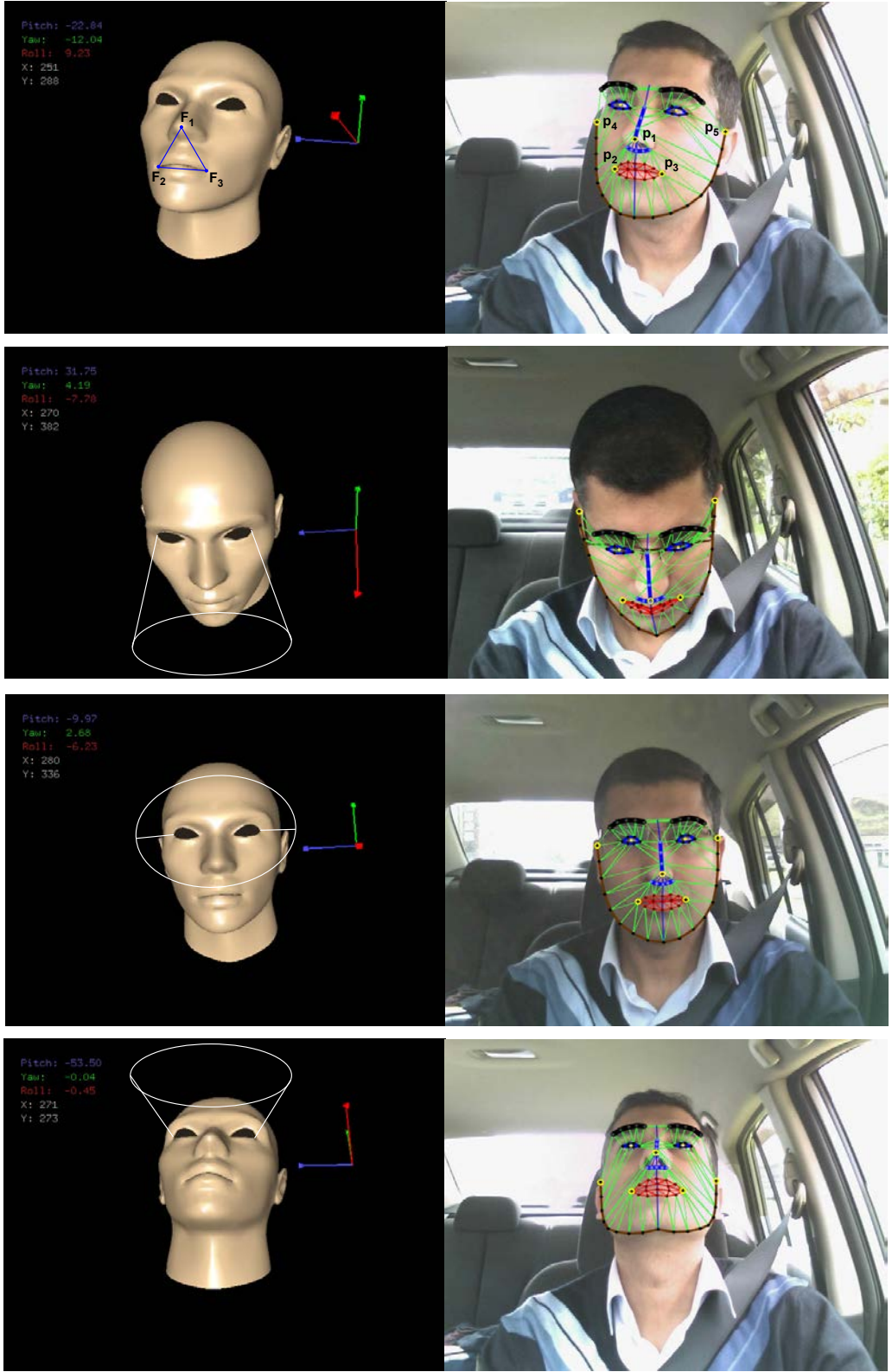
Monocular IN-monitoring

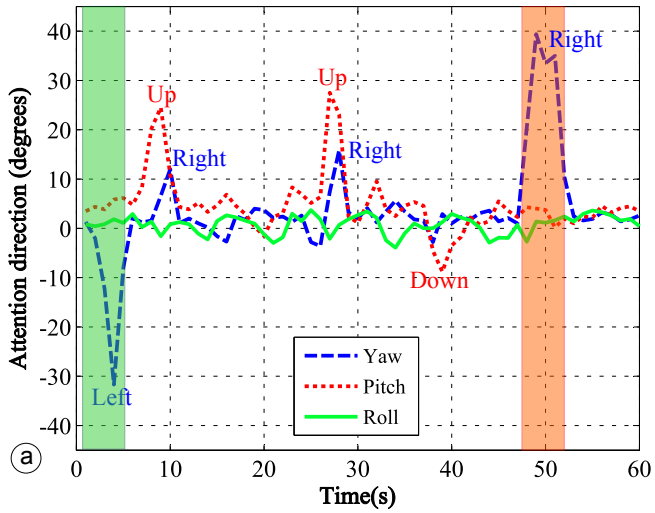
Specify head pose and detect eye gaze
using 2D face and eye detection and 3D head pose

