AI 2008, Auckland

Stereo-Vision-Based Driver Assistance

- Predict Situations and Automatically Adapt to Traffic -

Reinhard Klette

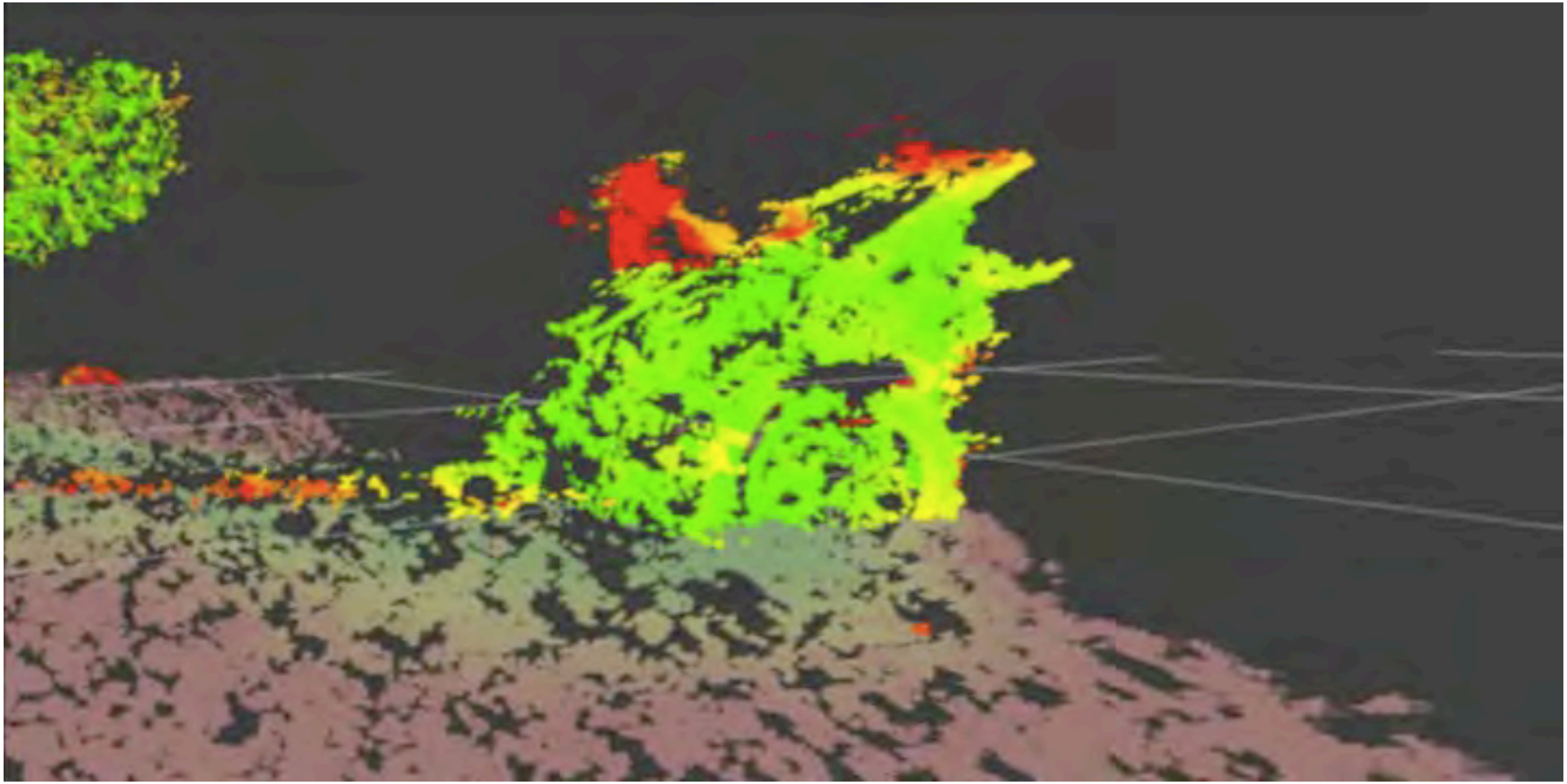
Zhifeng Liu, Shushi Guan, Tobi Vaudrey, Sandino Morales, Jorge Sanchez, and James Milburn

The University of Auckland, Tamaki campus

.*enpeda*.. Project

Abstract: This talk is about ...

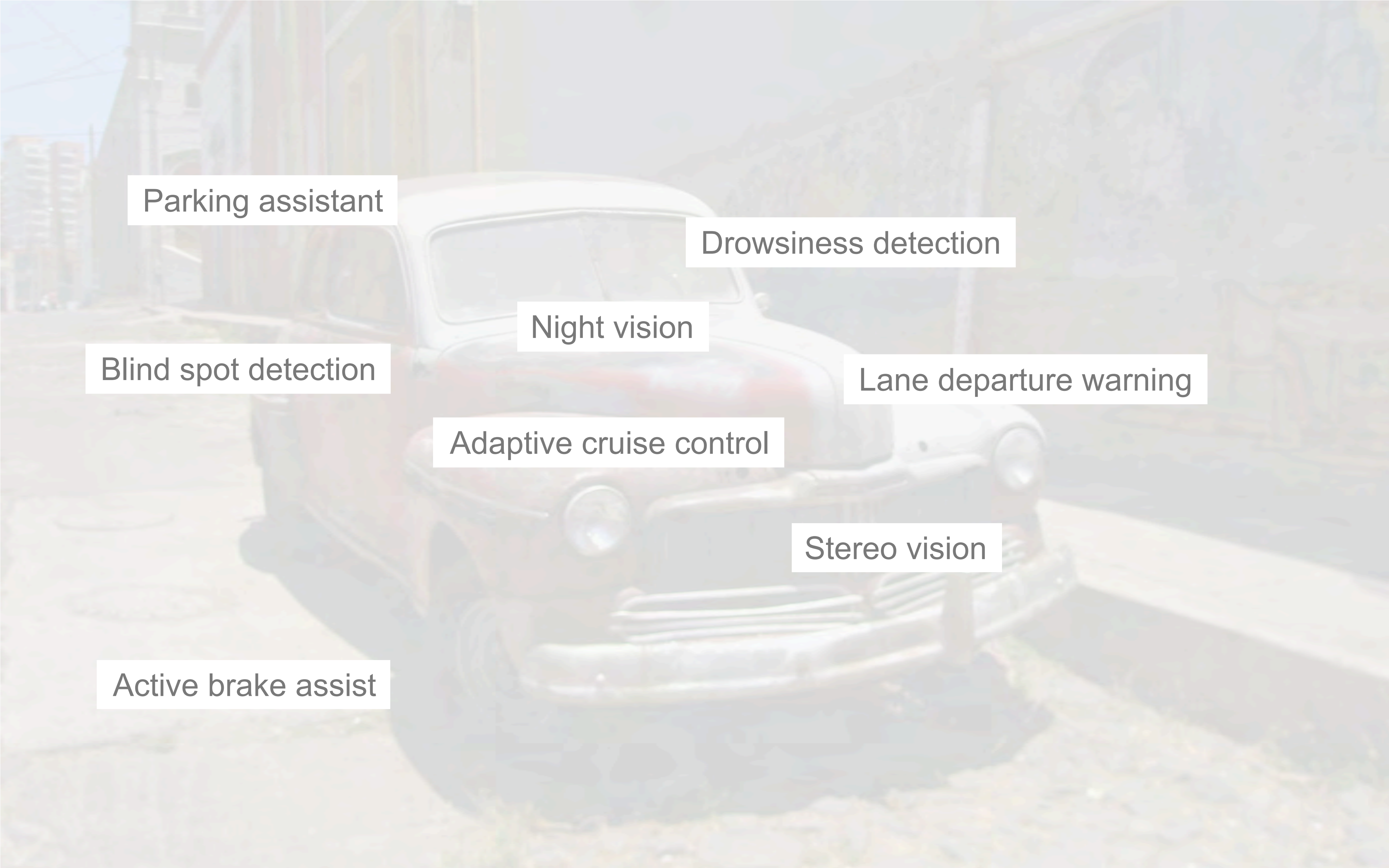
Stereo and Motion Analysis in DAS towards Analysis and Understanding of Traffic



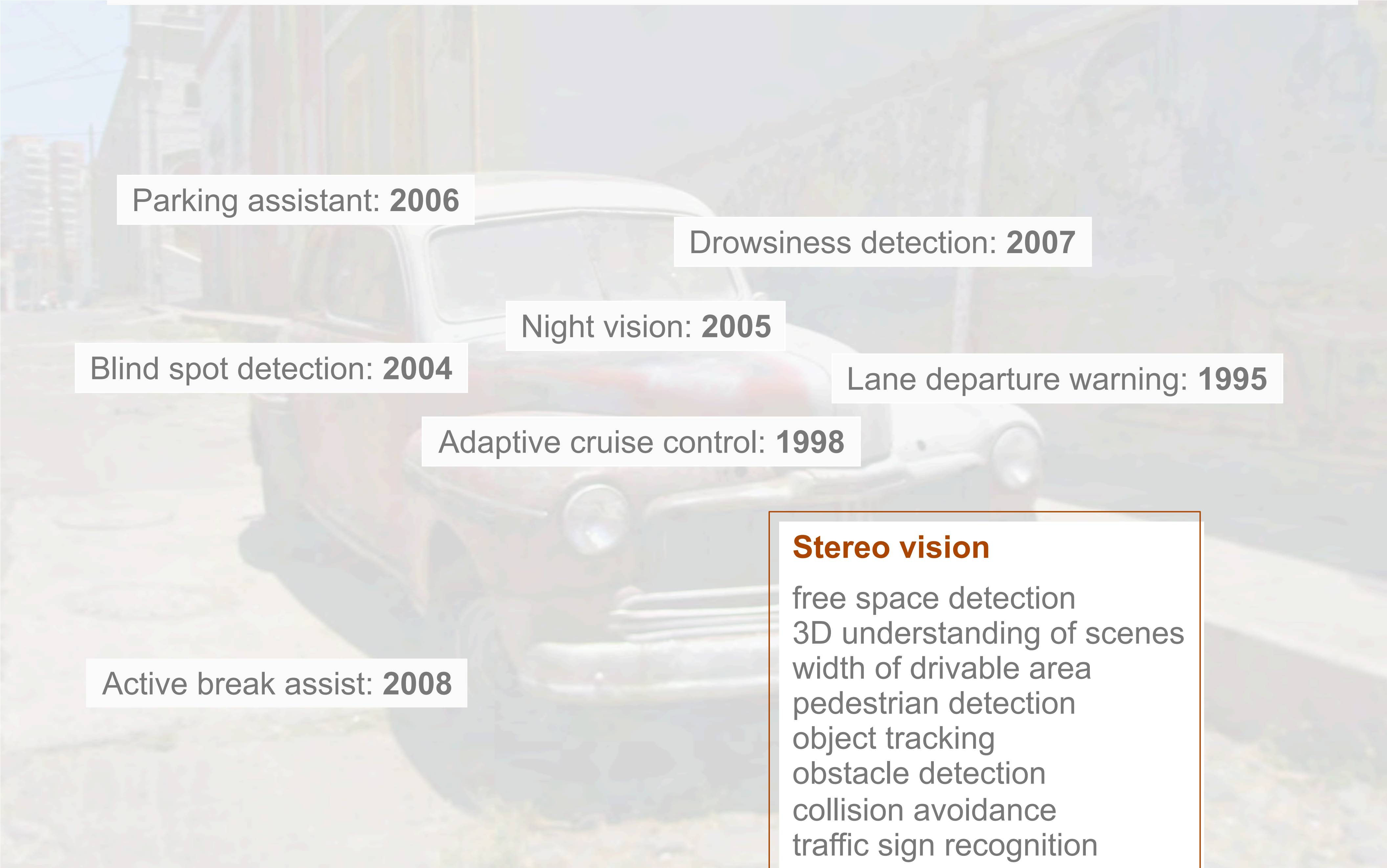
General motivation:

Visual Perception for Modern Cars

Standard sensors in the **car of the future**: radar, range-finder, cameras, ultra-sound



Applications in **cars today**



Parking assistant: **2006**

Drowsiness detection: **2007**

Night vision: **2005**

Blind spot detection: **2004**

Lane departure warning: **1995**

Adaptive cruise control: **1998**

Active break assist: **2008**

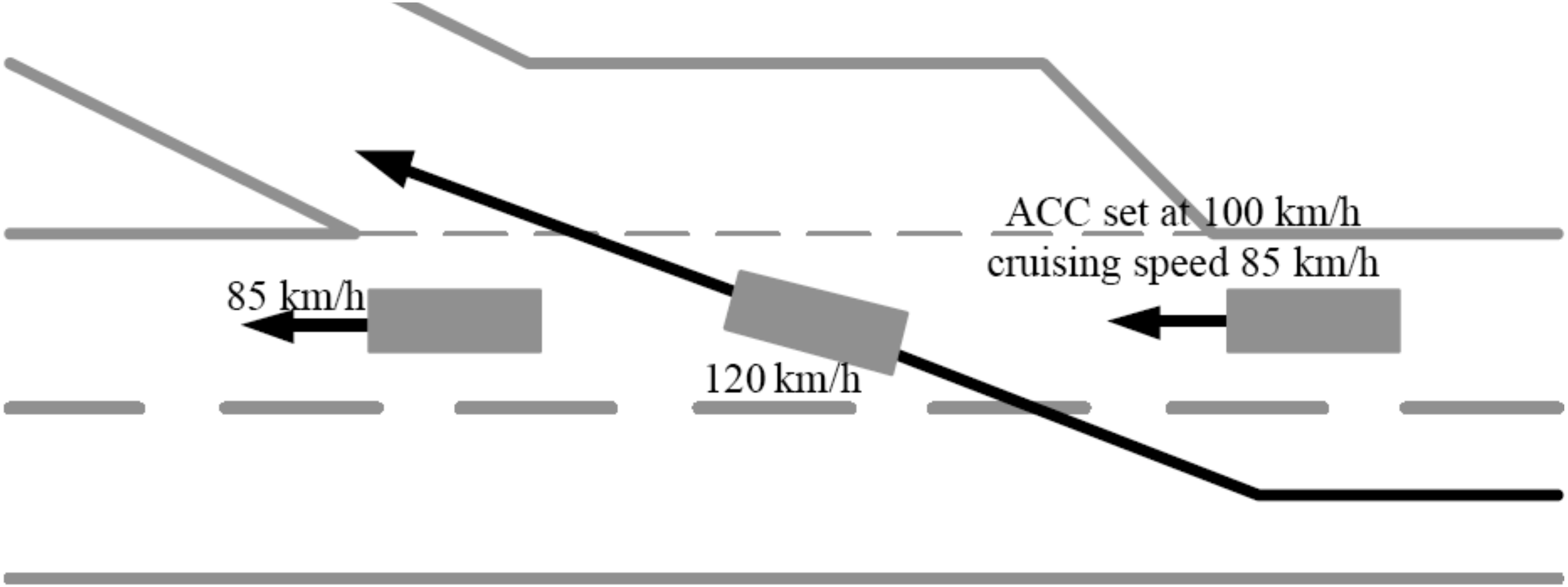
Stereo vision

free space detection
3D understanding of scenes
width of drivable area
pedestrian detection
object tracking
obstacle detection
collision avoidance
traffic sign recognition
...

" Very few innovations can be expected to have as high an impact on future cars as the capability to **perceive** the environment and to **plan** and **conduct** appropriate driving behavior. "

Ü. Özgüner, Ch. Stiller, K. Redmill, Proc. IEEE, 2007

Just one **example**: adaptive cruise control (since 1998)



H.M. Jagtman, E. Wiersma, 16th ICTCT workshop, 2003

ACC is distance based, thus there **should** be no problem. --- However, (stereo) vision might also help to exclude (any very low chance of) failure in such a case.

Visual perception influenced by

- day or night
- sun, shade, sun stroke
- rain, snow, hail
- heavy or light traffic
- urban or open road
- fast or slow ego-motion
- fast or slow motion of other objects
- etc.

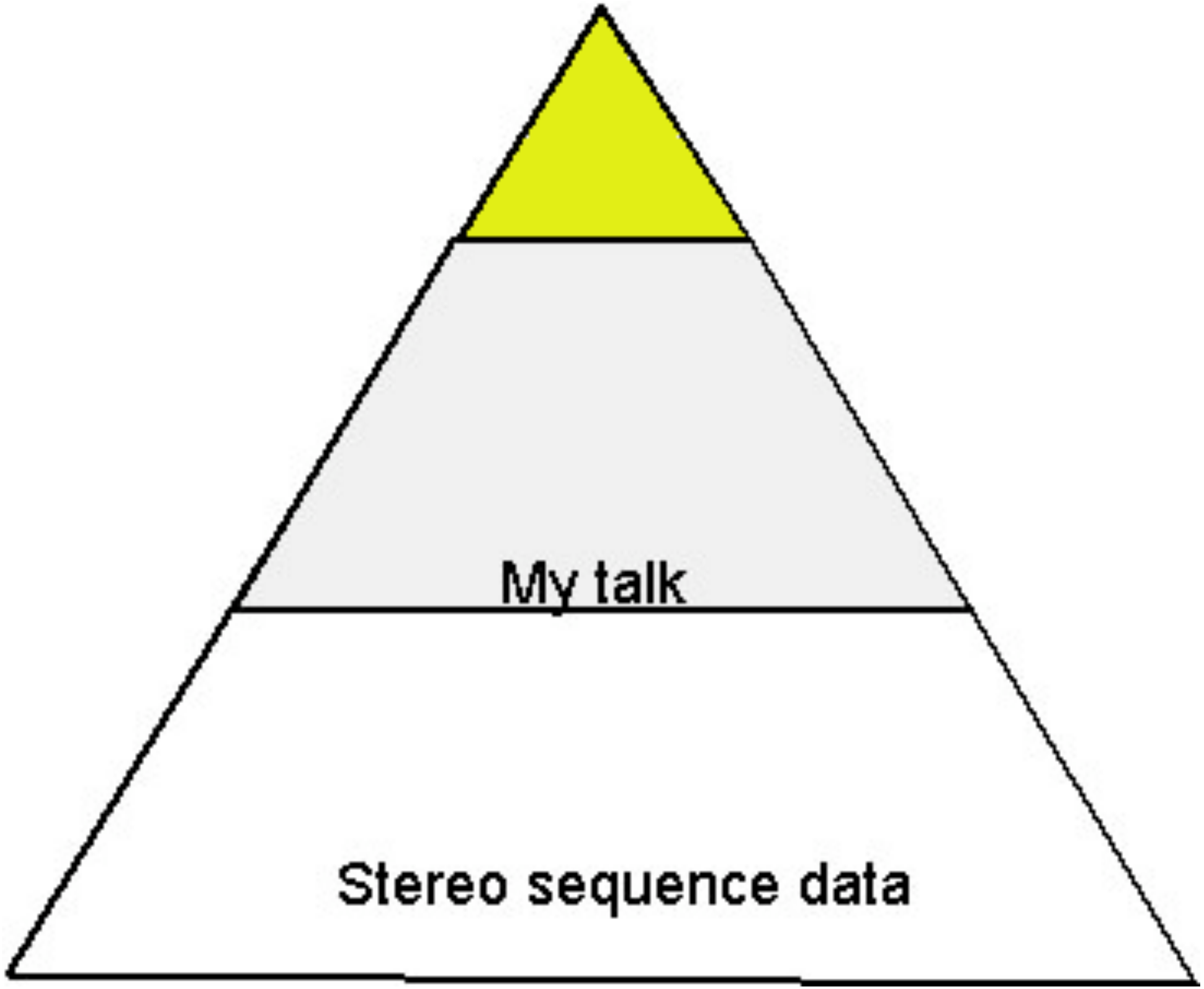
Performance of stereo vision sensors/programs is evaluated by using specific **metrics**: so far only by a few simple metrics to advice with respect to a basically unlimited diversity of possible situations.

