Discrete Driver Assistance

Reinhard Klette,

Ruyi Jiang, Sandino Morales and Tobi Vaudrey

The University of Auckland, New Zealand

Passive safety systems (seat belt, air bag, ABS, ESP) are designed to minimize the consequences when a vehicle is already involved in a dangerous situation.

Safety systems, that perceive the environment around them and act accordingly, are the next step to assure safe driving conditions. Cameras and computer vision offer potentially more flexibility for such active safety systems then using only radar, ultrasound, or LIDAR.

DAS

Driver Assistance Systems

.enpeda..

Environment Perception and Driver Assistance

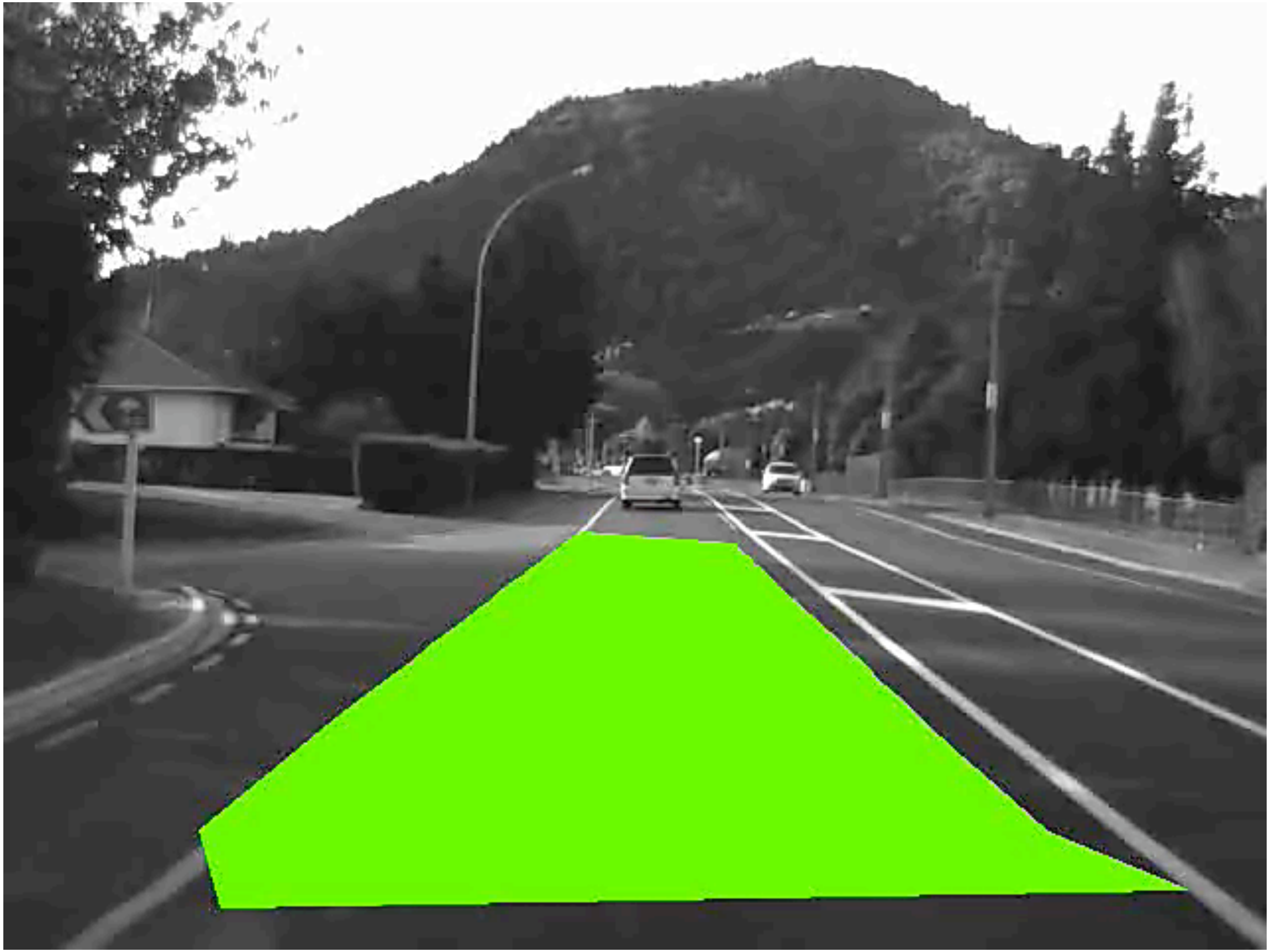
Active DAS are developed to

(i) predict traffic situations
(ii) adapt driving and car to current traffic situations
(iii) optimize for safety

Vision-based DAS applies one or multiple cameras for understanding the environment, to help achieve goals (i-iii).

Predicted space (corridor)

the car will drive in the next $\approx$2-3 seconds



The Ego-Vehicle

HAKA1

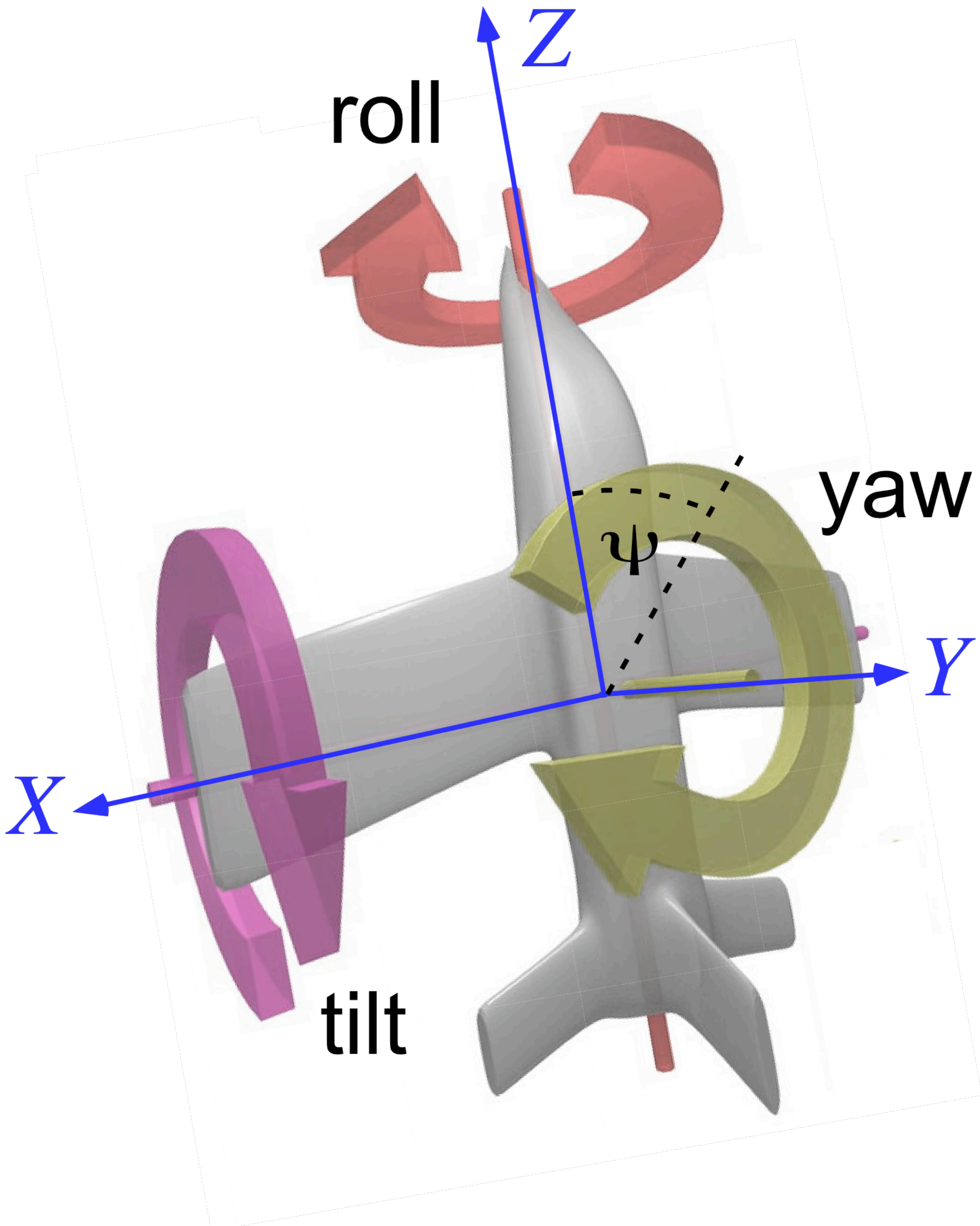
High Awareness Kinematic Automobile no. 1
test vehicle in the .*enpeda*.. project