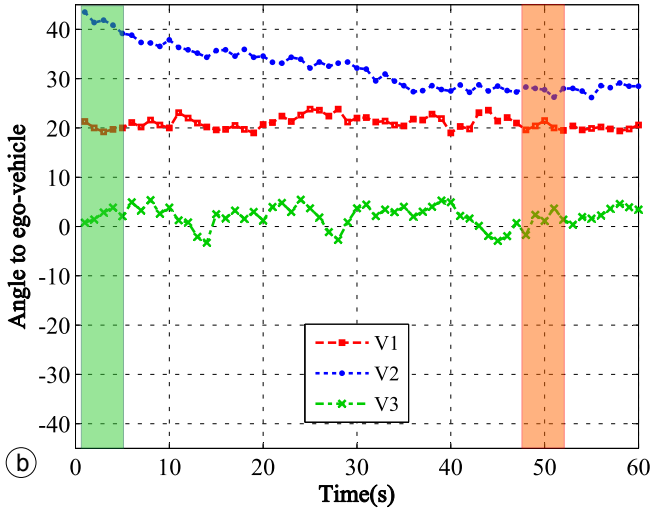
Combine both findings into one awareness model

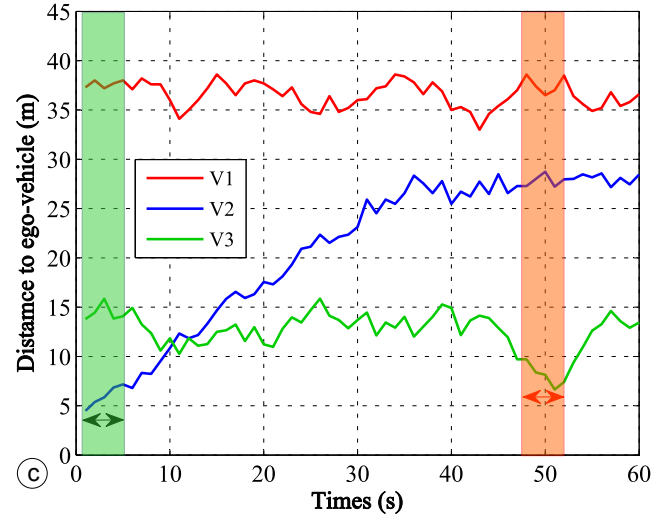
Fuzzy evaluation in a 7- (or more) dimensional
feature space