Imagine a driver with allochromasia, tunnel vision (limited peripheral viewing), and myopia (distance blur), but stereo vision:

Did you see the bird flying through?

At what distance to the car?

The common "AI pyramid":



The *.enpeda..* Project

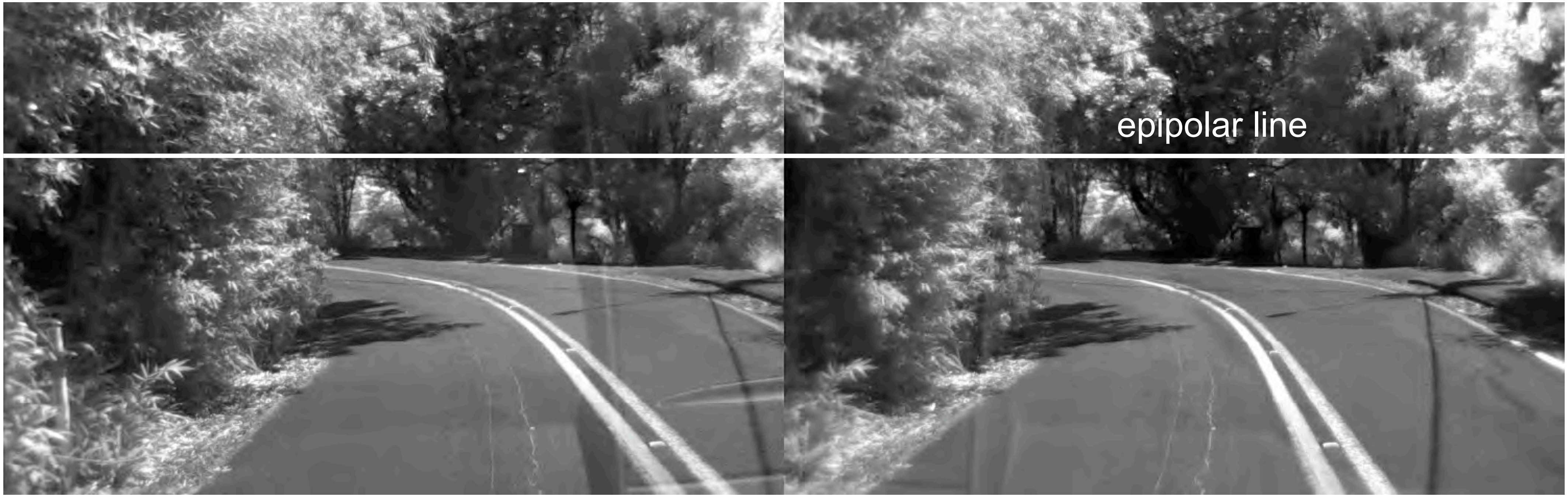
HAKA1
=
High
Awareness
Kinematic
Automobile
no. 1

Test vehicle for recording stereo sequences

two 752 x 480 gray level cameras, 10 bit per pixel

Sponsors: Mercedes Benz New Zealand & Coutts Cars North Shore

Computer installation, thanks to Manukau IT



HAKA1 in the Waitakeres: geometrically rectified stereo images

(see standard references
in computer vision)

with prior calibration
at Tamaki campus

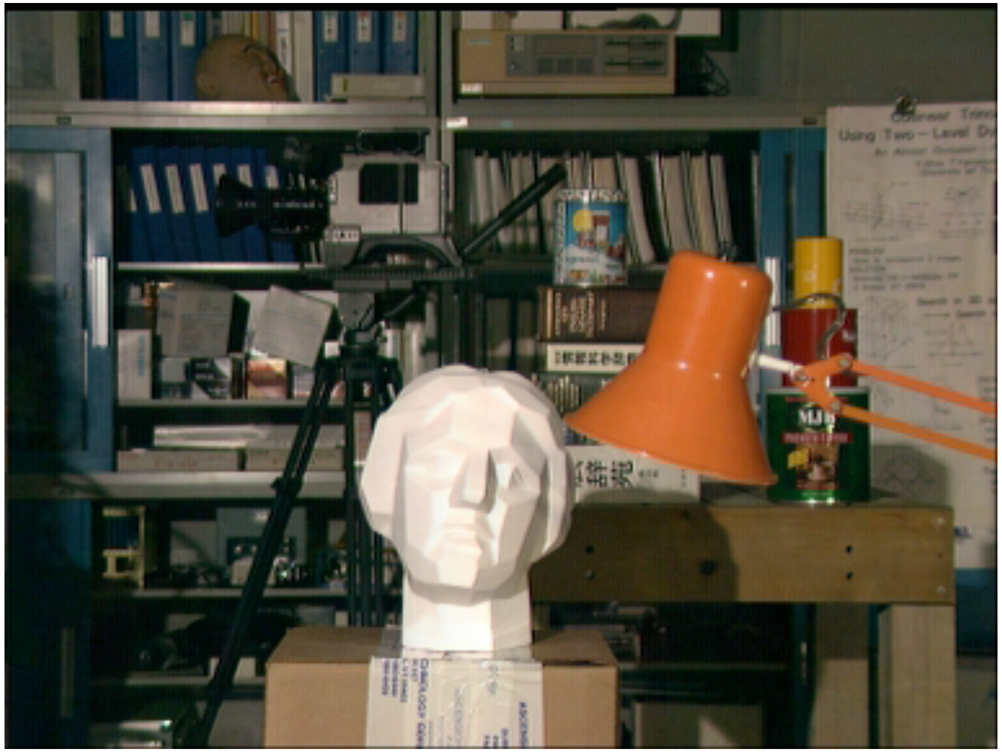
(see J.-Y. Bouguet,
Calibration Toolbox)

A very brief introduction into

Stereo and Motion Analysis

Stereo vision on rectified images

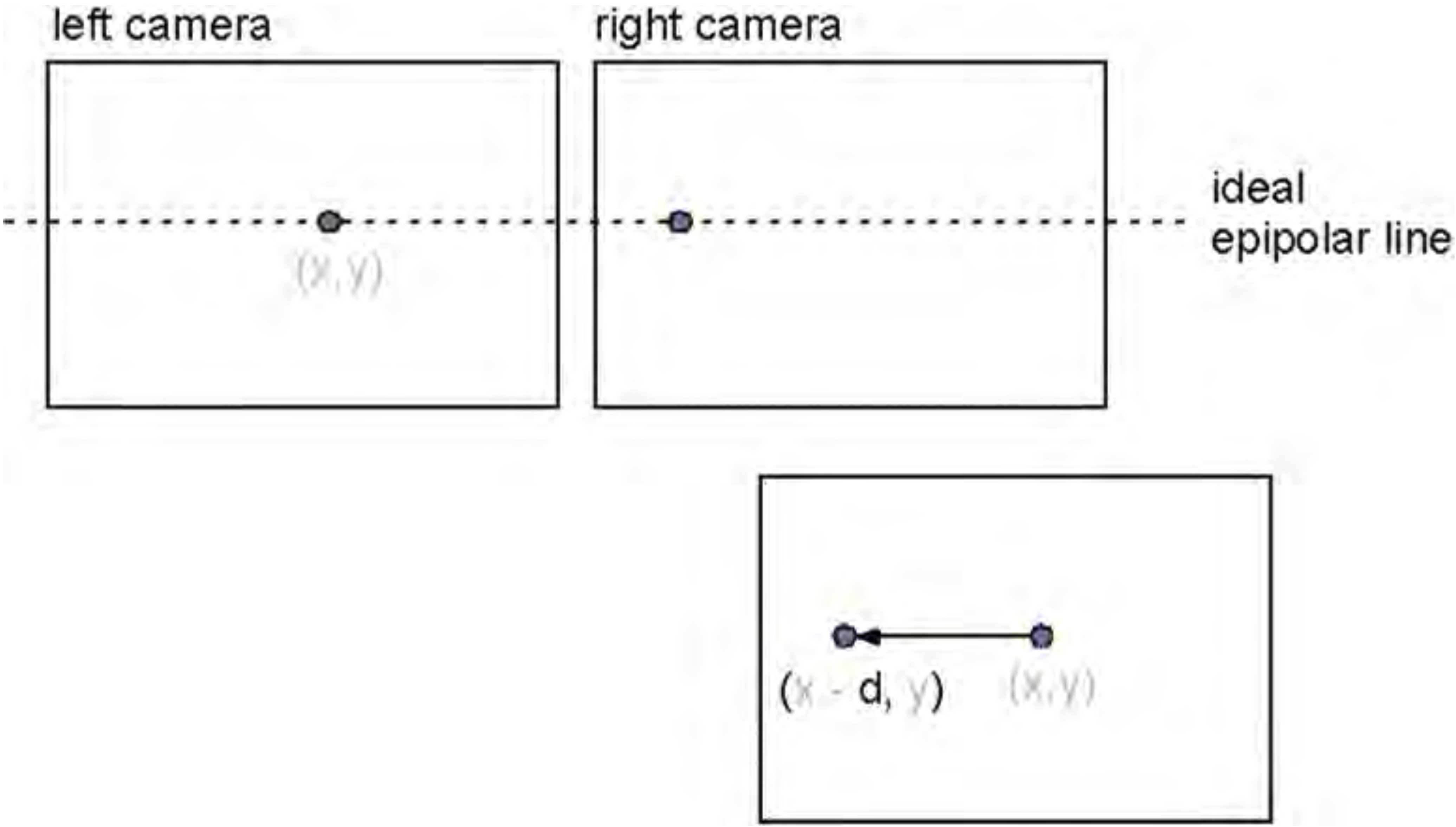
Left Image



Right Image

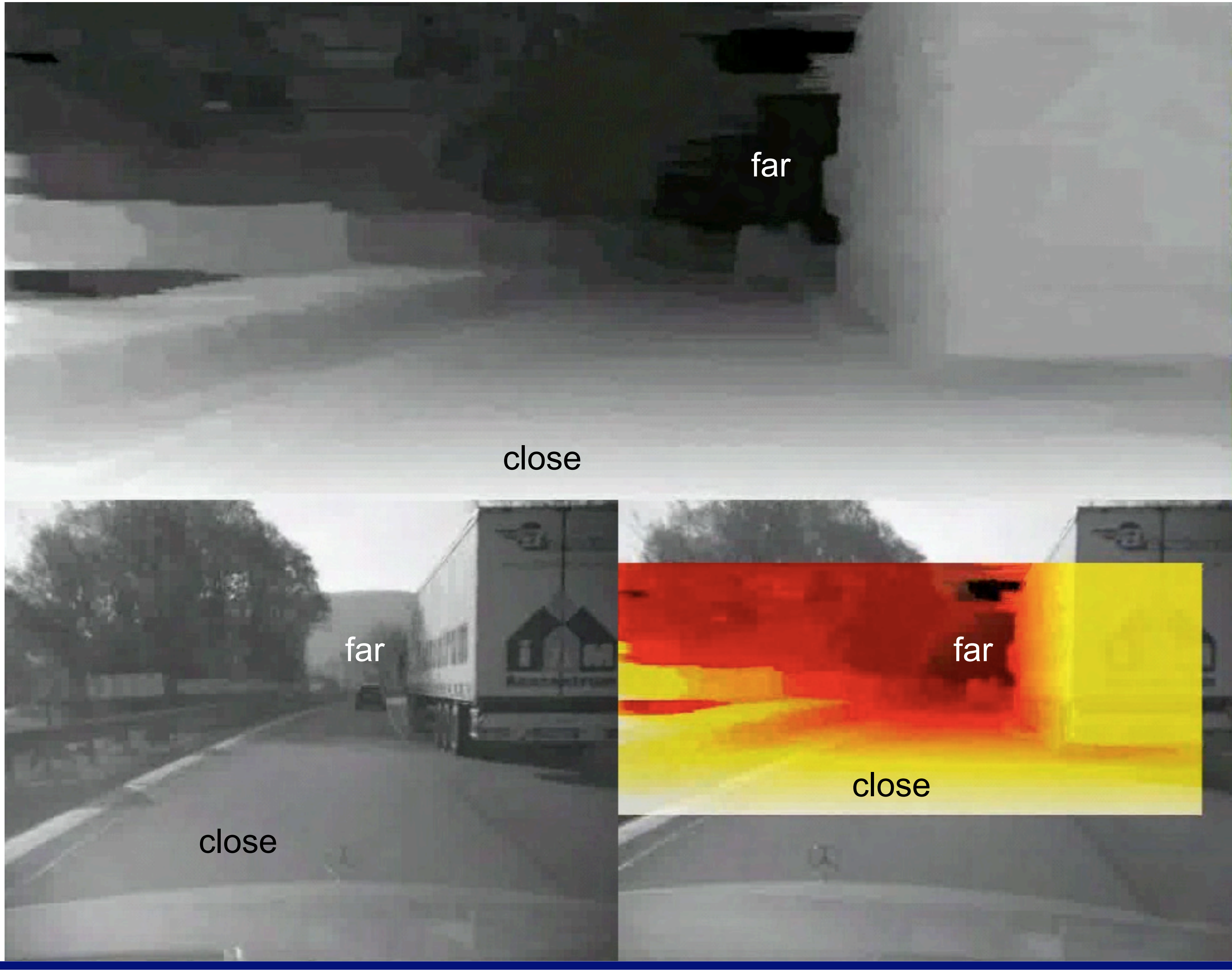
(Tsukuba stereo pair, see Middlebury stereo website)

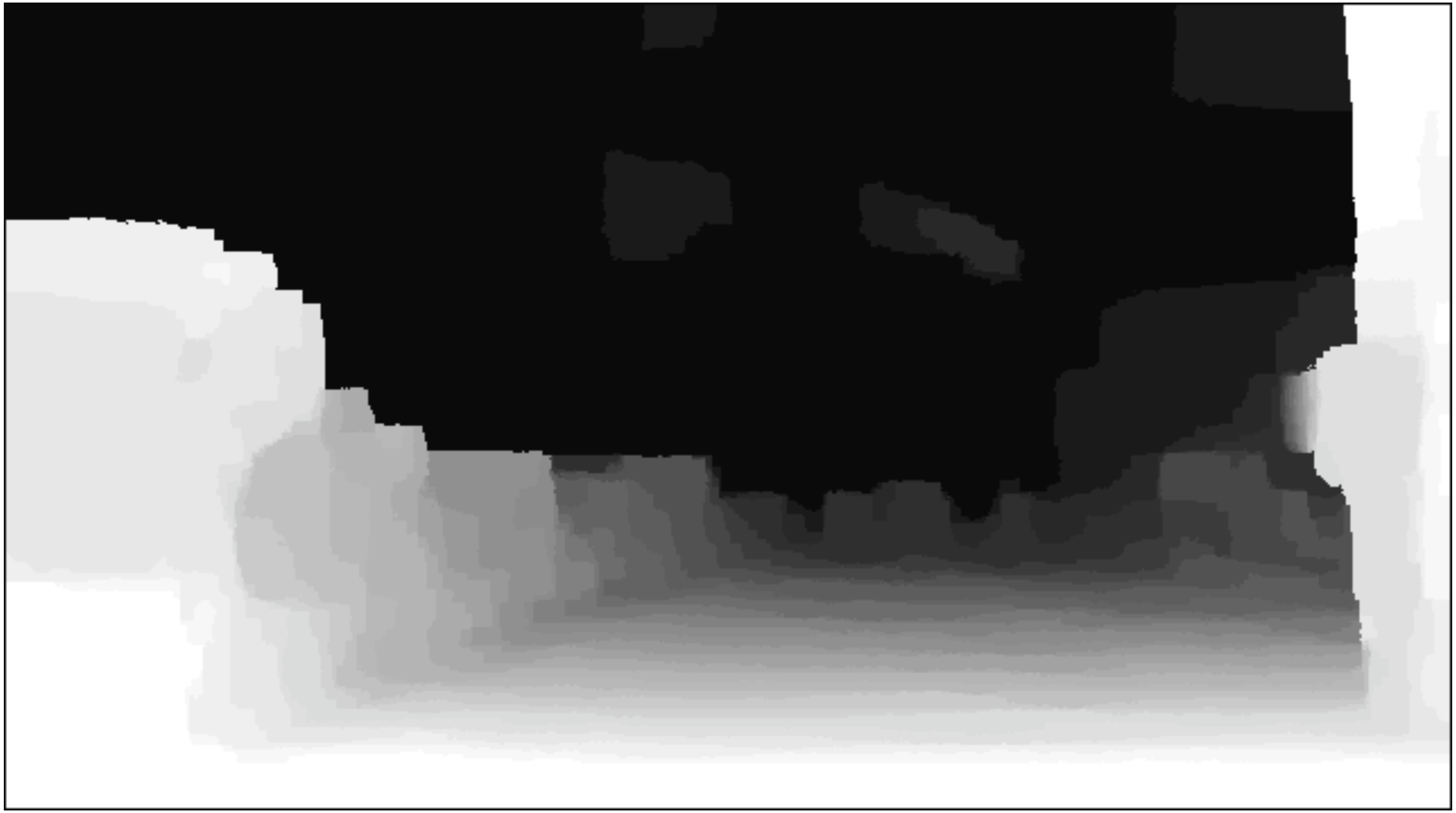
One-dimensional disparities between pairs of corresponding points



disparity *d* at *(x,y)* in left image

Disparities are mapped into depth and shown in gray scale or color.