Yaw ψ

steering angle

Tilt and roll

often `disturbing' ego-motion components



ego-vehicle

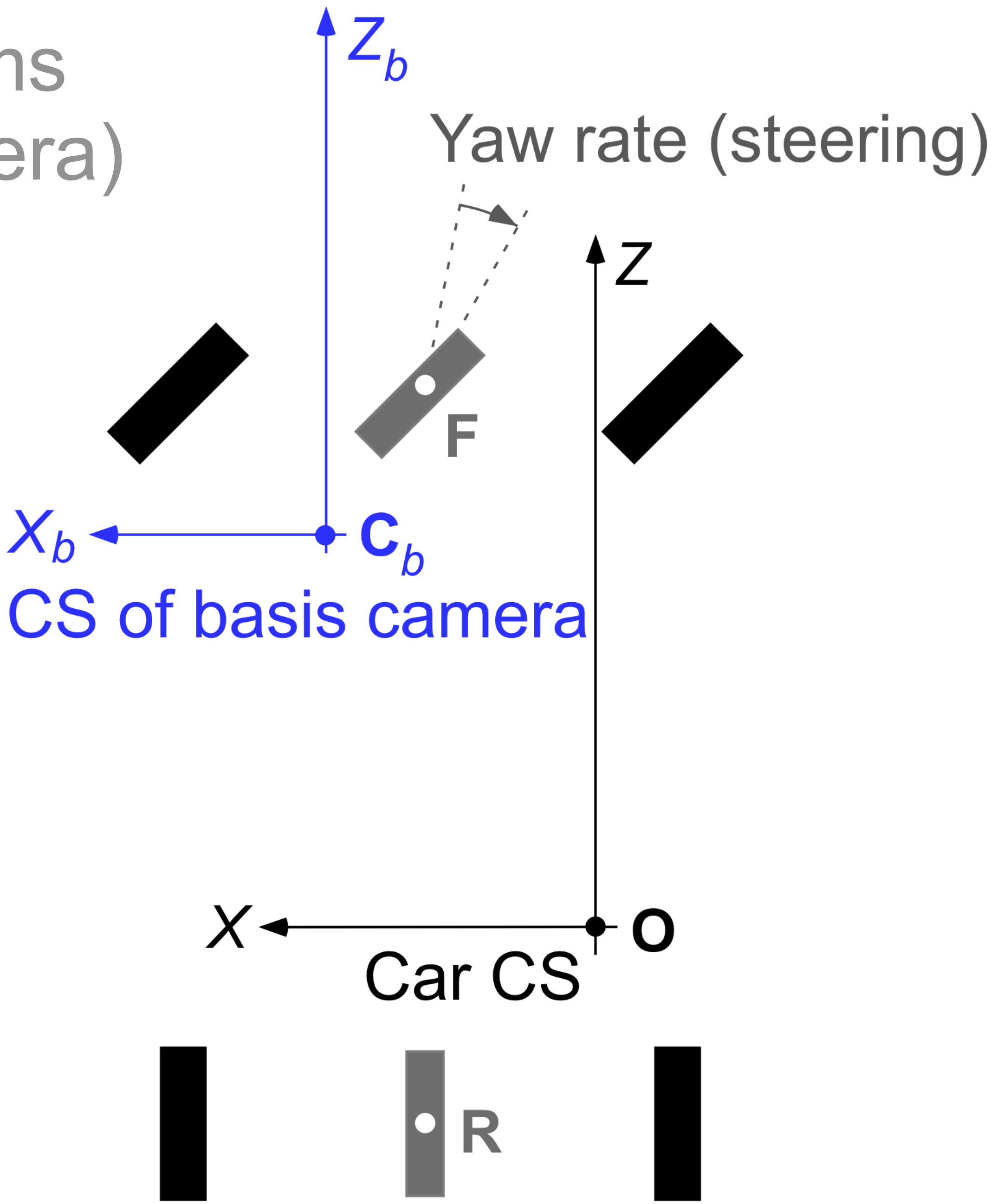
the car where the system is operating in

ego-motion

changes in yaw, tilt, roll and velocity

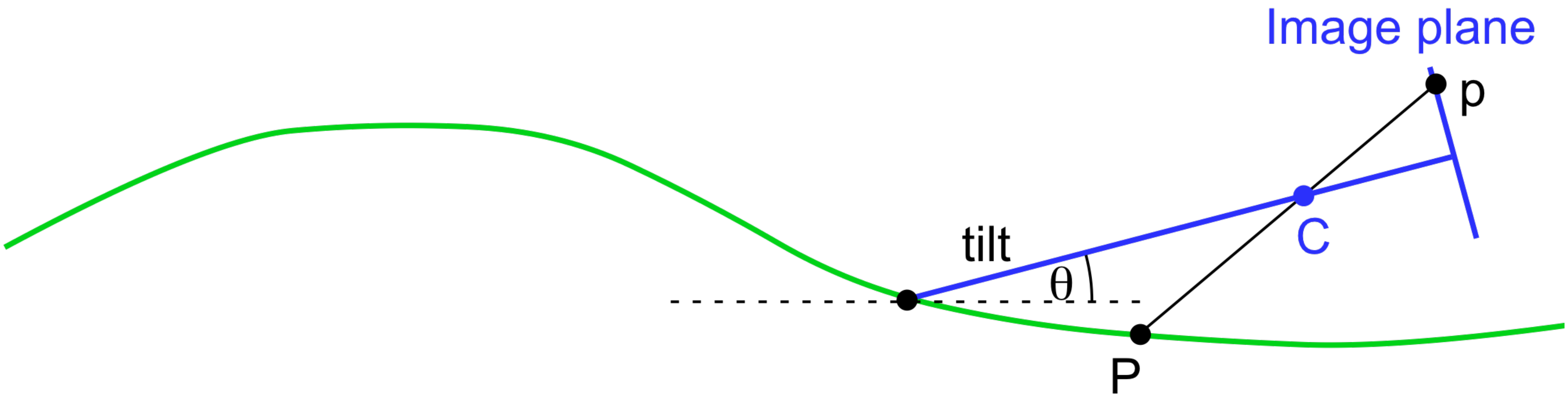
(on ground manifold)