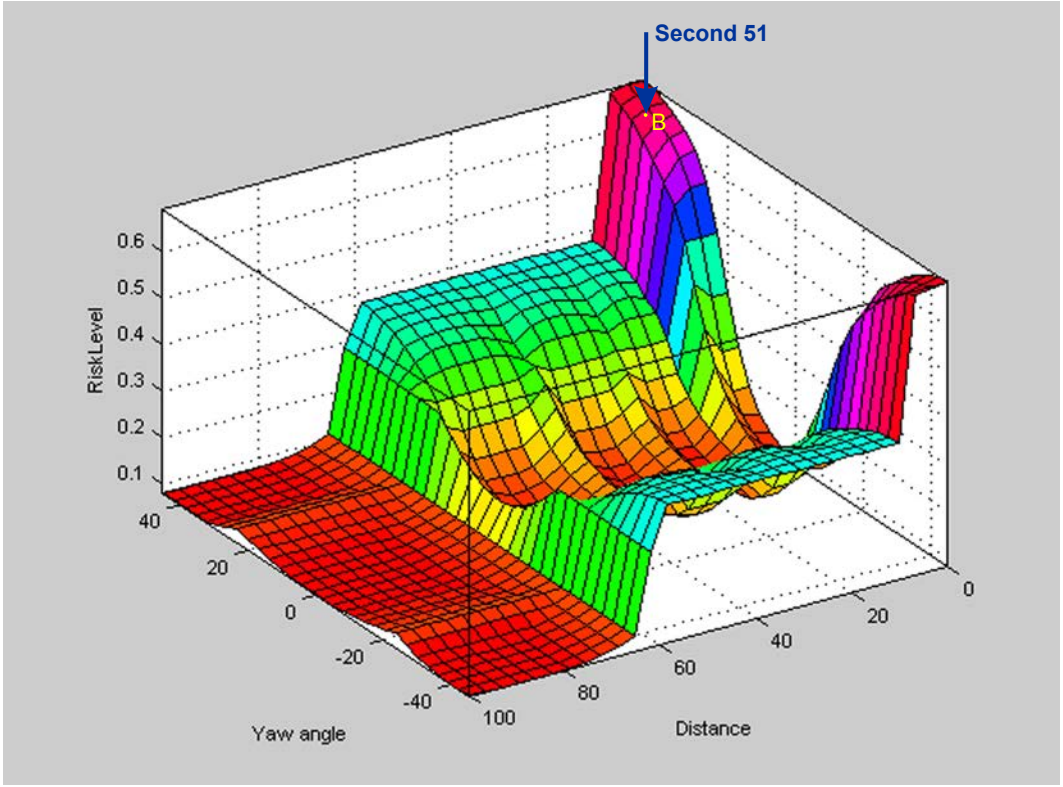
M. Rezaei and R. Klette, CVPR, 2014

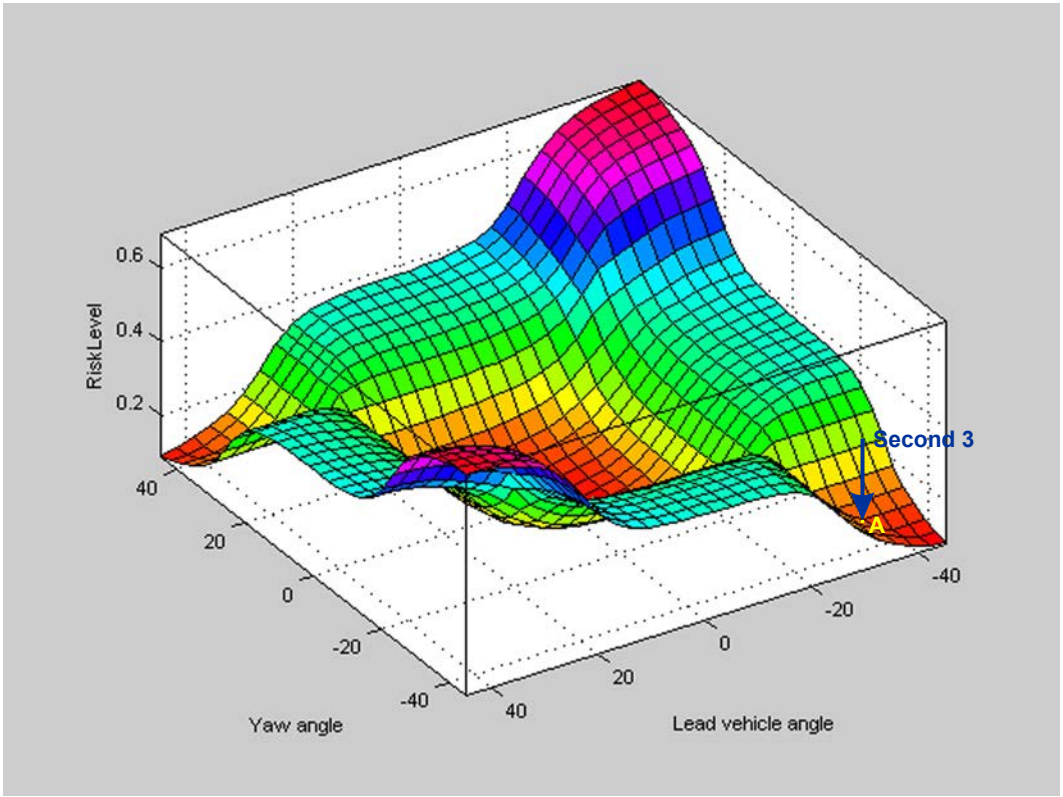

Pitch: -22.84
Yaw: -12.04
Roll: 9.23
X: 251
Y: 288
F1
F2
F3
p4
p1
p5
p2
p3
Pitch: 31.75
Yaw: 4.19
Roll: -7.78
X: 270
Y: 382
Pitch: -9.97
Yaw: 2.68
Roll: -6.23
X: 280
Y: 336
Pitch: -53.50
Yaw: -0.04
Roll: -0.45
X: 271
Y: 273