Optical flow analysis on image sequences

Image 1

Image 2

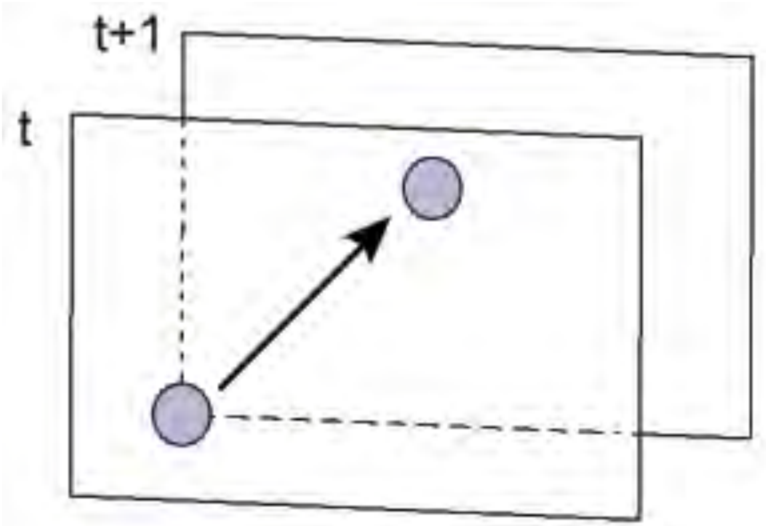
(engineered image sequence, Middlebury flow website)

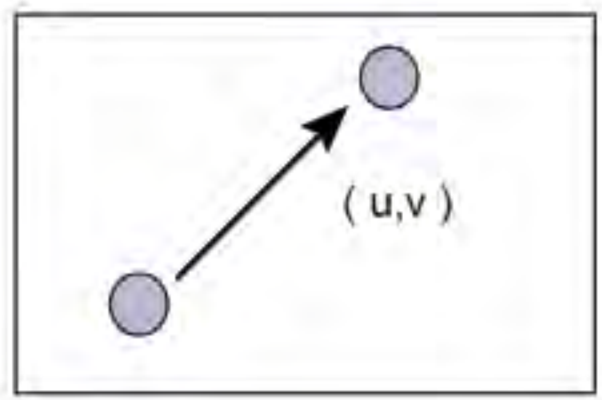
Two-dimensional flow vectors between pairs of corresponding points

Local motion (optic flow) estimation

flow key

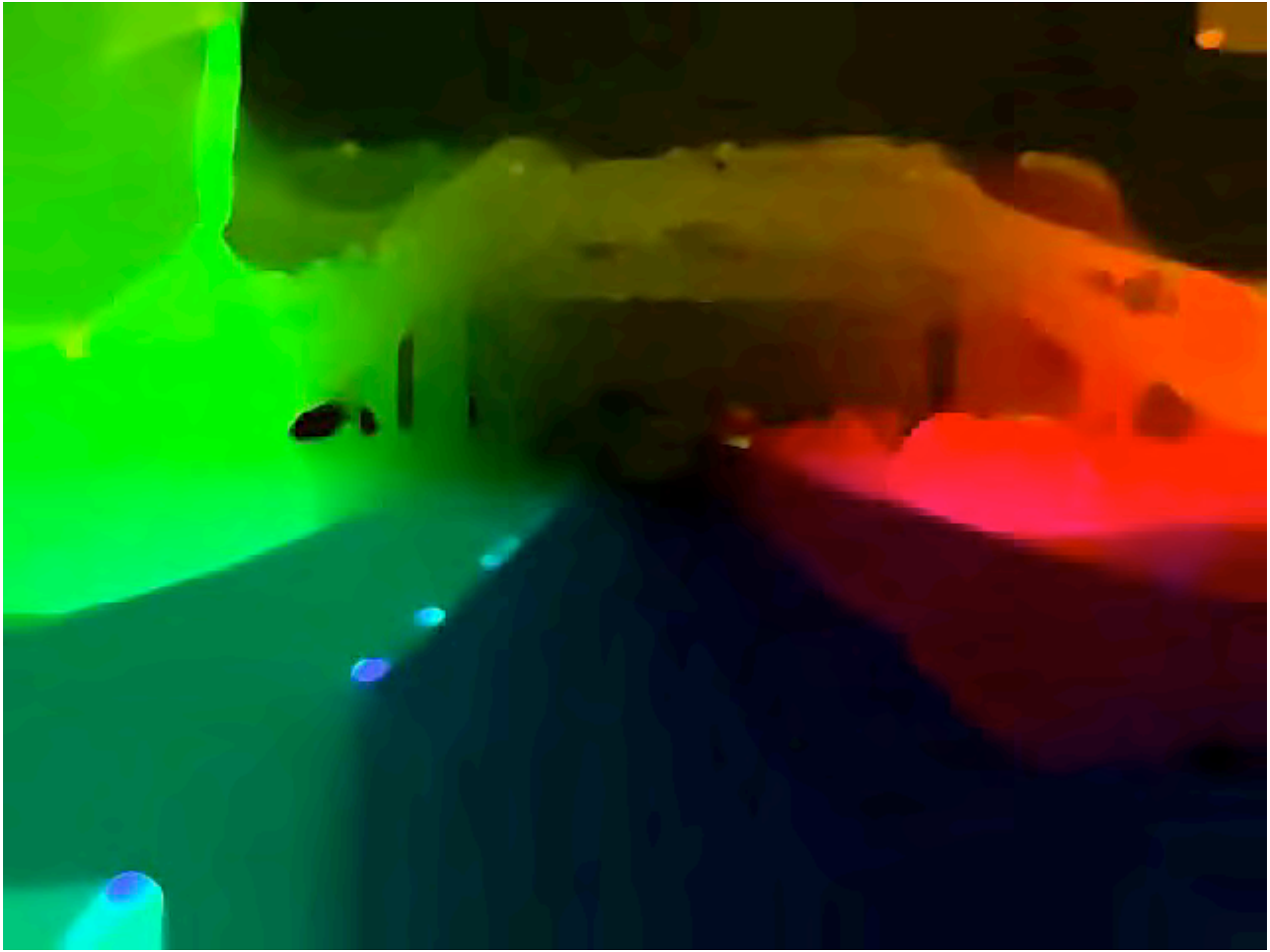
For each of the two sequences, motion vectors (u,v) are calculated by comparing frames along the time scale. Vectors at (x,y) are shown as one colored dot, representing direction by using the HSI scheme on the left; note that length is encoded by intensity (black = no motion)





BBPW algorithm on the Auckland Harbor Bridge

[Bruhn, Brox, Papenberg, Weickert, ECCV 2004]

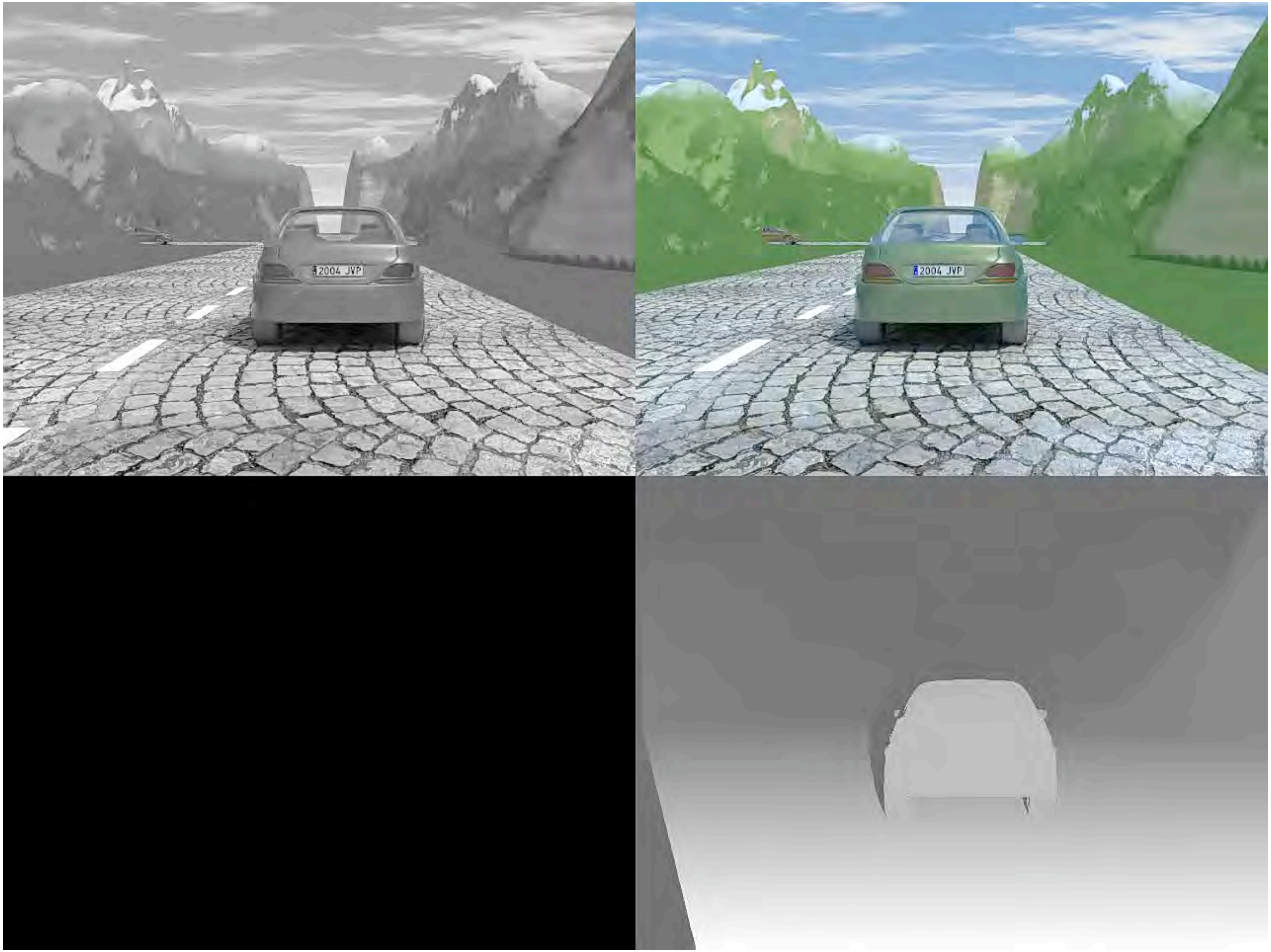


BBPW algorithm in the Waitakeres (left and right sequence separately)

Available data (with or without `ground truth') for

Performance Evaluation

Synthesized stereo sequence with ground truth

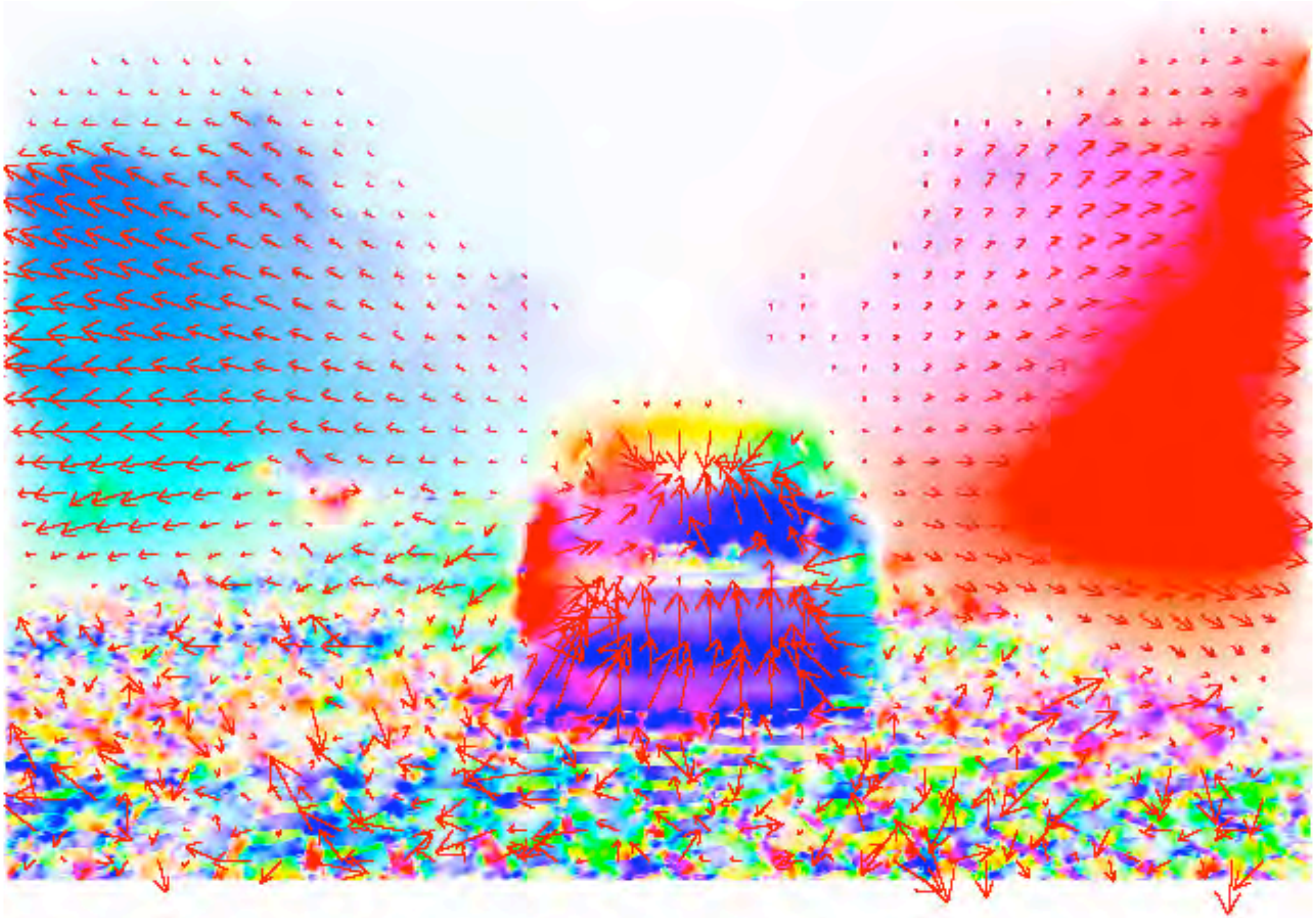


white =
no motion

CLG algorithm (combining local and global motion analysis)

[Bruhn, Weickert, Schnörr, IJCV 2005]

white =
no motion



applied to the synthetic stereo sequence (with ground truth) in Set 2 on

Performance of stereo vision sensors/programs are evaluated by using specific **metrics**: so far only a few simple metrics to advice with respect to a basically unlimited diversity of possible situations.

mean angular error (for ALL the *n* pixels)

$$E_{AE} = \frac{1}{n} \sum \arccos \left(\frac{\mathbf{u} \cdot \mathbf{u}_T}{|\mathbf{u}||\mathbf{u}_T|} \right)$$

endpoint error (for ALL the *n* pixels)

$$E_{EP} = \sqrt{(u - u_T)^2 + (v - v_T)^2}$$

Note: actually, only errors in **relevant regions** of interest

Real-world or synthetic image sequence test data

(partially with ground truth)

MI tech

Multimedia Imaging Technology

Home EISATS Discussion Contacts Login

Categories

- Home
- About MI
- Events
- Lectures
- Links
- People
- Projects
- Publications
- Positions and Scholarships
- Sponsors
- Technical reports

Topics

- Vision-Based Driver Assistance
- Panoramic Imaging
- Computer Vision
- Medical Imaging
- Enviromental Surveillance
- Geometry and Topology
- 3D TV
- Wireless Digital Video

.enpeda.. Image Sequence Analysis Test Site (EISATS)

.enpeda.. | Set 1 | Set 2 | Set 3 || vision Middlebury

This web site of the *.enpeda..* (Environmental Perception and Driver Assistance) project offers sets of geometrically rectified stereo image sequences for the purpose of comparative performance evaluation of stereo, motion, or 6D analysis techniques.

.enpeda..

Information on Data (READ THIS FIRST)

All grey-scale images are in PGM (Portable Grey Map) format and all colour images are in PPM (Portable Pixel Map) format. All binary data is Big Endian (as the current standard). Image depths are either 8-bit or 12-bit. (i.e., PGM has 1 channel of 8/12-bit, PPM has 3 channels of 8/12-bit).

These can be viewed through good image viewers (such as the freeware Infranview, which can also batch convert to other image formats) and can be read by OpenCV (C++) and Matlab. All data is compressed using 7-Zip format.

Set 1 Night Vision Stereo Sequences

These seven stereo night vision sequences (12 bit, between 220 and 300 pairs of frames each) have been provided by Daimler AG, Germany, in June 2007 (group of Dr. Uwe Franke). These sequences come with ego-motion data and time stamps for each frame.

Set 2 Synthesized Stereo Sequences

These synthesized stereo sequences (with ground truth) have been provided by Tobi Vaudrey (*.enpeda..*) and Clemens Rabe (Daimler AG).

Set 3 Color Stereo Sequences

These three day-time, color stereo sequences have been provided by the European *Drivsco* project.