Coordinate systems (car and one camera)



F and **R** define the *bicycle model* of the car

Camera and ground manifold



The Image Data

multi-camera, gray-level

currently ≈ 640 x 480 10 bit 25 Hz

e.g., up to 7 cameras in or on top of HAKA1

If human vision then it would be

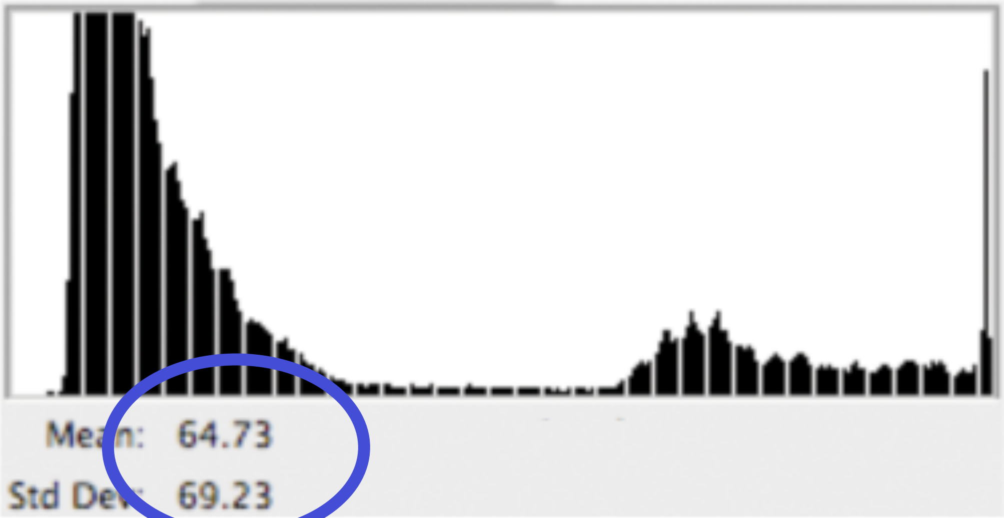
allchromasia, tunnel vision, myopia (distance blur), ...

Real-world data

noise, brightness differences, lighting artifacts, ...

any time of day, weather, traffic situation,...

Gray-level histograms



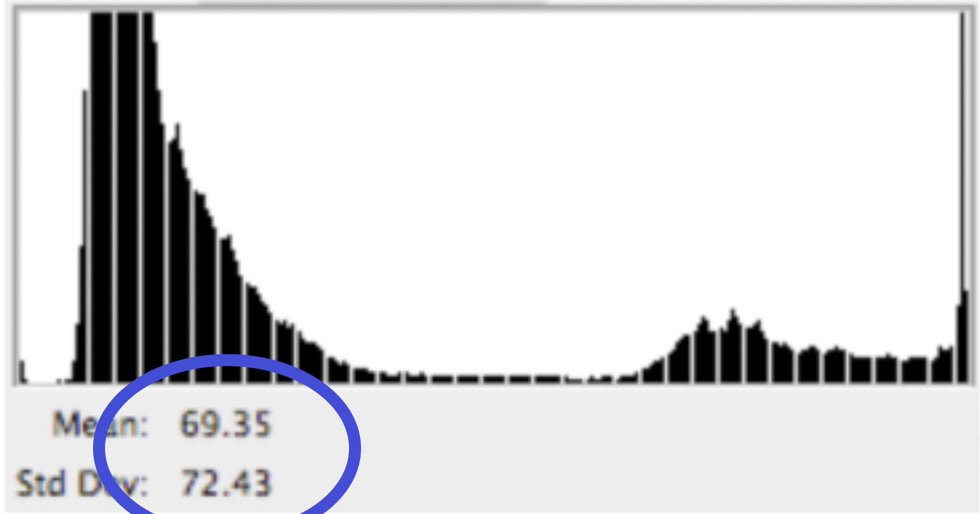


Photo-consitency is often violated

Two corresponding 297 x 208 windows : occlusions

Example of a more difficult situation

Dense night traffic

no rain, no snow,... – it could be much worse

Size of input space $\geq 614{,}400^{1024}$ stero pairs

and .....

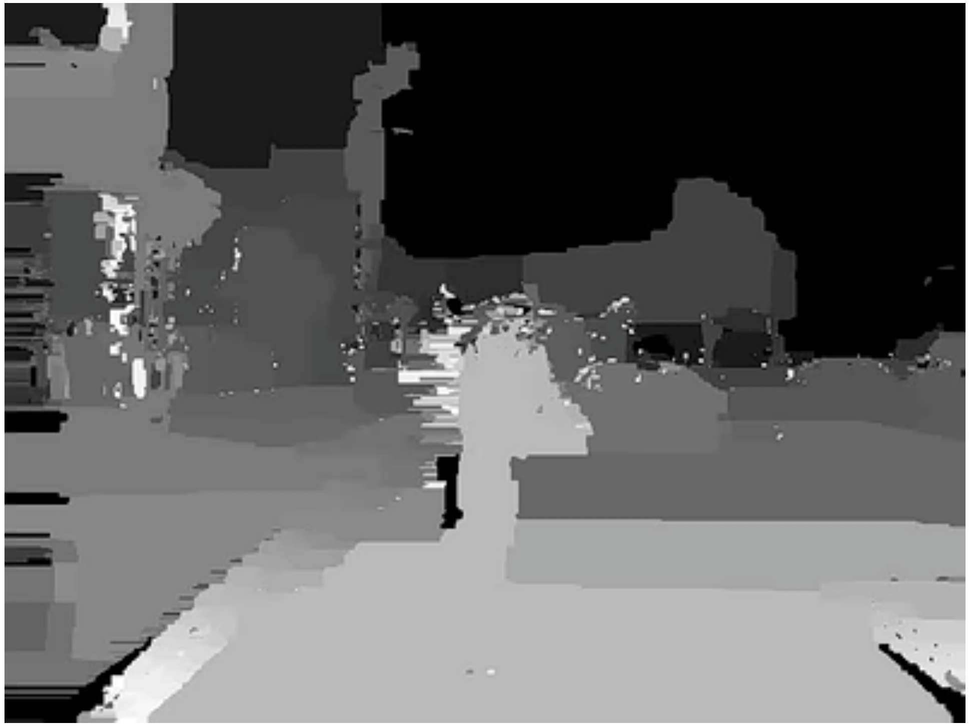
noise (e.g., photo-inconsistency) - no way to aim at a general optimization !

Situation: brightness diff.