Attention direction (degrees)
Time(s)
Up
Right
Up
Right
Right
Down
Left
Yaw
Pitch
Roll
a


Angle to ego-vehicle
Time(s)
V1
V2
V3
b


Distamce to ego-vehicle (m)
Times (s)
V1
V2
V3
c





7- (or more) dimensional feature space using a fuzzy evaluation system for identifying critical issues

M. Rezaei and R. Klette, CVPR, 2014

Conclusions for Monocular VbDA

Monocular camera data are able to provide solutions for a fairly wide spectrum of tasks

The talk illustrated a few monocular solutions; wide-angle lenses could contribute to new opportunities

The use of more sensors (stereo vision, radar, laser range-finder, ...) provides

more opportunities, e.g. for

- 3D roadside modeling,
- pedestrian detection,
- ...

PSIVT | 7th PACIFIC-RIM SYMPOSIUM ON IMAGE AND VIDEO TECHNOLOGY

23-27, November, 2015. Auckland, New Zealand

Call for Papers

The **seventh Pacific-Rim Symposium on Image and Video Technology (PSIVT 2015)** will be held in Auckland, the beautiful "City-of-Sails" in New Zealand, from November 23rd to November 27th, 2015.

PSIVT provides a forum for researchers to present the latest research and developments in image and video technology. Both academic and industrial research contributions are encouraged to stimulate discussion for the mutual benefit of both. Researchers, artists, developers, educators, performers, and practitioners of image and video technology from the Pacific Rim and around the world are invited to participate.

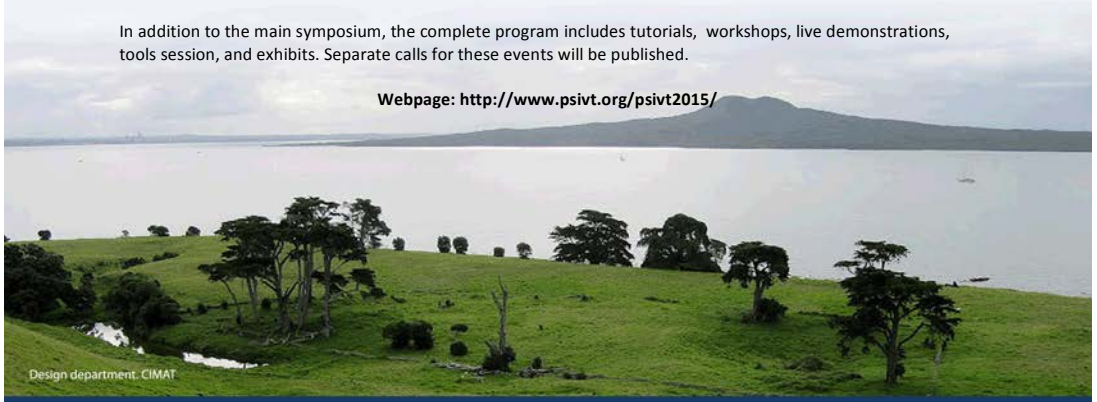
High-quality contributions from prospective authors are solicited in all aspects of image and video technology including, but not limited to:

- Imaging and Graphics Hardware and Visualization
- Image/Video Coding and Transmission
- Image/Video Processing and Analysis
- Image/Video Retrieval and Scene Understanding
- Applications of Image and Video Technology
- Biomedical Image Processing and Analysis
- Computational Photography and Arts
- Computer and Robot Vision
- Pattern Recognition
- Video Surveillance
- Automotive Vision

The **PSIVT 2015** proceedings are planned to be published by Springer-Verlag in the Lecture-Notes-in-Computer-Science series.
All papers will be subject to a double-blind review process which will be carried out by at least two experienced researchers in the field, supervised by area chairs, and including a brief rebuttal period. As a guide: Previous issues of PSIVT all had an acceptance rate of about 40 percent.

In addition to the main symposium, the complete program includes tutorials, workshops, live demonstrations, tools session, and exhibits. Separate calls for these events will be published.

Webpage: http://www.psivt.org/psivt2015/



IAPR

Lecture Notes in Computer Science
LNCS LNAI LNBI

This book discusses some of the presented subjects