Set 4 Gray-Level Stereo Sequences

These day-time, gray-level stereo sequences have been provided by Hella Aglaia A.G., Germany ... (soon to come).

www.mi.auckland.ac.nz/EISATS

#1: Construction-site sequence

#2: Save-turn sequence

#3: Squirrel sequence

#4: Dancing-light sequence

#5: Intern-on-bike sequence

#6: Traffic-light sequence

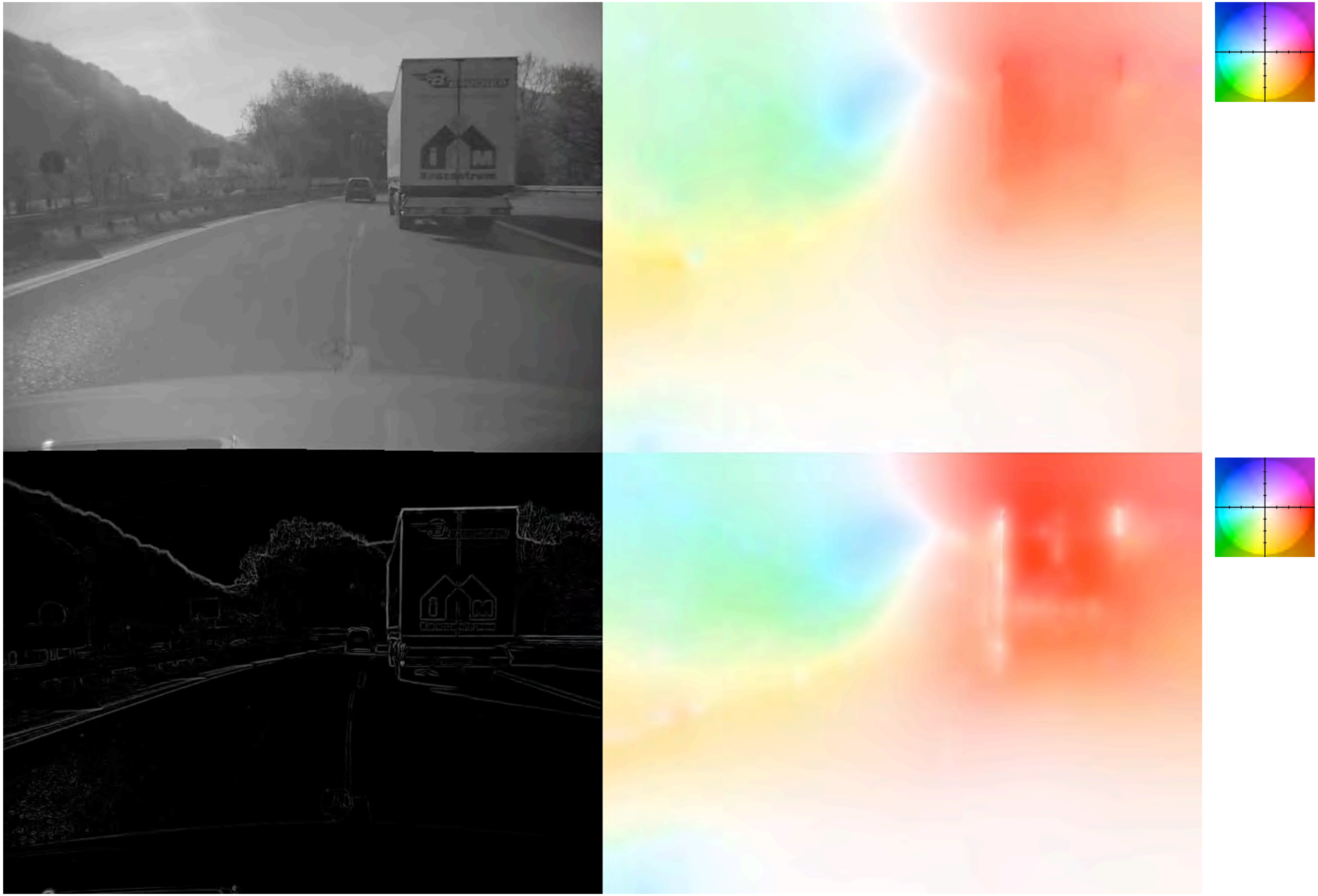
Set 1:

Seven stereo sequences

Daimler A.G. 2007

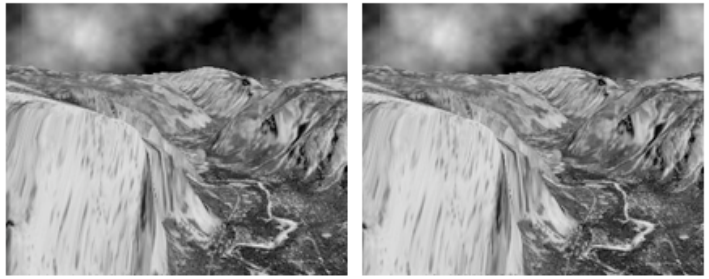
#7: Crazy-turn sequence

Original dancing light sequence Pyramid Horn-Schunck



Sobel edge image Pyramid HS on edge image

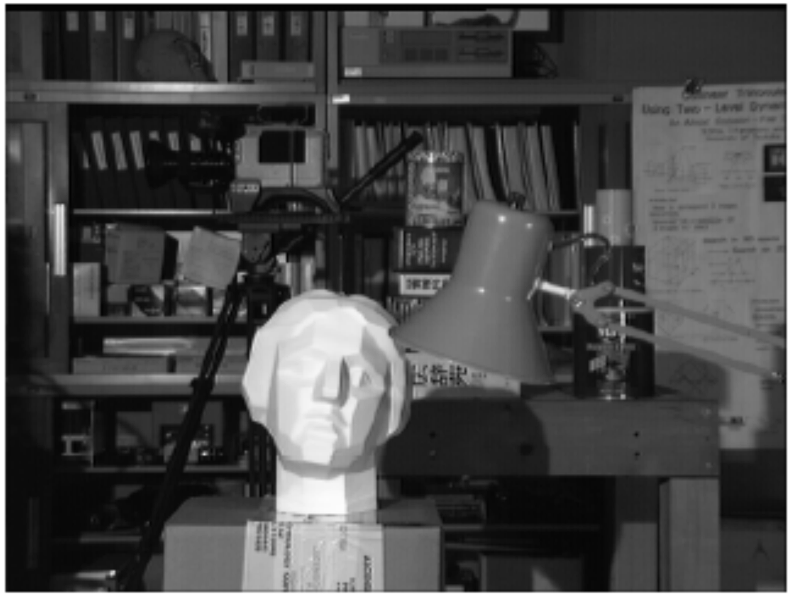
Comparative performance evaluation



- **Stereo algorithms**
  - An experimental comparison of stereo algorithms
    R. Szeliski & R. Zabih (2000)
  - Middlebury Stereo Test Site
    D. Scharstein and R. Szeliski (2002)
  - ....
- **Flow algorithms**
  - Performance of optical flow techniques
    J. Barron, D. Fleet, S. Beauchemin (1994)
  - Middlebury Flow Test Site
    S. Baker *et al.* (2007)
  - ...

Middlebury computer vision website

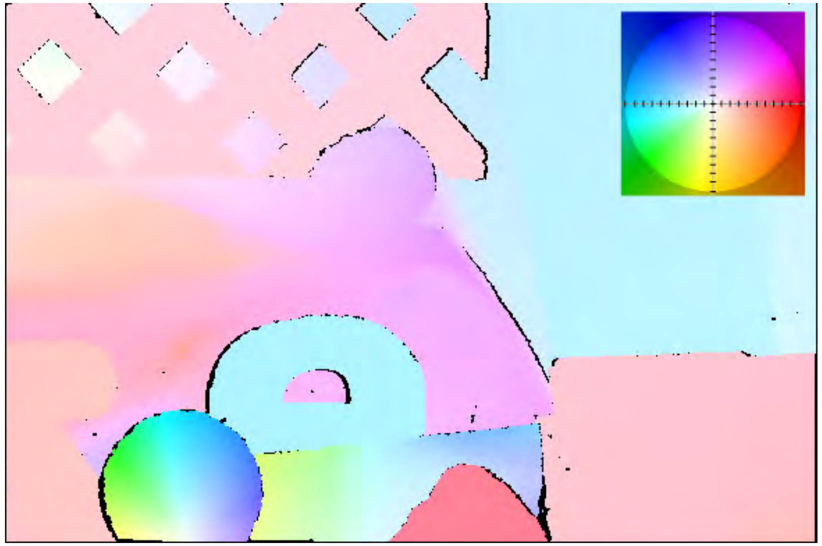
Stereo



Original Image

Ground Truth

Optical Flow



use of two or three metrics for defined subregions (non-occluded, textured, ...)

The Top 50 Stereo Algorithms on 1 Dec 2008

ranking by total accuracy of calculated disparities within whole images

engineered or synthetic stereo pairs

given with `ground truth'

each test set of relatively small cardinality (say, less than 10 images)

Similar approach for Flow Algorithms

Stereo | Evaluation · Datasets · Code · Submit

Middlebury Stereo Evaluation - Version 2

New features and main differences to version 1.
Submit and evaluate your own results.

Open a new window for each link

Error Threshold = 1 | Error Threshold... | Sort by nonocc | Sort by all | Sort by disc

Algorithm	Avg. Rank	Tsukuba ground truth nonocc	Tsukuba all	Tsukuba disc	Venus ground truth nonocc	Venus all	Venus disc	Teddy ground truth nonocc	Teddy all	Teddy disc	Cones ground truth nonocc	Cones all	Cones disc	Average percent of bad pixels (explanation)
AdaptingBP [17]	3.2	1.11 7	1.37 3	5.79 8	0.10 1	0.21 3	1.44 2	4.22 3	7.06 2	11.8 3	2.48 1	7.92 3	7.32 2	4.23
CoopRegion [41]	3.2	0.87 1	1.16 1	4.61 1	0.11 2	0.21 2	1.54 4	5.16 7	8.31 4	13.0 6	2.79 4	7.18 1	8.01 5	4.41
DoubleBP [35]	4.4	0.88 3	1.29 2	4.76 3	0.13 5	0.45 8	1.87 7	3.53 2	8.30 3	9.63 1	2.90 5	8.78 11	7.79 3	4.19
OutlierConf [42]	5.1	0.88 2	1.43 5	4.74 2	0.18 9	0.26 5	2.40 11	5.01 5	9.12 6	12.8 5	2.78 3	8.57 7	6.99 1	4.60
SubPixDoubleBP [30]	6.8	1.24 12	1.76 14	5.98 9	0.12 4	0.46 9	1.74 6	3.45 1	8.38 5	10.0 2	2.93 6	8.73 10	7.91 4	4.39
AdaptOvrSegBP [33]	11.7	1.69 24	2.04 22	5.64 7	0.14 6	0.20 1	1.47 3	7.04 17	11.1 9	16.4 13	3.60 13	8.96 13	8.84 12	5.59
PlaneFitBP [32]	12.6	0.97 6	1.83 15	5.26 6	0.17 8	0.51 11	1.71 5	6.65 11	12.1 15	14.7 8	4.17 23	10.7 22	10.6 21	5.78
SymBP+occ [7]	12.7	0.97 5	1.75 13	5.09 5	0.16 7	0.33 6	2.19 9	6.47 10	10.7 8	17.0 17	4.79 27	10.7 23	10.9 22	5.92
Undr+OvrSeg [48]	12.9	2.44 33	2.93 29	10.1 34	0.11 3	0.22 4	1.34 1	6.67 12	9.73 7	16.4 12	3.08 9	8.32 5	8.14 6	5.79
AdaptDispCalib [36]	13.8	1.19 10	1.42 4	6.15 11	0.23 11	0.34 7	2.50 13	7.80 22	13.6 23	17.3 20	3.62 14	9.33 14	9.72 17	6.10
Segm+visib [4]	14.3	1.30 17	1.57 6	6.92 21	0.79 25	1.06 22	6.76 28	5.00 4	6.54 1	12.3 4	3.72 15	8.62 9	10.2 19	5.40
C-SemiGlob [19]	14.9	2.61 35	3.29 30	9.89 32	0.25 14	0.57 13	3.24 18	5.14 6	11.8 10	13.0 6	2.77 2	8.35 6	8.20 7	5.76
SO+borders [29]	14.9	1.29 16	1.71 10	6.83 18	0.25 15	0.53 12	2.26 10	7.02 16	12.2 16	16.3 10	3.90 18	9.85 18	10.2 20	6.03
DistinctSM [27]	16.3	1.21 11	1.75 12	6.39 13	0.35 16	0.69 18	2.63 15	7.45 21	13.0 19	18.1 22	3.91 19	9.91 20	8.32 9	6.14
OverSegmBP [26]	16.9	1.69 25	1.97 19	8.47 27	0.50 21	0.68 17	4.69 23	6.74 13	11.9 14	15.8 9	3.19 10	8.81 12	8.89 13	6.11
CostAggr+occ [39]	17.2	1.38 19	1.96 18	7.14 22	0.44 19	1.13 24	4.87 24	6.80 14	11.9 12	17.3 19	3.60 12	8.57 8	9.36 15	6.20
SegmentSupport [28]	17.3	1.25 13	1.62 8	6.68 15	0.25 13	0.64 16	2.59 14	8.43 26	14.2 24	18.2 23	3.77 16	9.87 19	9.77 18	6.44
RegionTreeDP [18]	18.3	1.39 21	1.64 9	6.85 19	0.22 10	0.57 13	1.93 8	7.42 20	11.9 13	16.8 15	6.31 36	11.9 29	11.8 26	6.56
EnhancedBP [24]	19.0	0.94 4	1.74 11	5.05 4	0.35 17	0.86 19	4.34 22	8.11 26	13.3 21	18.5 26	5.09 30	11.1 25	11.0 23	6.69
AdaptWeight [12]	20.7	1.38 19	1.85 16	6.90 20	0.71 23	1.19 25	6.13 25	7.88 23	13.3 22	18.6 28	3.97 22	9.79 17	8.26 8	6.67
InteriorPtLP [34]	20.8	1.27 14	1.62 7	6.82 17	1.15 29	1.67 28	12.7 39	8.07 24	11.9 11	18.7 29	3.92 21	9.68 15	9.62 16	7.26
SegTreeDP [22]	20.9	2.21 31	2.76 26	10.3 35	0.46 20	0.60 15	2.44 12	9.58 32	15.2 31	18.4 25	3.23 11	7.86 2	8.83 11	6.82
ImproveSubPix [25]	21.8	3.00 37	3.61 34	10.9 37	0.88 27	1.47 26	7.10 30	7.12 18	12.4 18	16.6 14	2.96 7	8.22 4	8.55 10	6.90
SemiGlob [6]	23.0	3.26 38	3.96 35	12.8 40	1.00 28	1.57 27	11.3 34	6.02 8	12.2 17	16.3 11	3.06 8	9.75 16	8.90 14	7.50
VariableCross [44]	26.3	1.99 28	2.65 25	6.77 16	0.62 22	0.96 21	3.20 17	9.75 33	15.1 29	18.2 24	6.28 35	12.7 32	12.9 34	7.60
RealtimeBP [21]	26.6	1.49 22	3.40 32	7.87 25	0.77 24	1.90 31	9.00 33	8.72 30	13.2 20	17.2 18	4.61 25	11.6 27	12.4 32	7.69
2OP+occ [37]	27.2	2.91 36	3.56 33	7.33 24	0.24 12	0.49 10	2.76 16	10.9 38	15.4 32	20.6 35	5.42 33	10.8 24	12.5 33	7.75
CCH+SegAggr [47]	27.7	1.74 26	2.11 23	9.23 30	0.41 18	0.94 20	3.97 21	8.08 25	14.3 25	19.8 33	7.07 39	12.9 33	16.3 39	8.07
FastAggreg [45]	28.3	1.16 8	2.11 24	6.06 10	4.03 43	4.75 42	6.43 26	9.04 31	15.2 30	20.2 34	5.37 32	12.6 31	11.9 28	8.24
GC+occ [2]	28.5	1.19 9	2.01 21	6.24 12	1.64 35	2.19 33	6.75 27	11.2 39	17.4 38	19.8 32	5.36 31	12.4 30	13.0 35	8.26
Layered [5]	29.2	1.57 23	1.87 17	8.28 26	1.34 32	1.85 29	6.85 29	8.64 29	14.3 26	18.5 27	6.59 38	14.7 37	14.4 37	8.24
MultiCamGC [3]	29.2	1.27 15	1.99 20	6.48 14	2.79 41	3.13 38	3.60 20	12.0 40	17.6 39	22.0 38	4.89 28	11.8 28	12.1 29	8.31
ConvexTV [46]	30.4	3.61 39	5.72 41	18.0 46	1.16 30	2.50 35	12.4 38	6.10 9	15.7 33	16.8 16	3.88 17	14.4 36	11.5 25	9.30
AdaptPolygon [43]	30.8	2.29 32	2.88 28	8.94 29	0.80 26	1.11 23	3.41 19	10.5 37	15.9 34	21.3 37	6.13 34	13.2 34	13.3 36	8.32
GenModel [20]	32.3	2.57 34	4.74 38	13.0 41	1.72 36	3.08 37	16.9 41	6.86 15	15.0 28	19.2 30	4.64 26	14.9 38	11.4 24	9.50
TensorVoting [9]	32.9	3.79 40	4.79 39	8.86 28	1.23 31	1.88 30	11.5 35	9.76 34	17.0 37	24.0 41	4.38 24	11.4 26	12.2 30	9.25
RealTimeGPU [14]	33.3	2.05 30	4.22 37	10.6 36	1.92 38	2.98 36	20.3 43	7.23 19	14.4 27	17.6 21	6.41 37	13.7 35	16.5 40	9.82
CostRelax [11]	34.5	4.76 45	6.08 43	20.3 47	1.41 34	2.48 34	18.5 42	8.18 27	15.9 35	23.8 39	3.91 20	10.2 21	11.8 27	10.6
ReliabilityDP [13]	35.6	1.36 18	3.39 31	7.25 23	2.35 40	3.48 41	12.2 37	9.82 35	16.9 36	19.5 31	12.9 47	19.9 46	19.7 42	10.7
TreeDP [8]	37.0	1.99 29	2.84 27	9.96 33	1.41 33	2.10 32	7.74 31	15.9 44	23.9 44	27.1 45	10.0 43	18.3 42	18.9 41	11.7
GC [1d]	37.9	1.94 27	4.12 36	9.39 31	1.79 37	3.44 40	8.75 32	16.5 45	25.0 46	24.9 42	7.70 40	18.2 41	15.3 38	11.4
BP+MLH [40]	38.6	4.17 43	6.34 44	14.6 43	1.96 39	3.31 39	16.8 40	10.2 36	18.9 40	24.0 40	4.93 29	15.5 39	12.3 31	11.1
DP [1b]	42.7	4.12 42	5.04 40	12.0 38	10.1 50	11.0 50	21.0 44	14.0 41	21.6 41	20.6 35	10.5 44	19.1 43	21.1 44	14.2
PhaseBased [31]	44.6	4.26 44	6.53 45	15.4 44	6.71 45	8.16 45	26.4 47	14.5 42	23.1 42	25.5 43	10.8 46	20.5 47	21.2 45	15.3
RegionalSup [38]	44.8	3.99 41	6.05 42	14.2 42	8.14 46	9.68 47	36.8 49	18.3 48	26.7 48	32.1 47	9.16 41	19.3 44	19.9 43	17.0
SSD+MF [1a]	44.9	5.23 48	7.07 46	24.1 48	3.74 42	5.16 43	11.9 36	16.5 46	24.8 45	32.9 48	10.6 45	19.8 45	26.3 47	15.7
STICA [16]	46.4	7.70 49	9.63 50	27.8 49	8.19 47	9.58 46	40.3 50	15.8 43	23.2 43	37.7 49	9.80 42	17.8 40	28.7 49	19.7
SO [1c]	46.9	5.08 47	7.22 48	12.2 39	9.44 49	10.9 49	21.9 45	19.9 49	28.2 50	26.3 44	13.0 48	22.8 49	22.3 46	16.6
PhaseDiff [23]	47.8	4.89 46	7.11 47	16.3 45	8.34 48	9.76 48	26.0 46	20.0 50	28.0 49	29.0 46	19.8 50	28.5 50	27.5 48	16.8
Infection [10]	48.0	7.95 50	9.54 49	28.9 50	4.41 44	5.53 44	31.7 48	17.7 47	25.1 47	44.4 50	14.3 49	21.3 48	38.0 50	20.7