Winner: belief-propagation (BP) stereo, preprocessed (residual, 3x3 mean, 40 iterations) sequence

Situation: close objects

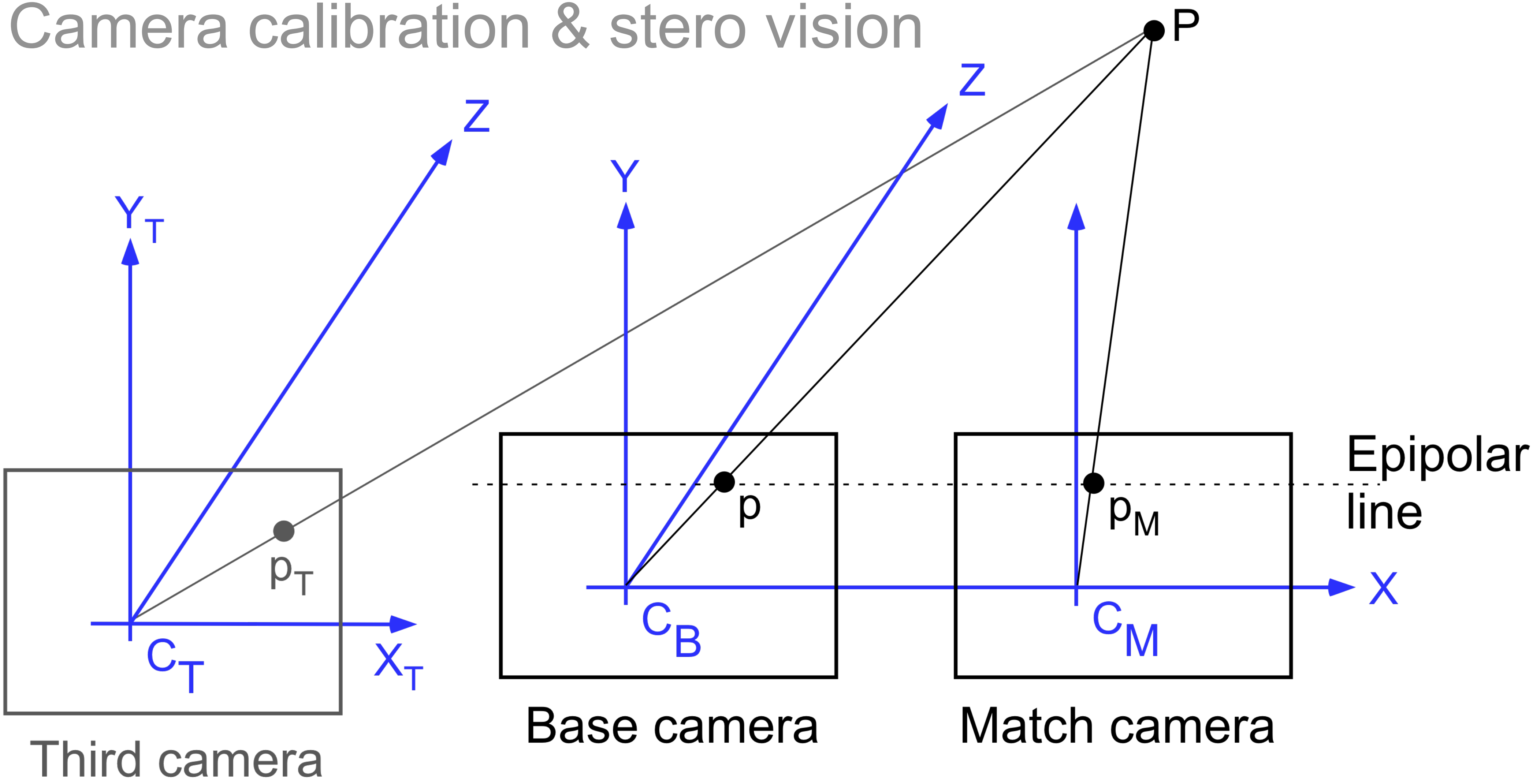
Winner: graph-cut stereo (GC) on preprocessed (residual, 3x3 mean, 40 iterations) sequence



enpeda..

Stereo Matching (mainly discrete math.)

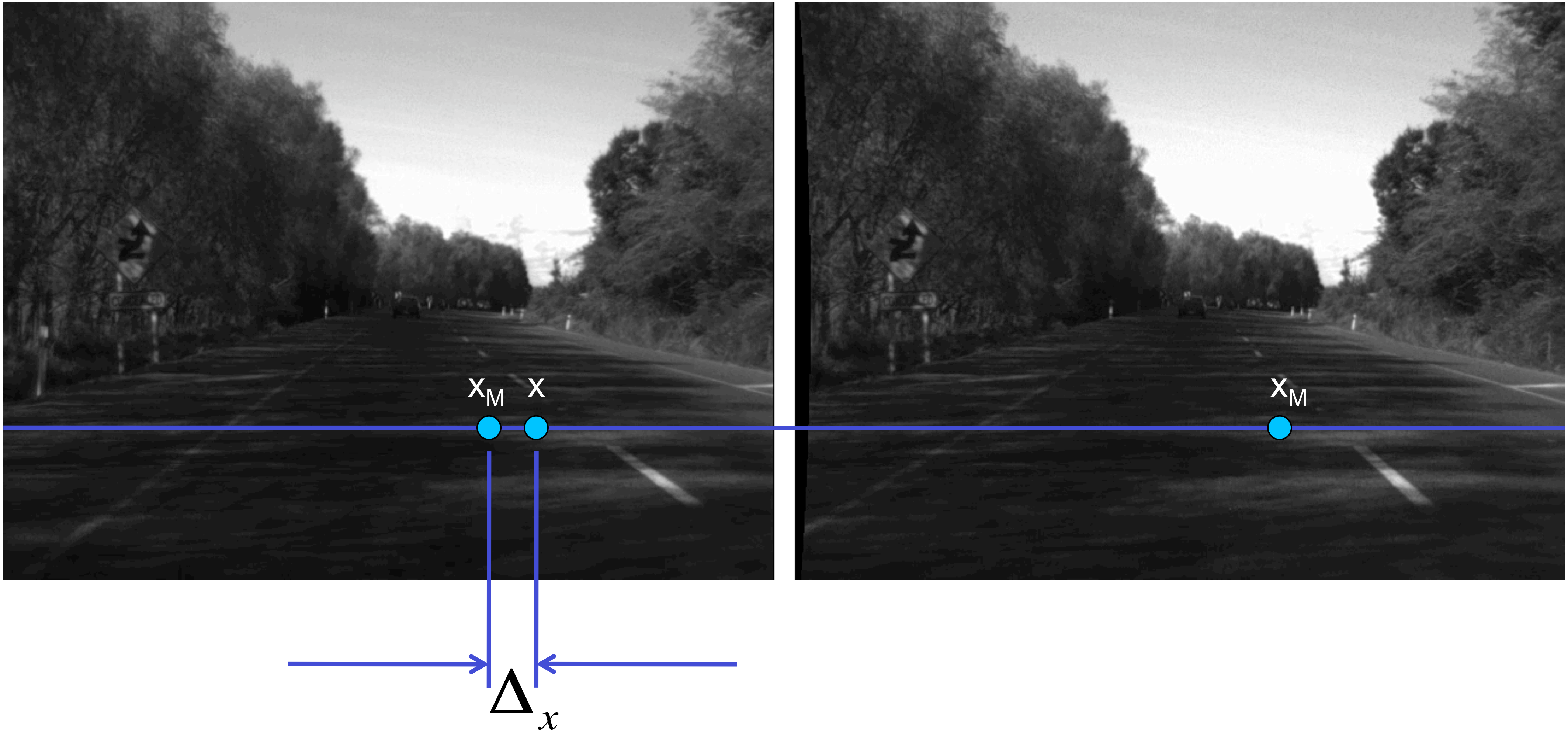
Camera calibration & stero vision



see, e.g.,
[J.-Y. Bouguet. Calibration Toolbox]