The Middlebury Effect

- Great leaps forward in stereo and flow algorithms!
- Very accurate results for these specific scenes
- Very strong influence on differential methods
- no focus on using multiple time frames (e.g., Kalman filtering)

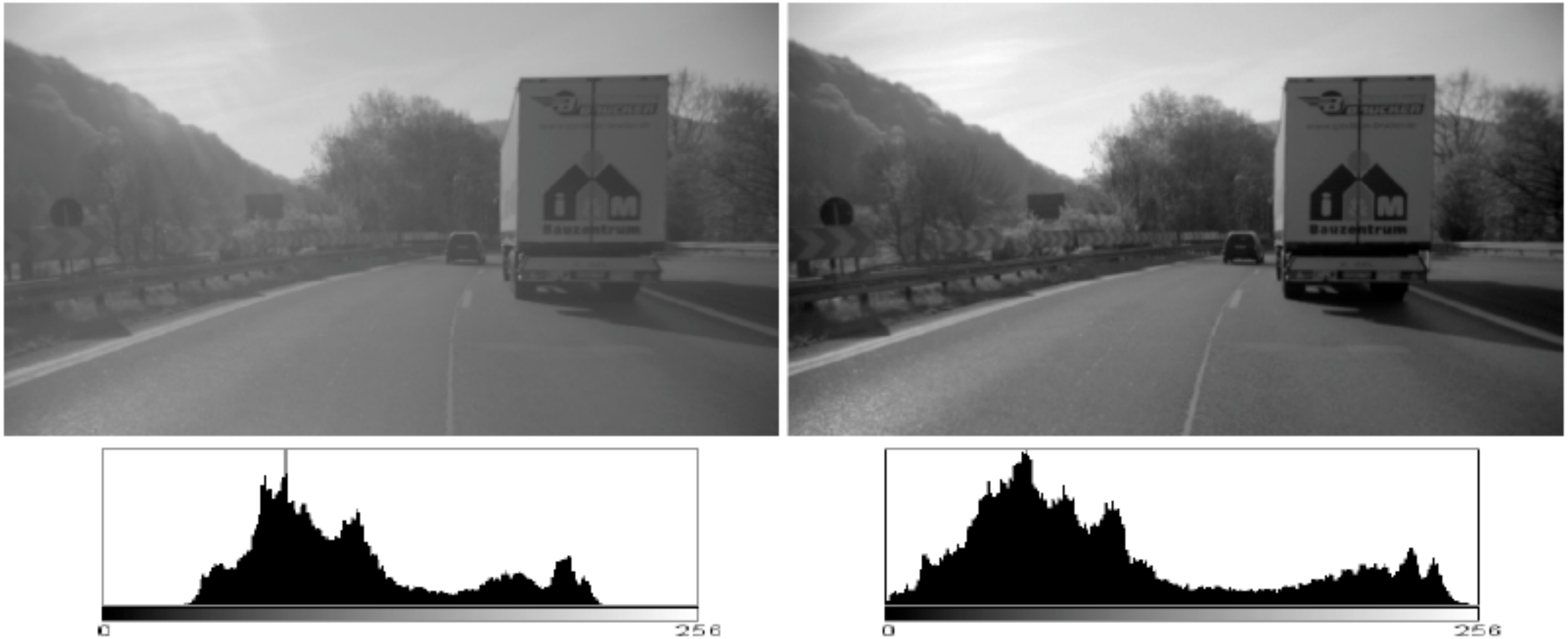
Differences in image data

Real-world images

low contrast, gray-value, but 10 or 12 bit
homogeneous regions, specularities
differences in brightness (e.g., left to right)
etc. - just REAL WORLD

Synthetic images
(engineered or rendered)

Brightness differences between left and right image

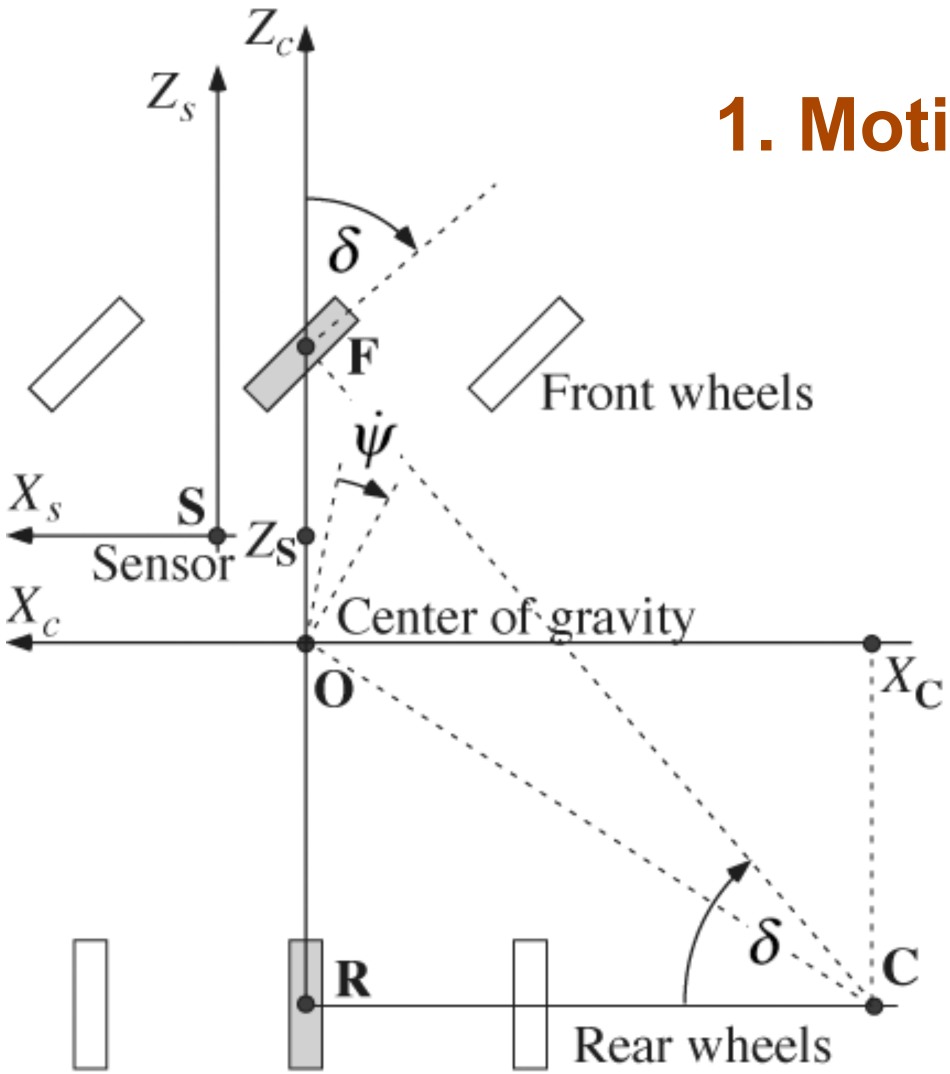


causes BP to fail (in difference to DP, SGM MI, or BT) - so far not discussed on Middlebury stereo page

Examples of approaches to provide ...

Ground Truth for Real-World Sequences

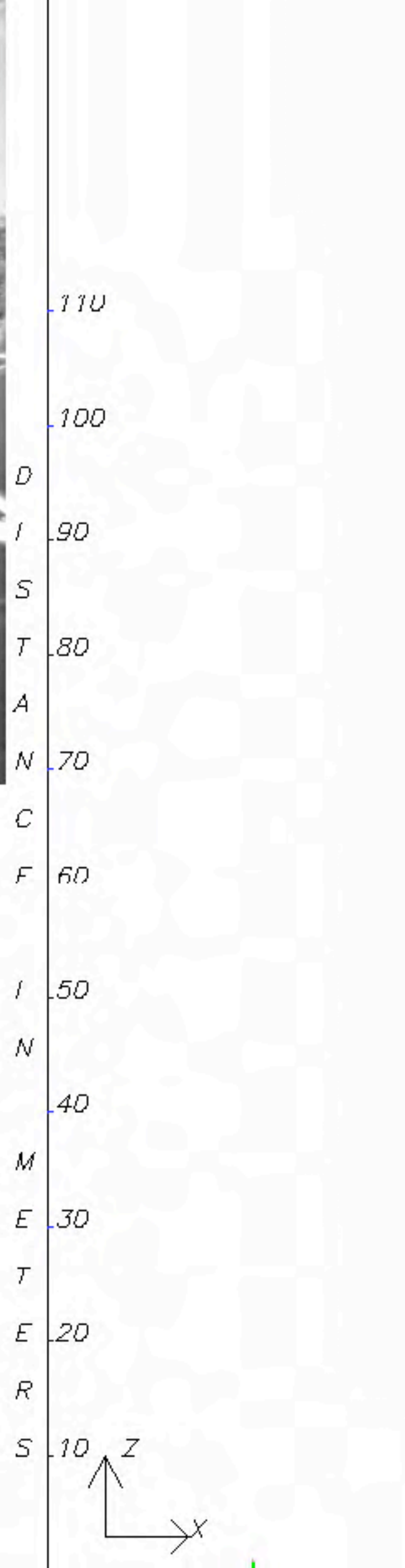
1. Motion parameters of ego-vehicle by reading sensor data



We read yaw rate and velocity at 25 Hz from the car's computer, thanks to the research department at

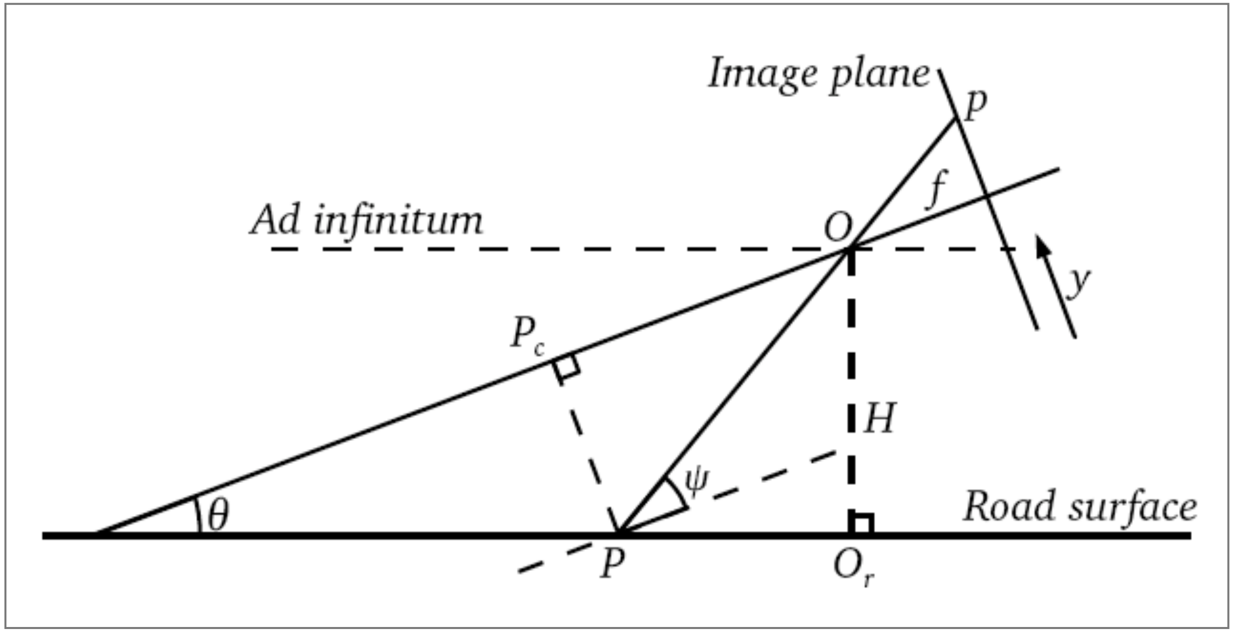
Daimler AG Germany

GPS: Blackhawk Tracking Devices Ltd



Trajectory of the ego-vehicle (assumed to drive on a plane) for the `bird sequence' as seen before

2. Adaptive tilt estimation based on disparities



d is the disparity at P
b is the base distance
θ tilt angle (pitch)

See [Liu and Klette, AI 2008]

$$\theta = \arcsin\left(\frac{H\cos\psi \cdot d}{b \cdot f}\right) - \psi$$

$$\psi = \arctan\left(\frac{(y_p - y_0)s_y}{f}\right)$$

Example: estimated mean tilt angles

for (subsequent, disjoint) intervals of

10 stereo frames (taken in 0.4 sec)

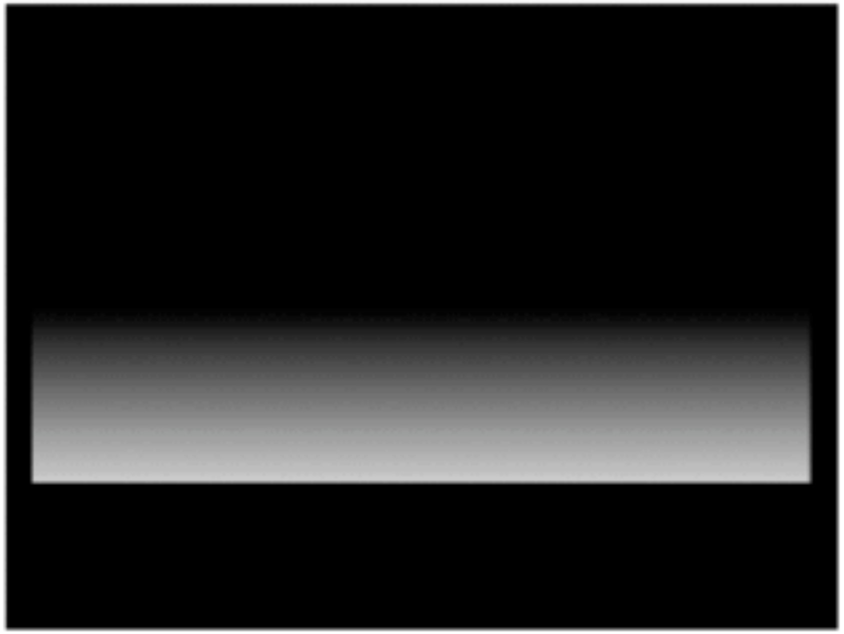
within a recorded stereo sequence

First pair of frames	1	11	21	31	41	51	61	71	81	91	101	111
Tilt angle (10^{-3} of a radian)	80	71	60	60	62	63	65	70	77	71	63	66
First pair of frames	121	131	141	151	161	171	181	191	201	211	221	231
Tilt angle (10^{-3} of a radian)	60	50	50	59	58	54	55	56	58	53	53	42

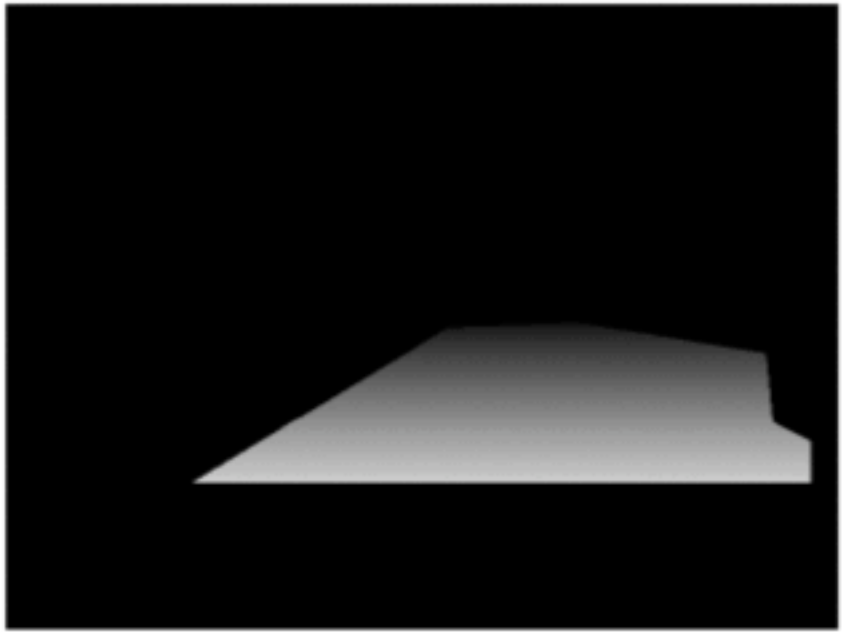
See [Liu and Klette, PSIVT 2009]

Image in a recorded sequence

Road mask

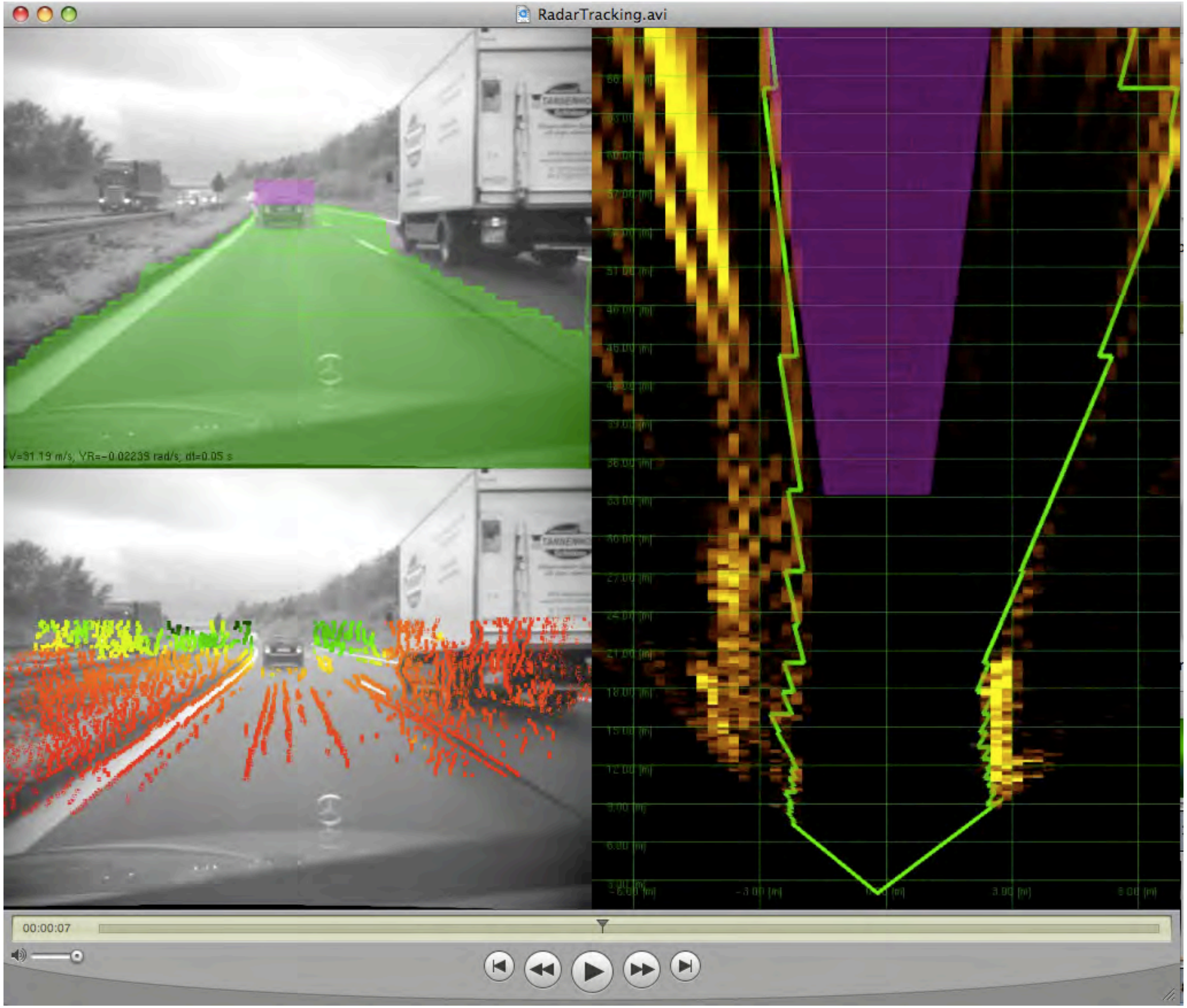


Disparities due to estimated tilt



Road mask with disparities
(assuming `local road planarity')

3. Radar for distance to lead vehicle

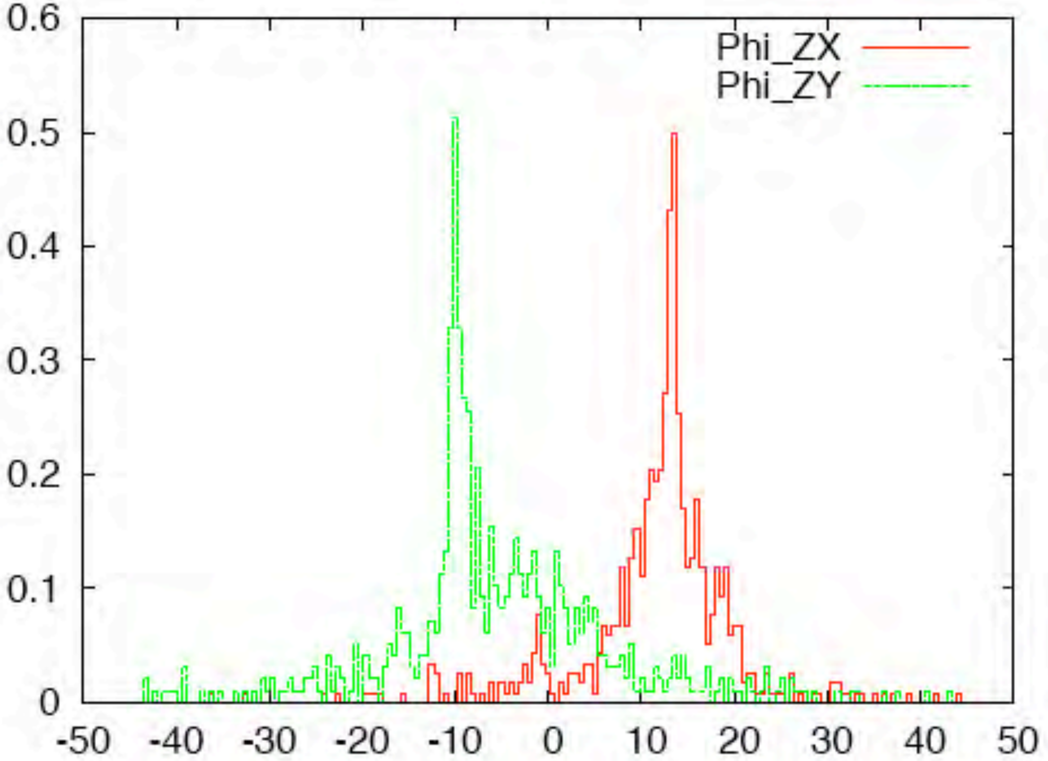


4. Indoor sequence recorded under defined conditions

constant and known camera viewing angles towards a ground plane

translation of calibrated camera with respect to a scene on this plane

Image of a recorded sequence



Example of estimated viewing angles
ground truth: -10 and 12 degrees

Examples of ...

Evaluations on Different Input Data

1. Improving dynamic programming stereo by propagation

Sequence from Set 1
www.mi.auckland.ac.nz/EISATS

Dynamic programming stereo with temporal propagation
[Liu and Klette, AI 2008]

Root Mean Squared Error and Bad Matches (> 1 pixel difference)

Sequence Name	Num of Frames	RMS	Bad Match %
1: 2007-03-06_121807	300	0.020271	2.68%
2: 2007-03-07_144703	300	0.023257	8.51%
3: 2007-03-15_182043	300	0.023400	23.11%
4: 2007-04-20_083101	250	0.067744	21.40%
5: 2007-04-27_145842	250	0.063743	17.50%
6: 2007-04-27_155554	250	0.071799	44.78%
7: 2007-05-08_132636	220	0.056440	35.75%

Original DP (without propagation) on road areas

Root Mean Squared Error and Bad Matches (> 1 pixel difference)

Sequence Name	Num of Frames	RMS	Bad Match %
1: 2007-03-06_121807	300	0.026014	15.12%
2: 2007-03-07_144703	300	0.053607	51.87%
3: 2007-03-15_182043	300	0.025122	40.37%
4: 2007-04-20_083101	250	0.069820	46.37%
5: 2007-04-27_145842	250	0.064231	24.85%
6: 2007-04-27_155554	250	0.074456	58.16%
7: 2007-05-08_132636	220	0.061994	50.80%

DP with spatial propagation on road areas

Root Mean Squared Error and Bad Matches (> 1 pixel difference)

Sequence Name	Num of Frames	RMS	Bad Match %
1: 2007-03-06_121807	300	0.019513	1.88%
2: 2007-03-07_144703	300	0.018198	3.28%
3: 2007-03-15_182043	300	0.022127	17.75%
4: 2007-04-20_083101	250	0.067528	19.24%
5: 2007-04-27_145842	250	0.063678	16.37%
6: 2007-04-27_155554	250	0.071739	45.28%
7: 2007-05-08_132636	220	0.054376	32.87%

DP with temporal propagation on road areas

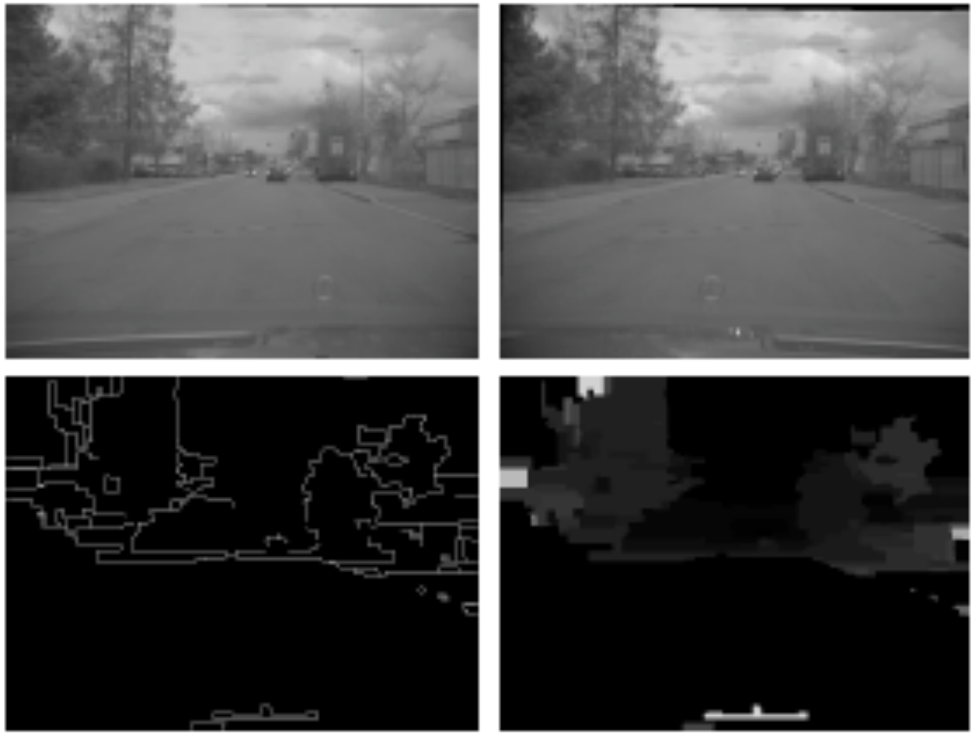
Root Mean Squared Error and Bad Matches (> 1 pixel difference)

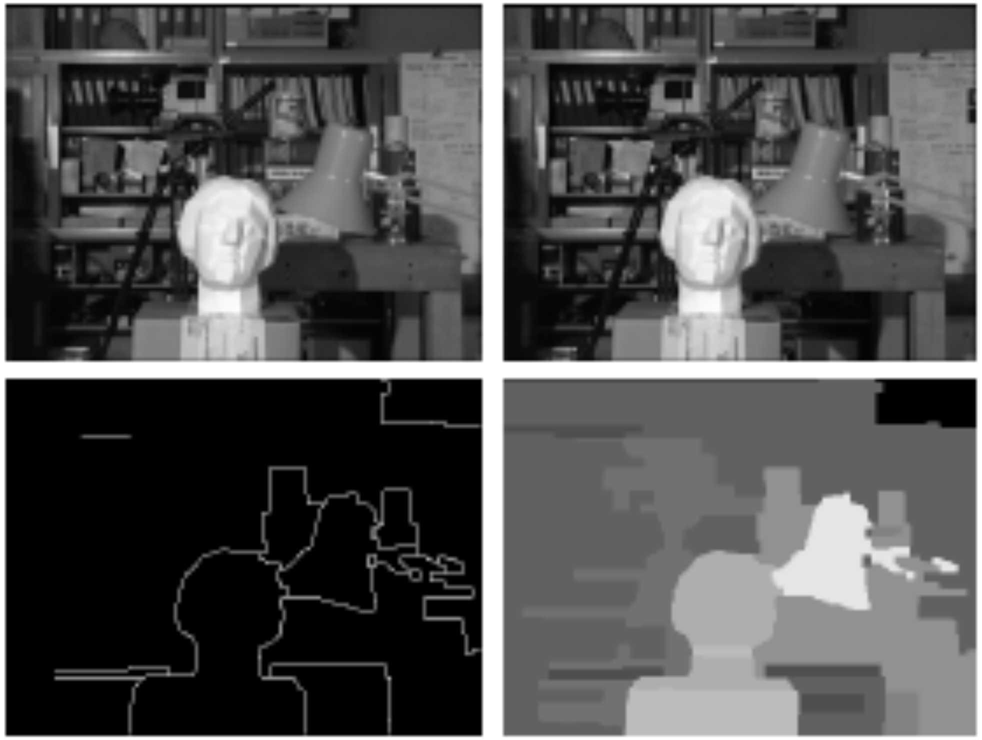
Sequence Name	Num of Frames	RMS	Bad Match %
1: 2007-03-06_121807	300	0.020348	5.54%
2: 2007-03-07_144703	300	0.038693	18.98%
3: 2007-03-15_182043	300	0.022827	24.36%
4: 2007-04-20_083101	250	0.067902	31.79%
5: 2007-04-27_145842	250	0.063755	16.19%
6: 2007-04-27_155554	250	0.072373	51.01%
7: 2007-05-08_132636	220	0.058989	41.79%

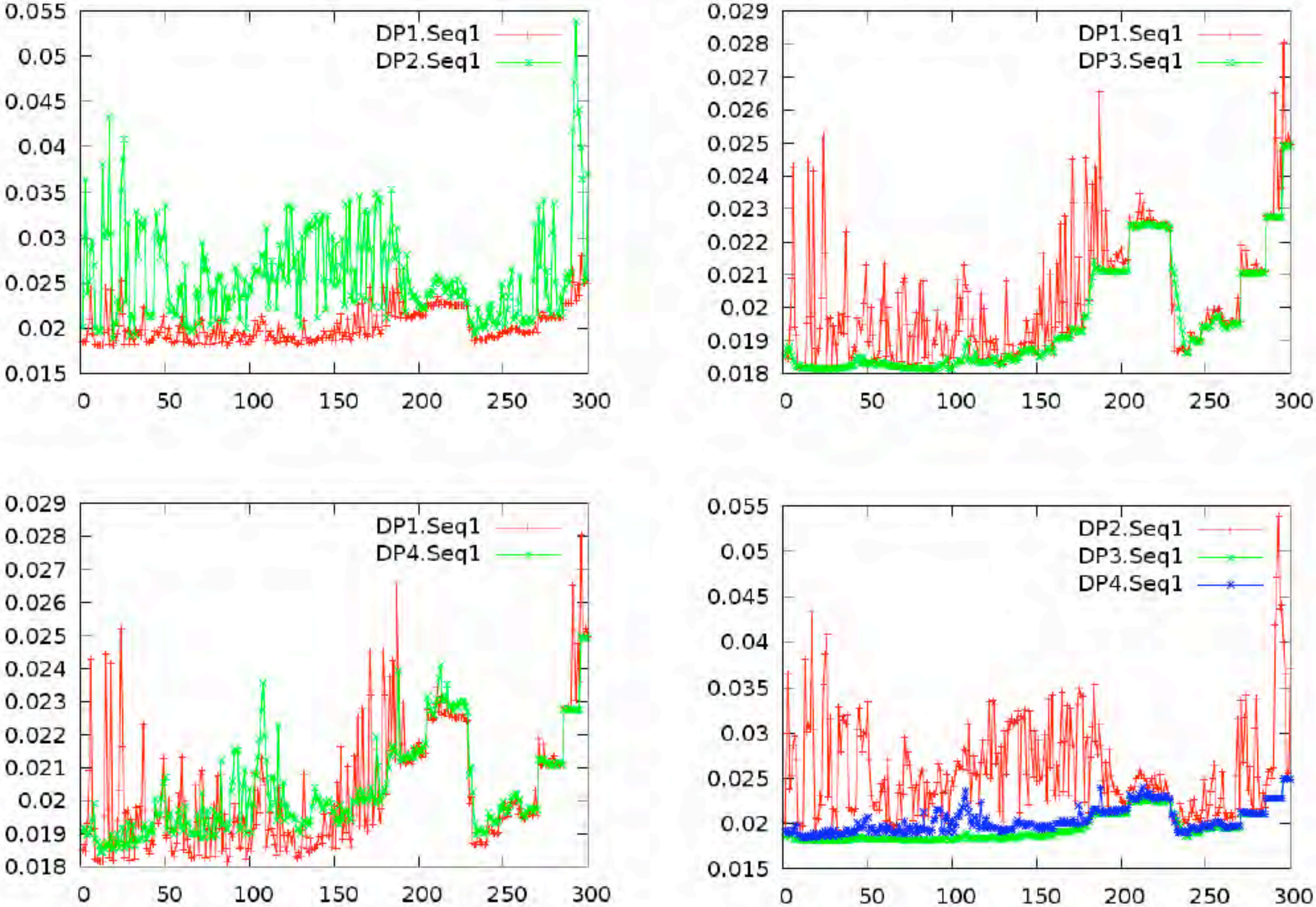
DP with spatio-temporal propagation on road areas

Sequence Name	Num of Frames	RMS	Bad Match %
1: 2007-03-06_121807	300	0.089806	60.98%
2: 2007-03-07_144703	300	0.105662	96.80%
3: 2007-03-15_182043	300	0.109850	81.04%
4: 2007-04-20_083101	250	0.125842	99.21%
5: 2007-04-27_145842	250	0.116894	94.63%
6: 2007-04-27_155554	250	0.135165	99.82%
7: 2007-05-08_132636	220	0.104936	99.43%

Birchfield-Tomasi
on road areas







Note: analysis along sequences allows a new quality (compared to small input sets)

2. Improving belief propagation stereo by preprocessing

Original left input sequence

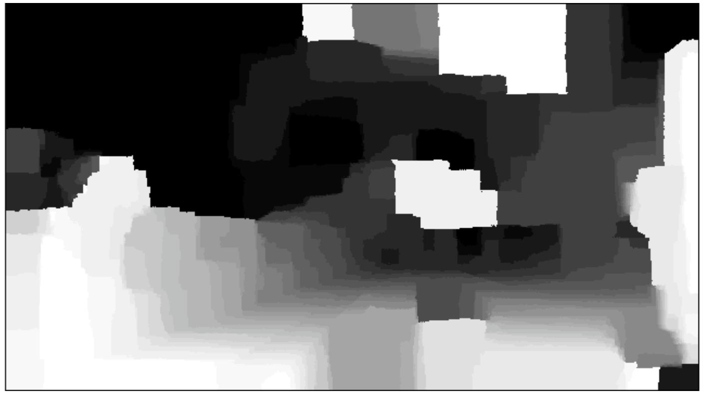
BP on original input sequences

Sobel of left input sequence

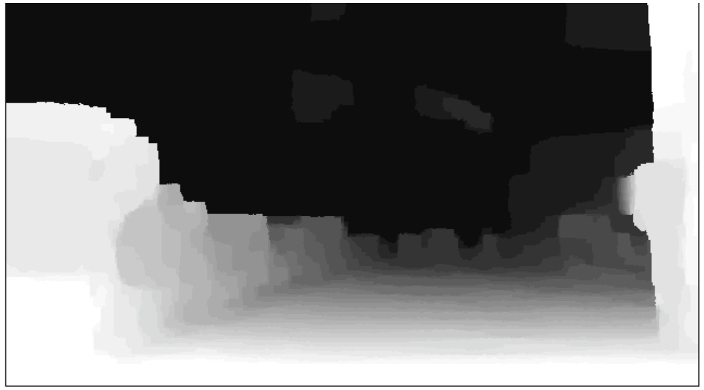
BP on Sobel input sequences

[Guan and Klette, Robot Vision 2008]