Stero matching

is a 1D (along epipolar line) correspondence problem



disparity - zero at infinity, finite range of disparities

Prediction error analysis for stereo triples [R. Szeliski, 1999]

calculate disparities for base and match sequence

warp base intensities into third camera view,

based on calculated disparities

compare those virtual images with third images

(i.e., images of the third camera)

using the normalized cross-correlation measure

$$N(t) = \frac{1}{|\Omega_t|} \sum_{p \in \Omega_t} \frac{\left[T_t(p) - \mu_{T,t}\right]\left[V_t(p) - \mu_{V,t}\right]}{\sigma_{T,t}\sigma_{V,t}}$$

Three cameras in HAKA1

Third | Base (left) | Match (right)

40 cm left of base camera | about 30 cm apart from each other

all three on one bar behind windscreen

left and right images are rectified for stereo matching

110 stereo frames > thus 110 NCC values

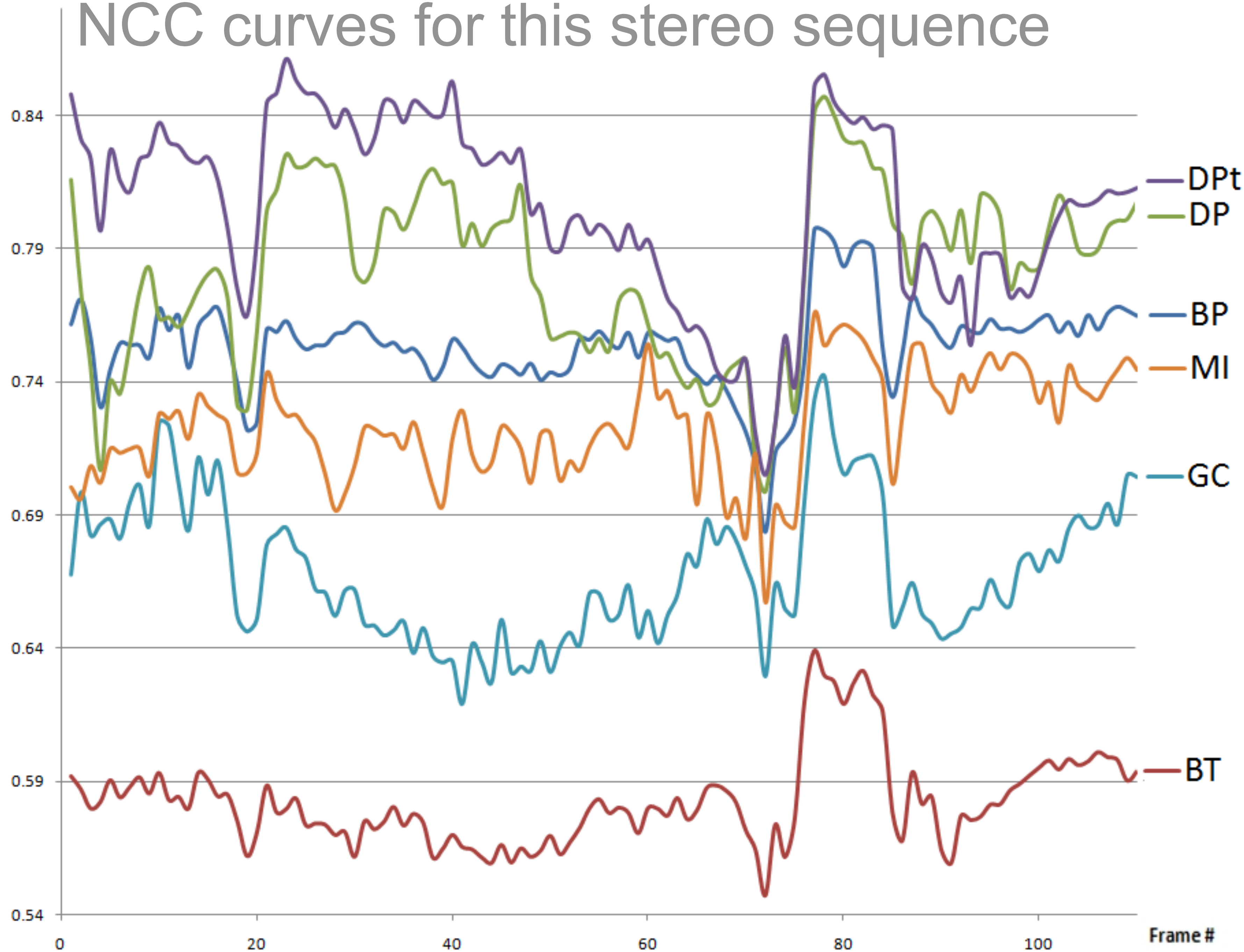
Virtual view

using Dynamic Programming (DP)

Third view

(reflections on screen cannot be predicted – but are constant for all comparisons)

NCC curves for this stereo sequence



Note: changes in rankings along the sequence

Disparity Calculation

Disparity calculation as a labeling problem

minimization of an error function (known from MRFs)

$$E(\Delta) = \sum_{p \in \Omega} \left(D_p(\Delta_p) + \sum_{q \in A(p)} C(\Delta_p, \Delta_q) \right)$$

labeling Δ for all pixels p in Ω

data term D e.g. $| L(x,y) - R(x-\Delta_p, y) |$

continuity term C between adjacent pixels; often

$$C\left(\left|\Delta_p - \Delta_q\right|\right)$$

[V.Kolmogorov & R.Zabih, 2002] state that a minimization of *E* is NP-hard (but without giving any proof)

Data term commonly assumes photo-consistency
(which does not hold for DAS image sequences)

Depth discontinuities at object or occlusion edges should not disappear due to *C*

BTW: a continuity term which enforces over-smoothing could allow a trivial solution; and this is not NP-hard

Dominant paradigms for energy optimization

Scanline optimization stereo matching

Dynamic programming stereo

single scanline (epipolar line) in one direction

Semi-global matching

multiple scanlines (DSLs) in both directions

Belief propagation stereo matching

general BP paradigm applied to stereo vision

Graph-cut stereo matching

general GC (of combinatorial optimization) applied to computer vision

[see paper in ISMM proceedings for outlines and references]

Dynamic programming stereo matching (DP)

Let $E(\Delta) = E_M(\Delta)$, and at stage $m \leq M$, optimize

$$E_m(\Delta) = \sum_{x=1}^{m} \left(D_x(\Delta_x) + \sum_{\hat{x} \in A(x)} C(\Delta_x, \Delta_{\hat{x}}) \right)$$

DP with temporal propagation (DPt)

at p in frame t, let disparity at p in frame t-1 contribute with some percentage (e.g., as 20%)

Virtual view for DP and situation `close objects'

Some streaking effects

Semi-global stero matching (SGM)

For each pixel, optimize energy along digital rays starting at this pixel. Uniform weights for all rays. Possibly add further cost functions (e.g., for surface curvature, see [S. Hermann, R. Klette, E. Destefanis, 2009]).

Data term: common functions are, for example, mutual information (SGM MI) using an entropy measure, or Birchfield-Tomasi (SGM BT) which is time-efficient and considers also interpolated values at subpixel positions.