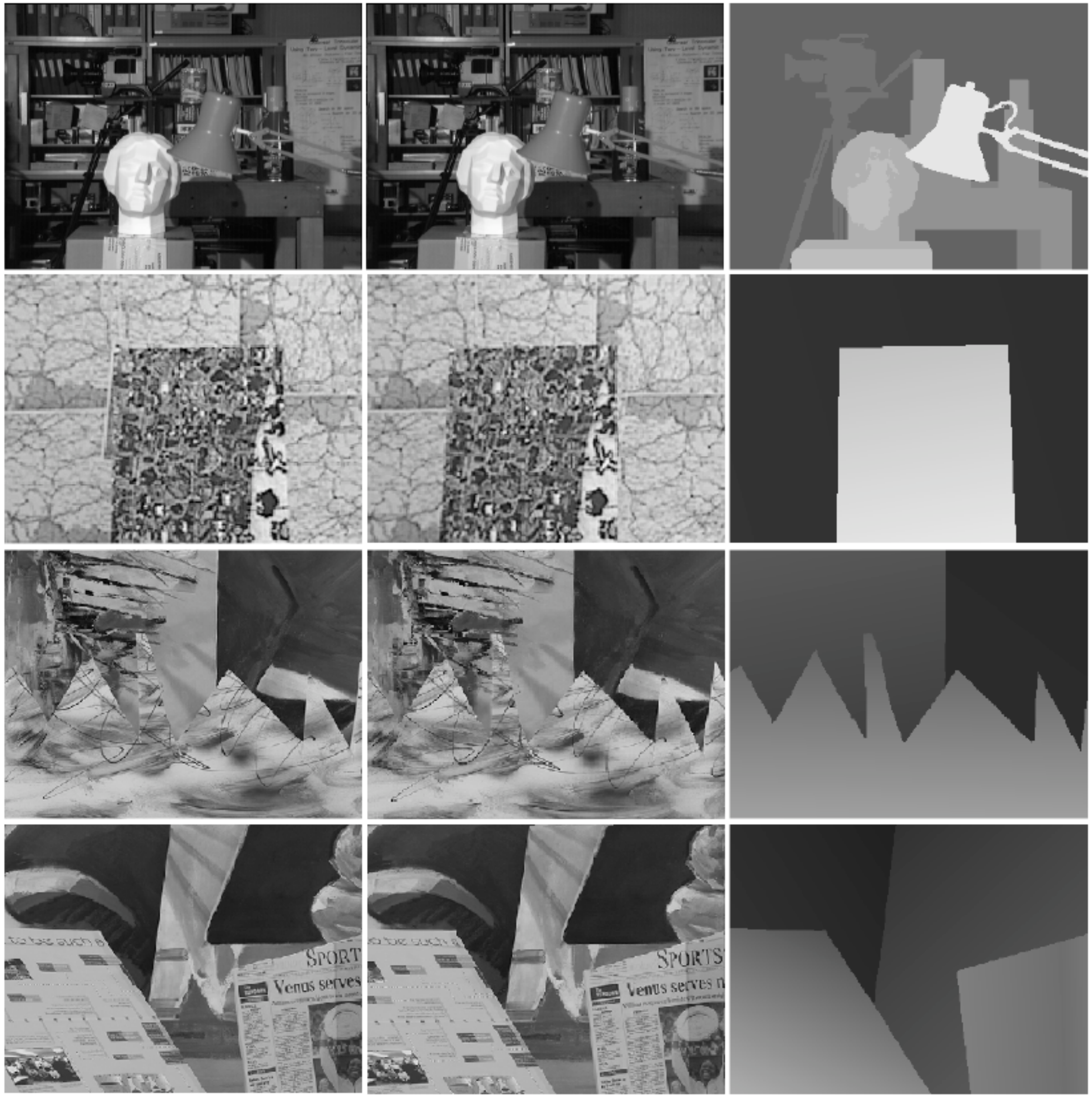
Original



BP



Edge BP



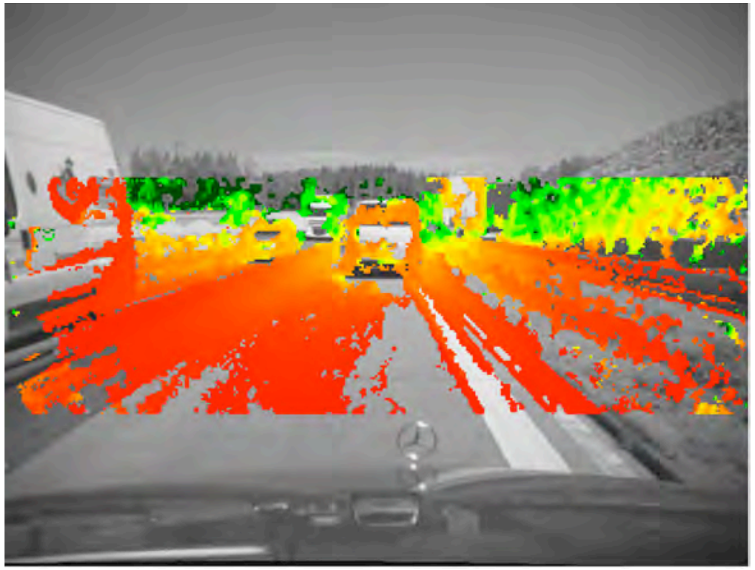
BP on Middlebury
stereo image pairs

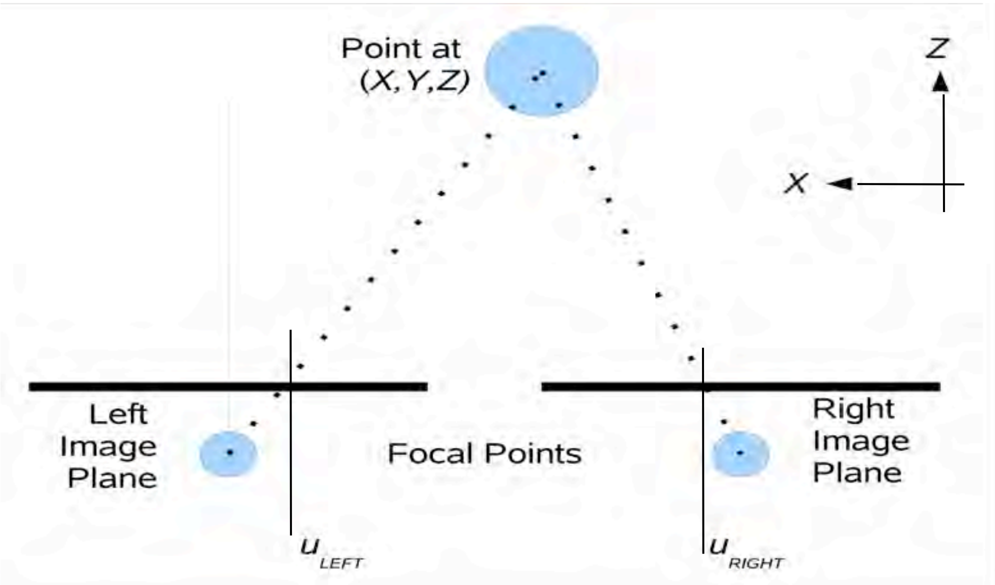
Table below: results get
slightly worse
for edge (Sobel) images
of those stereo pairs

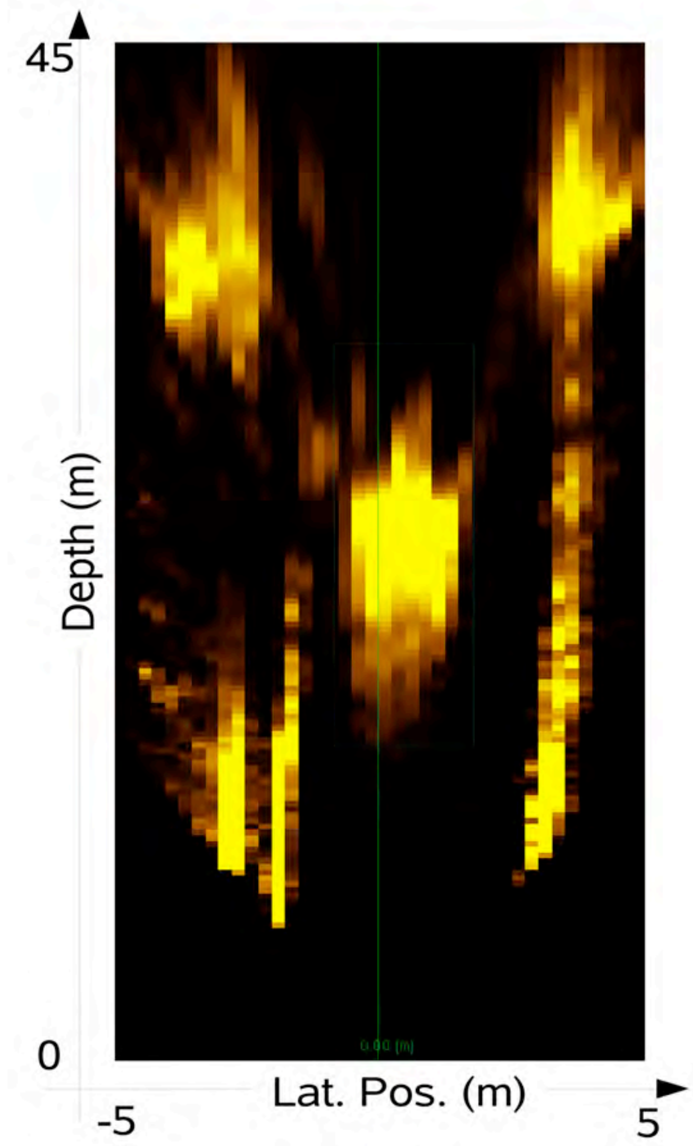
Image pair	Tsukuba	edge	Map	edge	Sawtooth	edge	Venus	edge
error	1.75	1.81	0.31	0.33	0.94	0.95	0.99	1.02

3. Improving occupancy grid by also using disparity rate in the Kalman filter

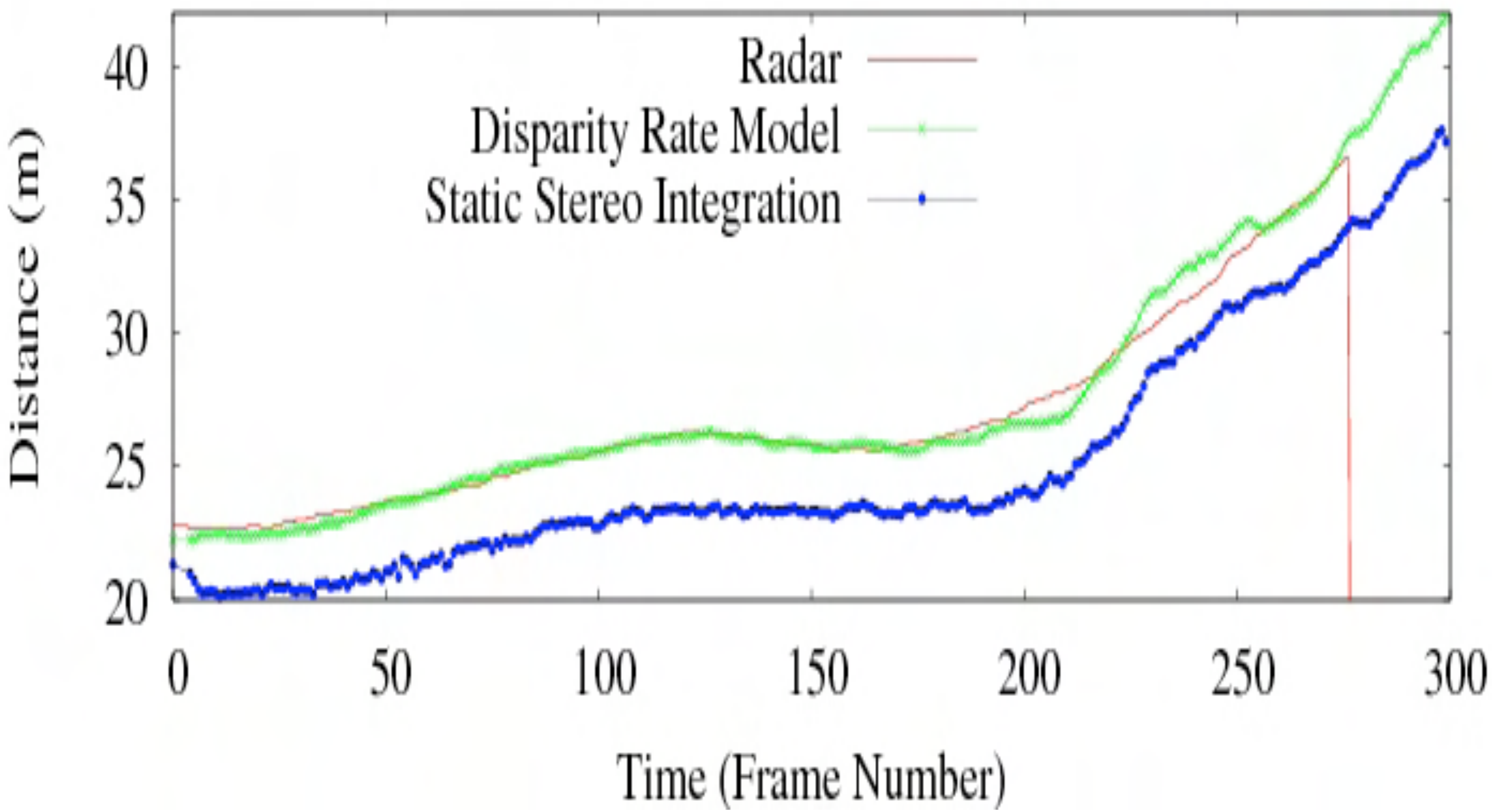
[Vaudrey, Badino, Gehrig, Robot Vision 2008]







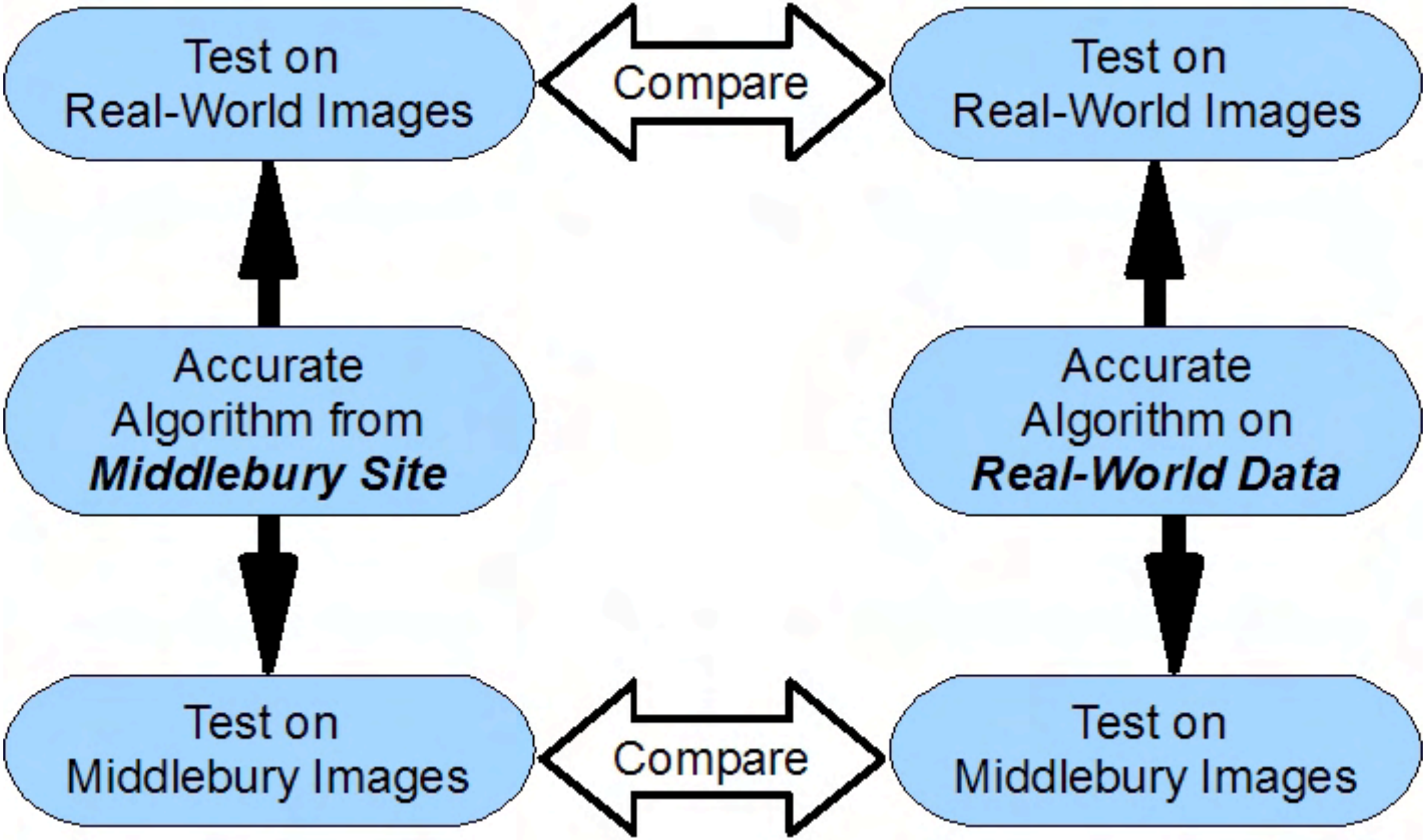

occ_grid_truck_vel_with_limg.mpg
Singen
Ehningen
60.00 [m]
40.00 [m]
20.00 [m]
-5.00 [m]
5.00 [m]
0:00:09
(pixel-wise) estimates
of speed of lead vehicle