Depth map for SGM MI and situation `night traffic’

use of eight rays

Belief-propagation stereo matching

General paradigm: message passing in a graph

Here in general: 4-adjaceny grid

Messages: `support' between adjacent pixels for particular labels (disparities)

Continuity function $C\left(\left|\Delta_p - \Delta_q\right|\right)$

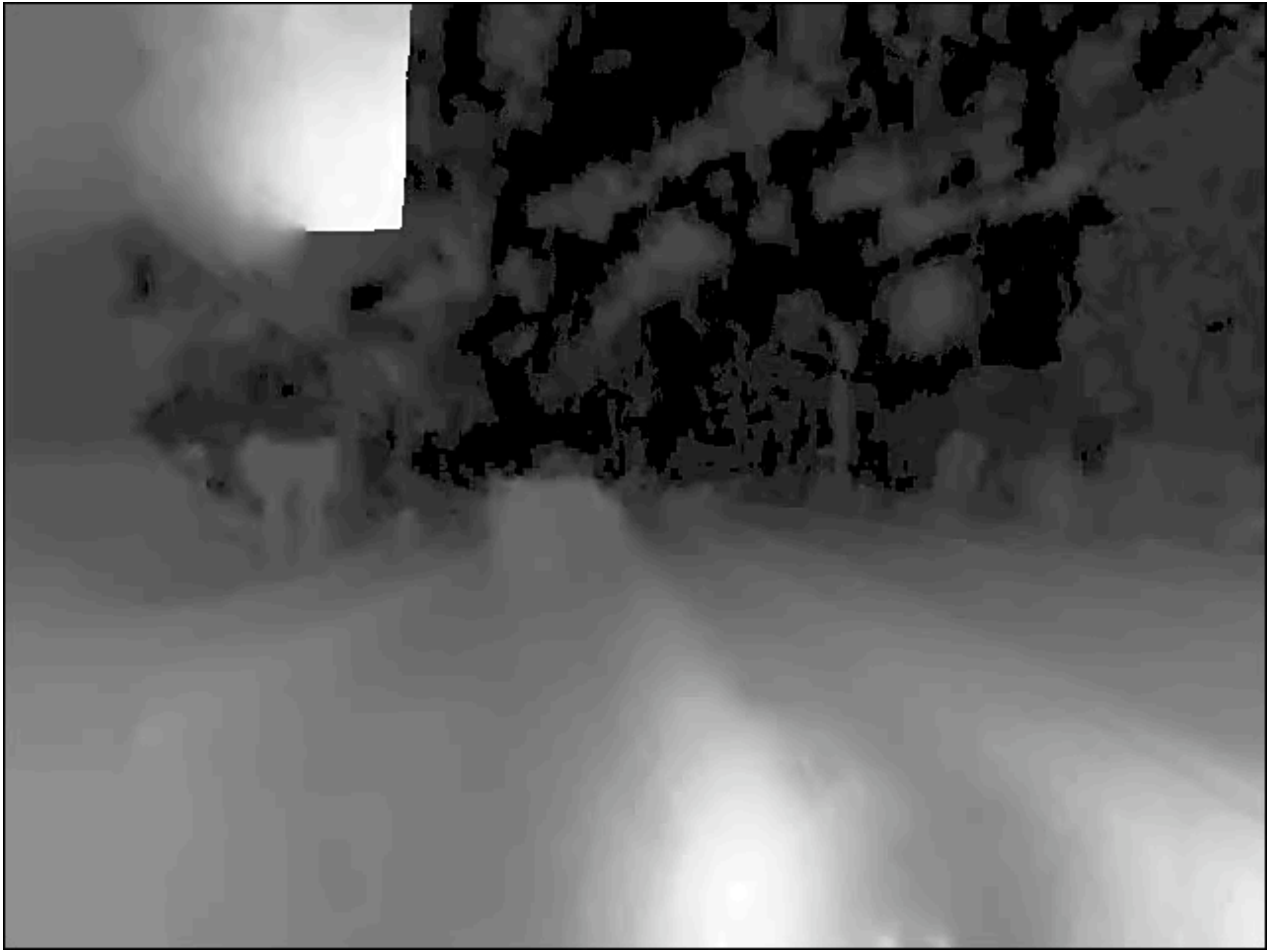
- (simple) binary Potts-model
- truncated linear function
- truncated quadratic function

use of lower envelope algorithm (as designed for EDT)

quality improvement and speed-up by hierarchical implementation

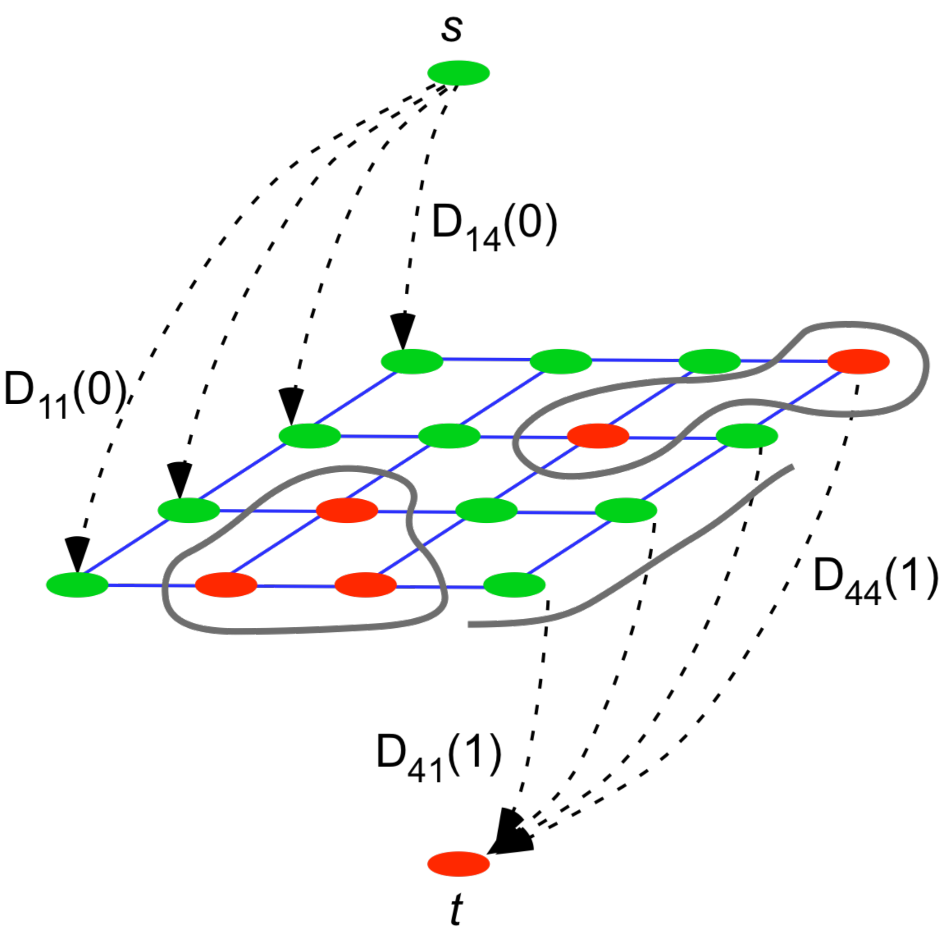
Depth map for BP and situation `default conditions'

quadratic cost function, 6 layers in the hierarchy

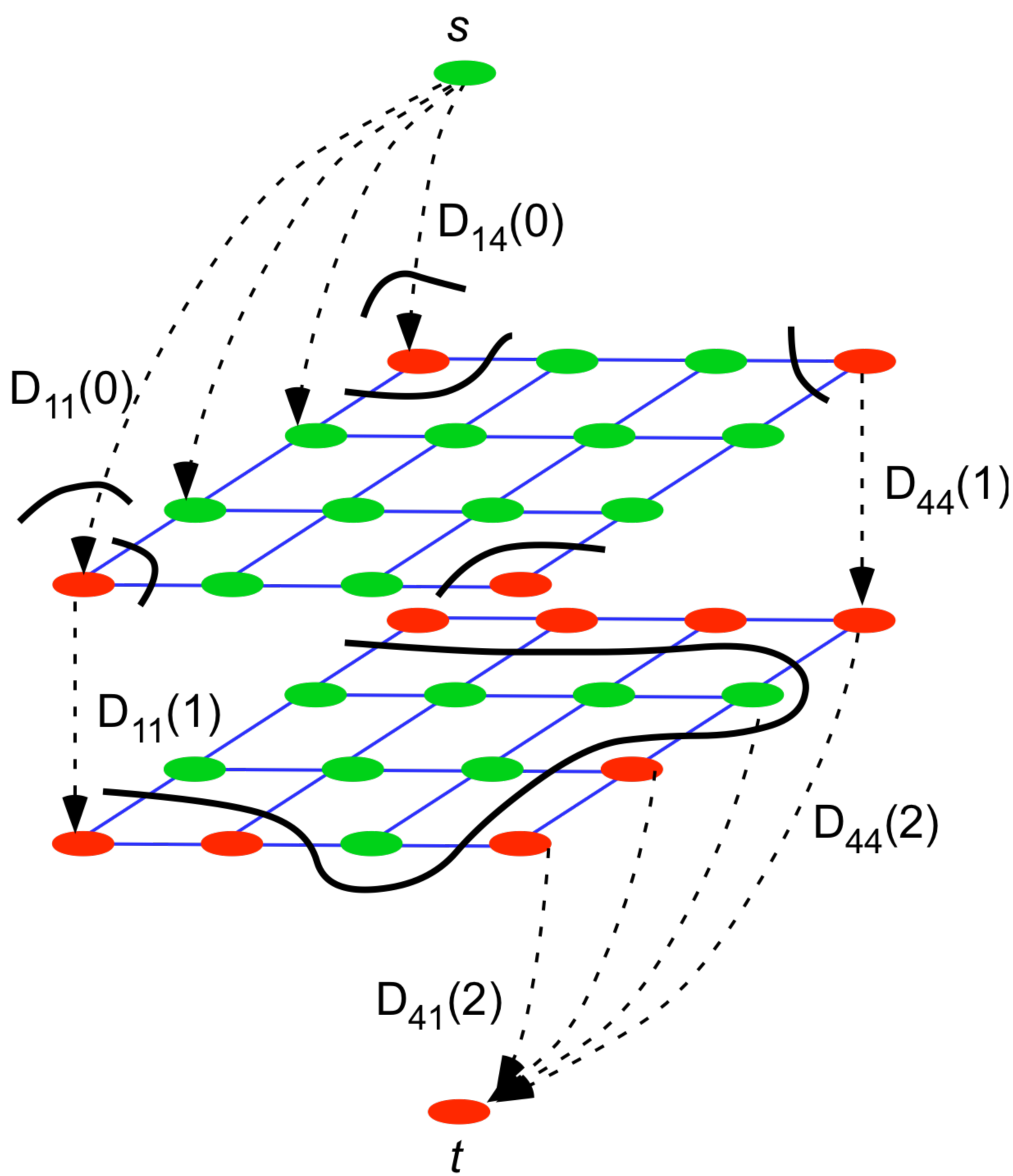


Graph-cut stereo matching

Min-cut between source *s* and target *t*



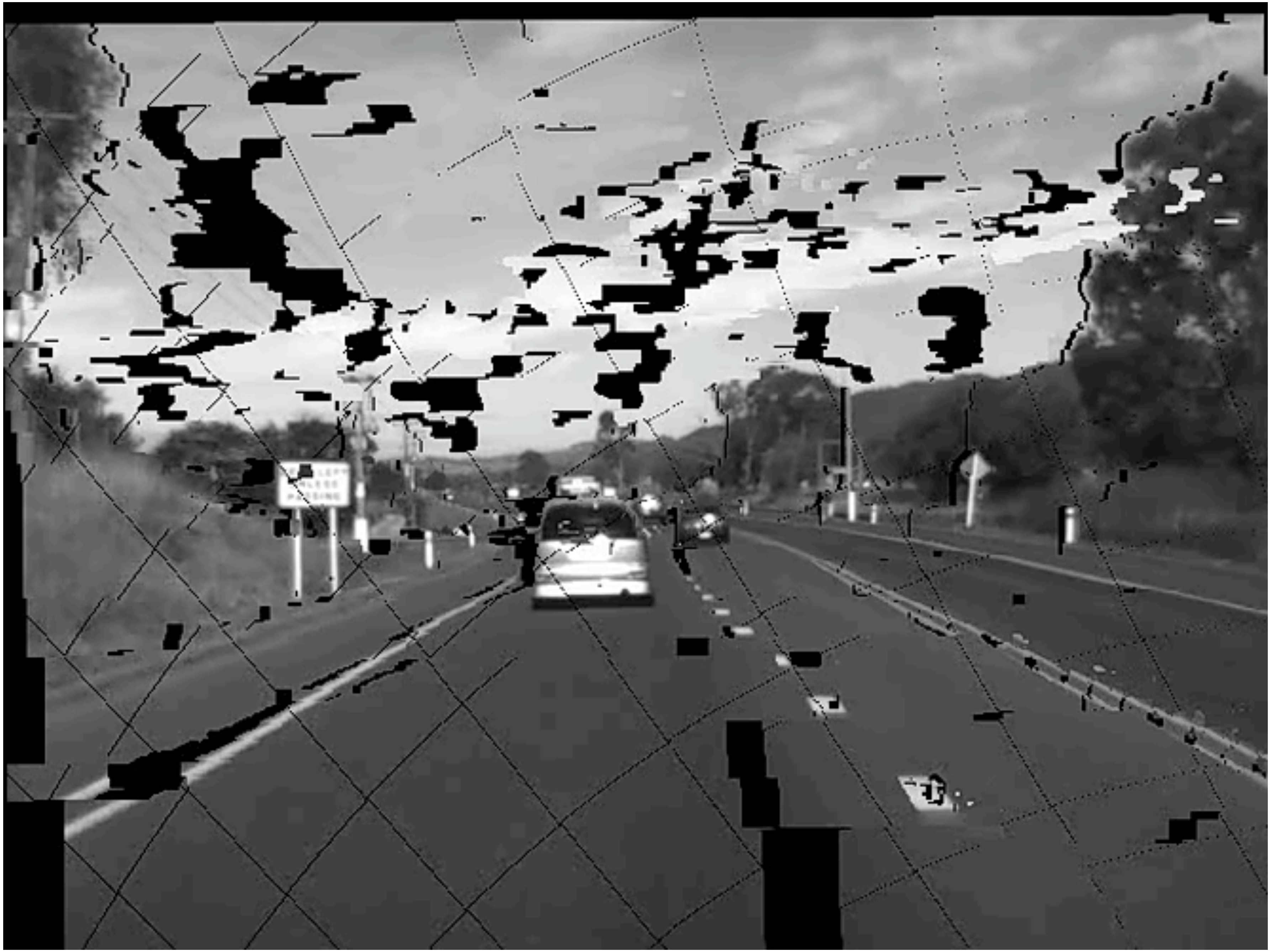
Two disparities 0 and 1



Three disparities 0, 1 and 2

Virtual view for GC and situation `default conditions'

creates somehow flat zones in depth map, and isothetic regions in V



Comments

No global winner, situations define local winners

Preprocessing of sequences sometimes useful (esp. residual images)