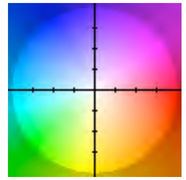


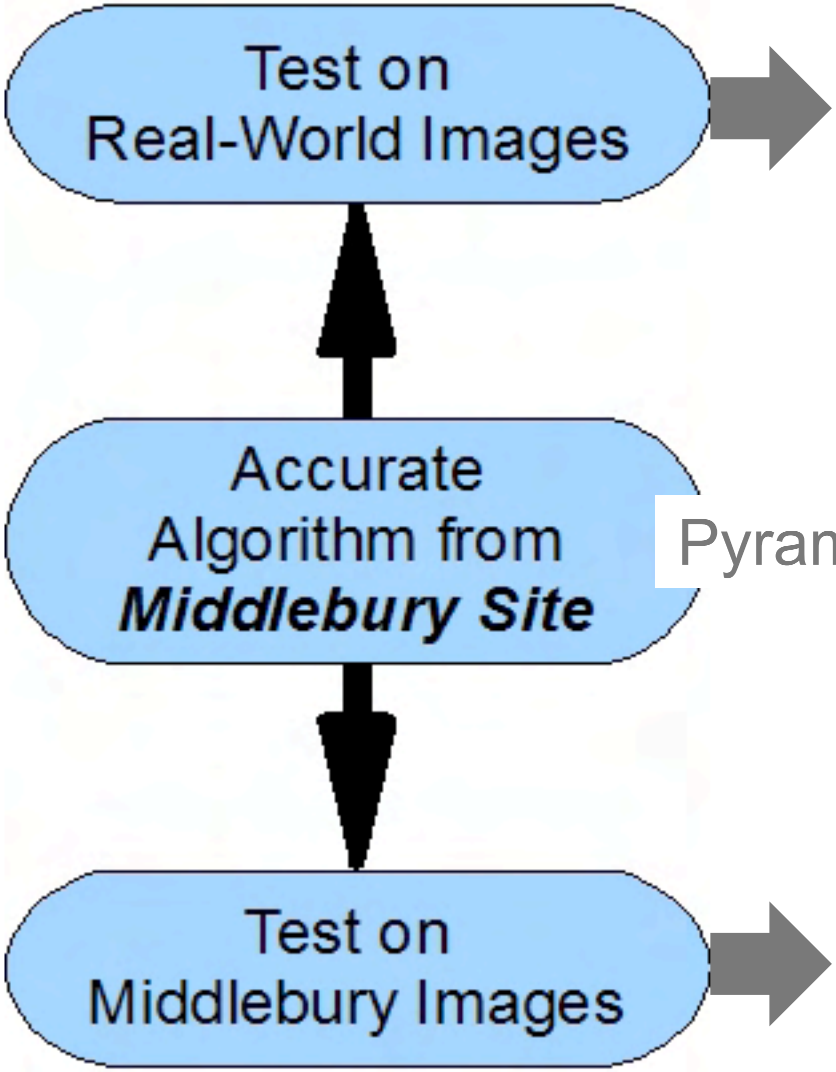
Incorporating disparity rate improved distance (speed) estimates to lead vehicle, compared to static stereo integration

[Vaudrey, Badino, Gehrig, Robot Vision 2008]

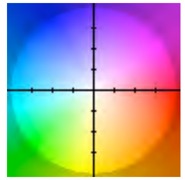
4. Evaluation for different types of input data

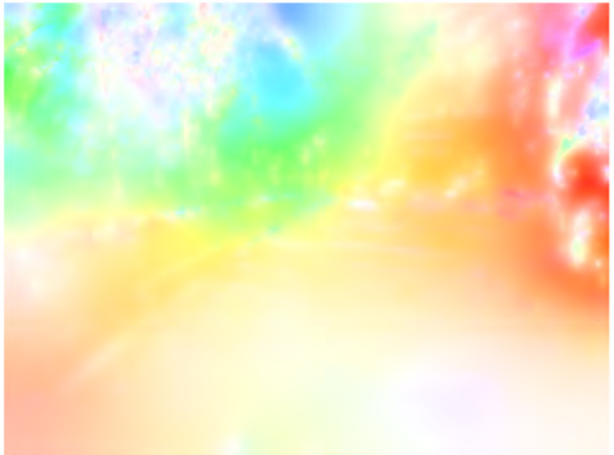




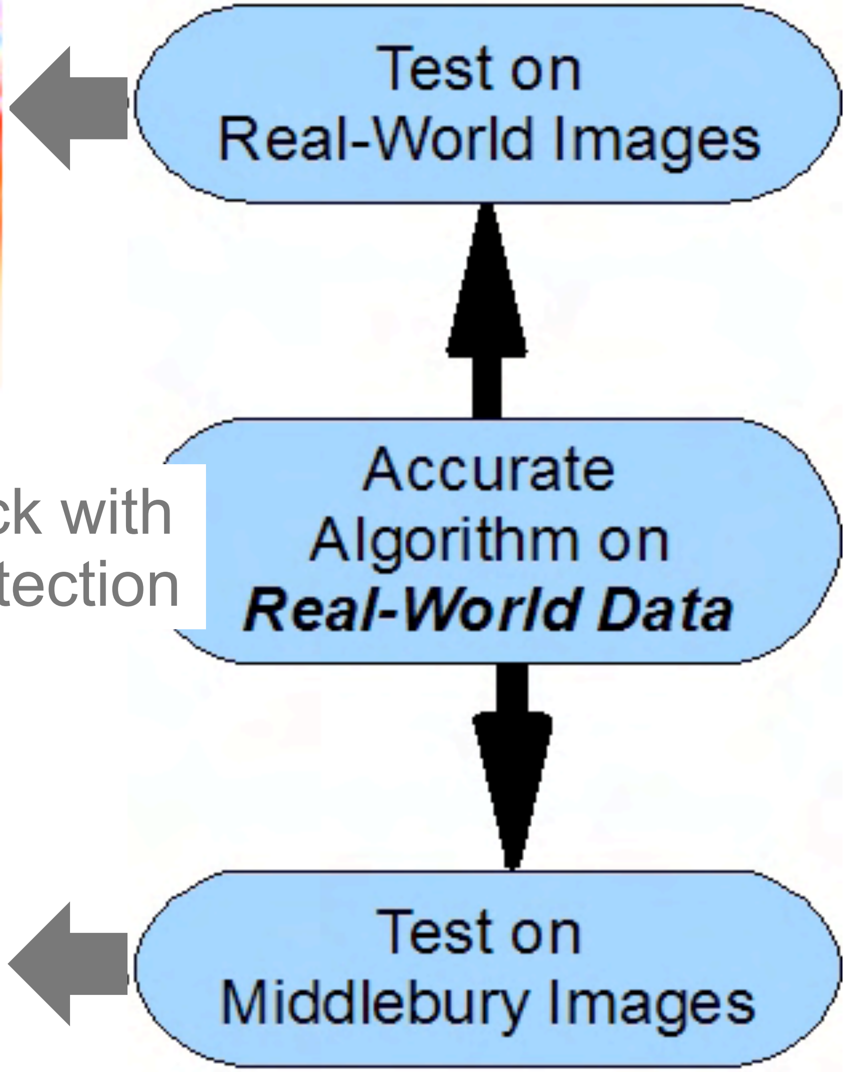


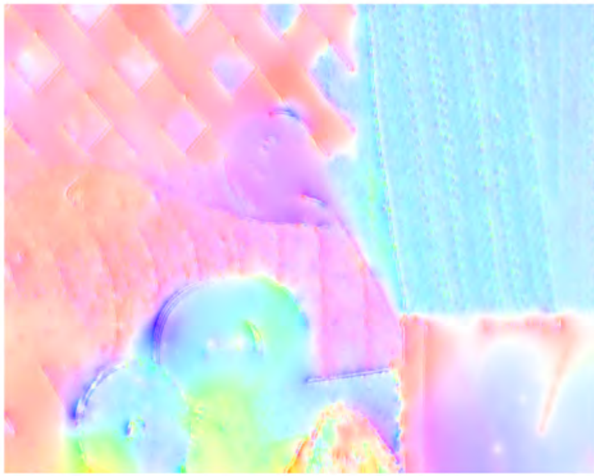
Pyramid Horn-Schunck





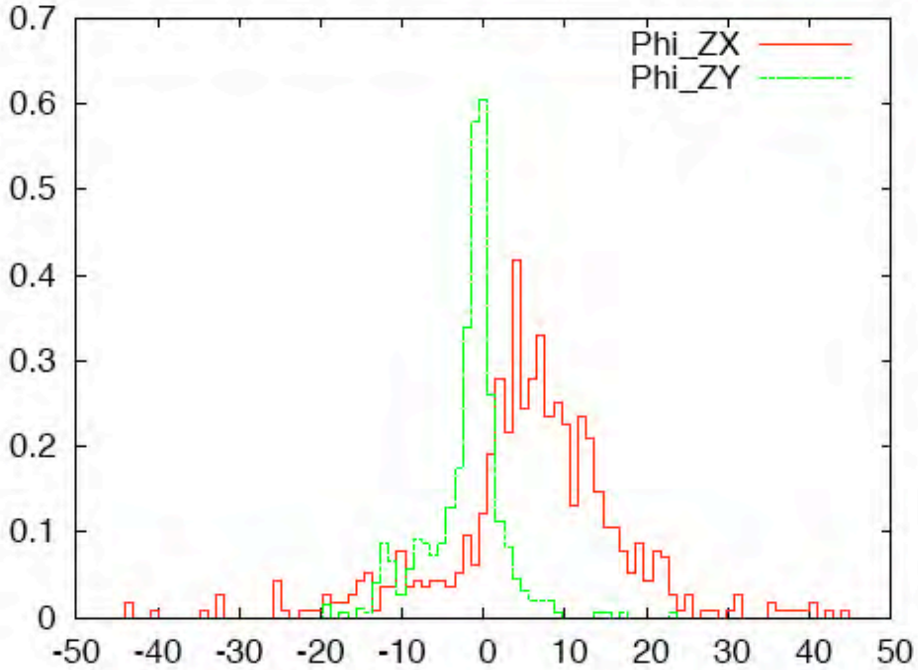
Pyramid Horn-Schunck with
Pre-Process Edge Detection



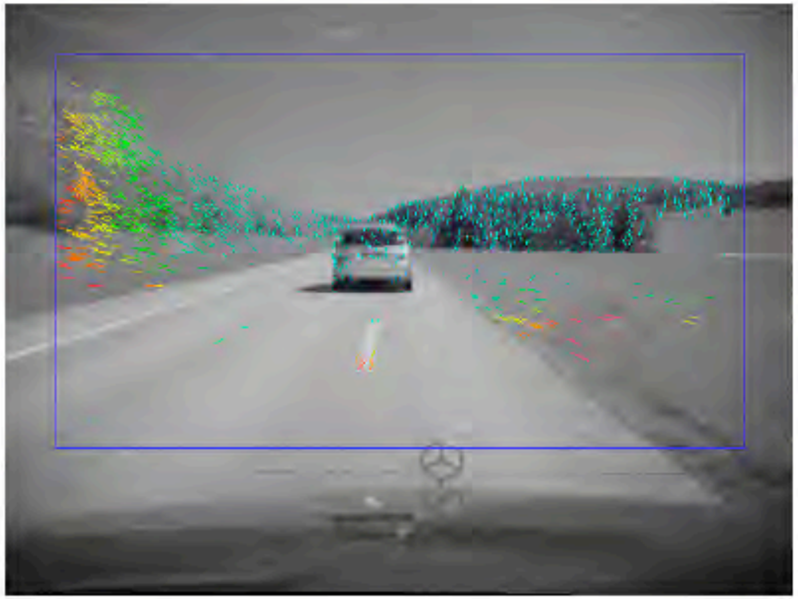


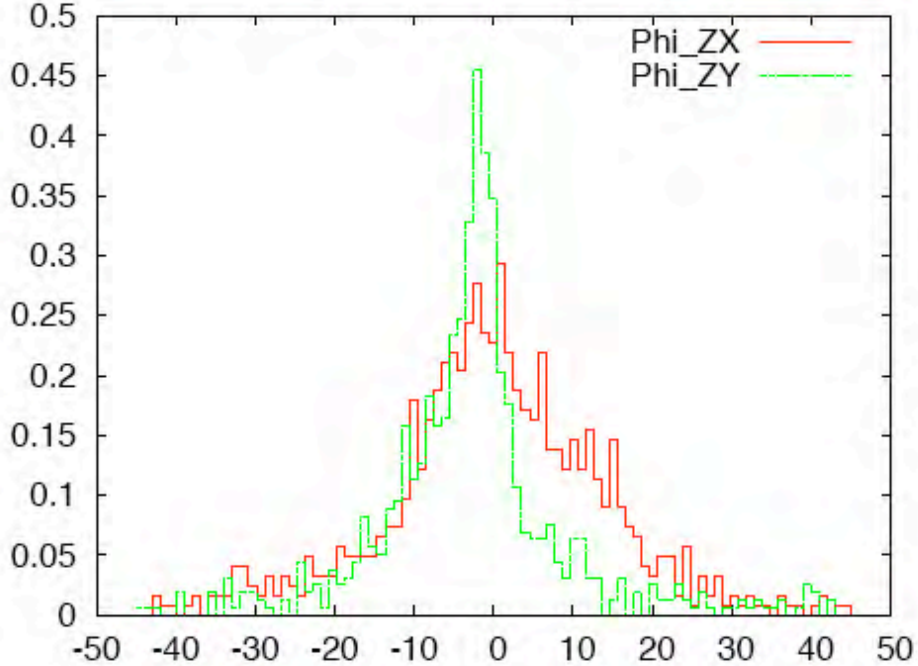
5. Evaluation of estimated navigation angles (roll, yaw)

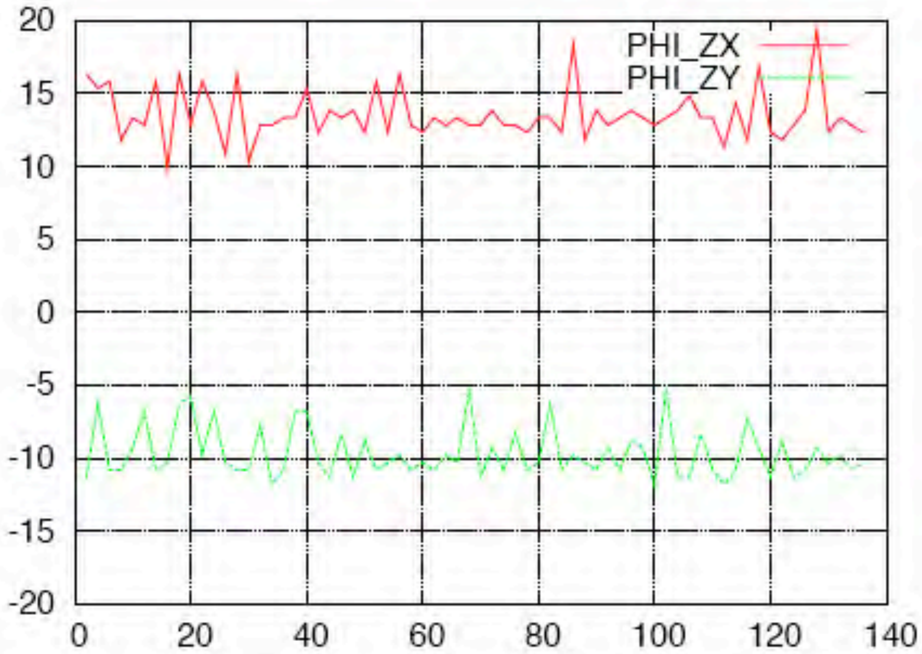
KLT tracker, max-response disks in scale space for 3D vector estimation



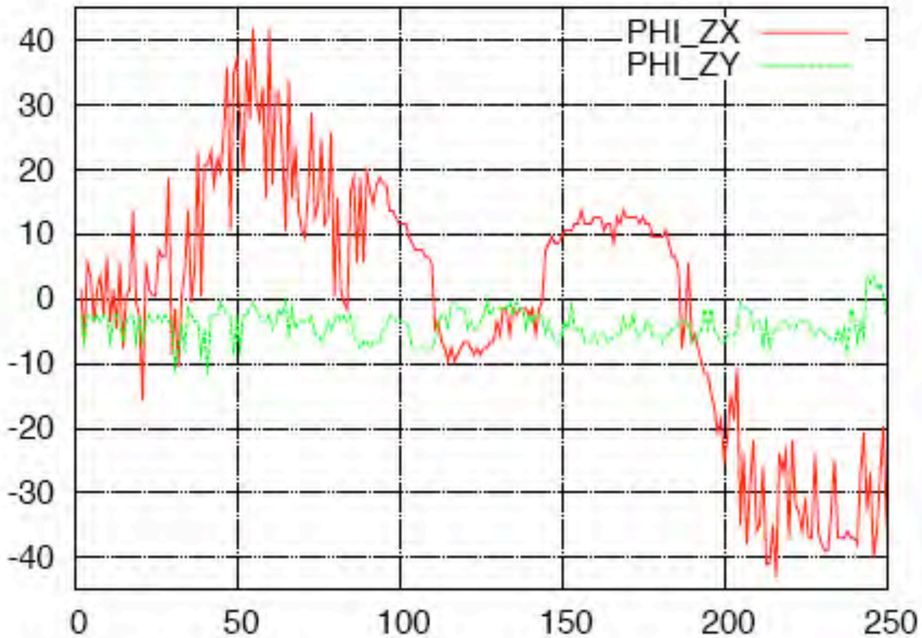
Mean navigation angles.







Mean navigation angles. Top: constant translation, ground truth: -10 and 12 degrees.



A few...

Concluding Remarks

- The ultimate goal: **predict** traffic situations about 3 seconds ahead (basically impossible, but how close can we go?)

- Current focus on accumulating (accurate) stereo and motion data may soon move on to **data analysis and scene understanding**: what is likeley to happen in the next 3 seconds?

- **Adapt** driving and car to current traffic situation

- For example, **prepare** the car for a collision if it appears to be immanent

- **Understand** complex traffic situations: moving or static people versus any other "less important object"

- Driver assistance scenes may be the most difficult for flow + stereo algorithms
- Algorithms performing well on one test data set, may not perform well on another set
- For long stereo sequences with high bits per pixel see test sets on www.mi.auckland.ac.nz/EISATS
- We **aim at optimizing** for safety, i.e. avoidance of any human accidental damage.
- **Needed:** model building of scenes, scenarios, or behaviorial patterns, inference of appropriate conclusions

The Next Evaluation? – Scene Flow