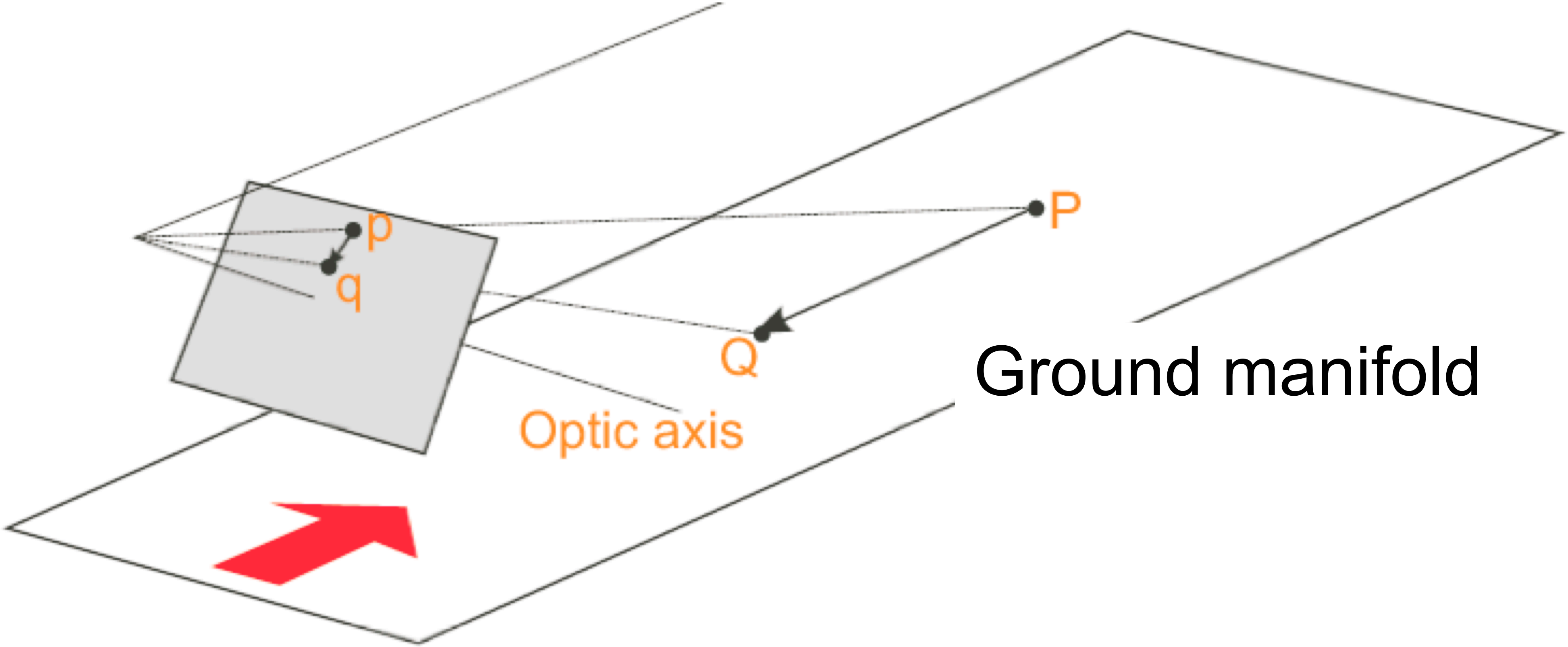
Methods considered to be time-inefficient a few years ago are now candidates for real-time (25 Hz) stereo processing

Specialized hardware or processors are now common

The other early-vision correspondence subject:

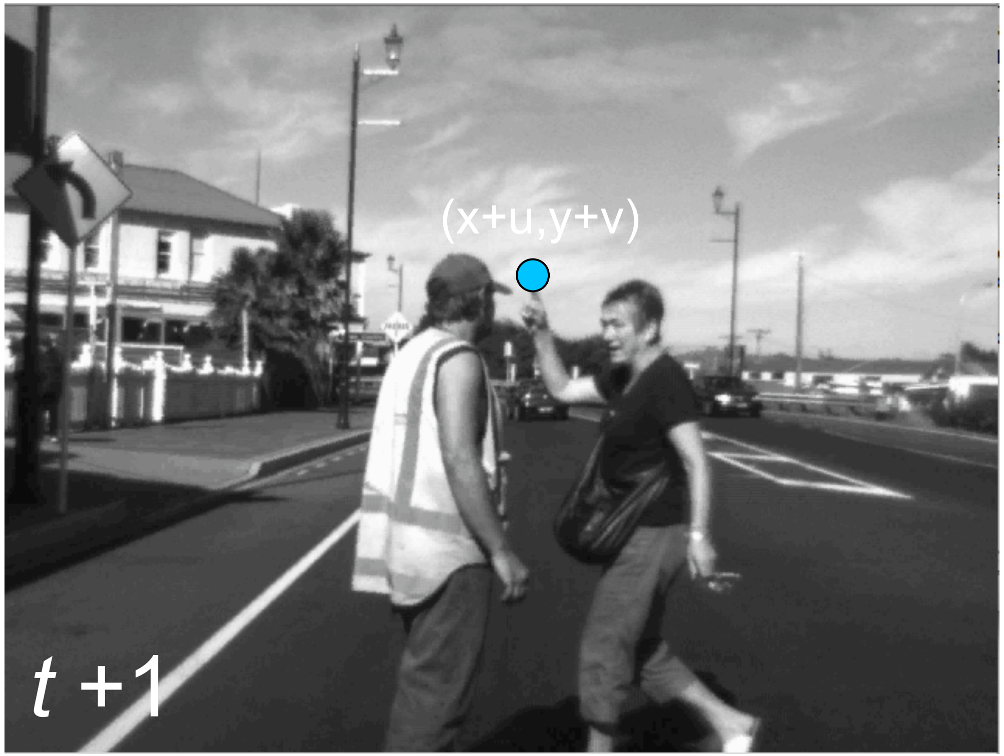
Motion Analysis (use of continuous models)

Motion analysis



Motion analysis

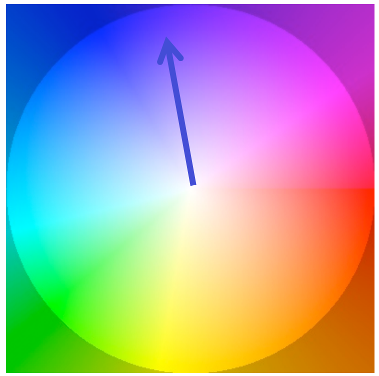
is a 2D (in image plane) correspondence problem



at 25 Hz

(u,v)

Color key

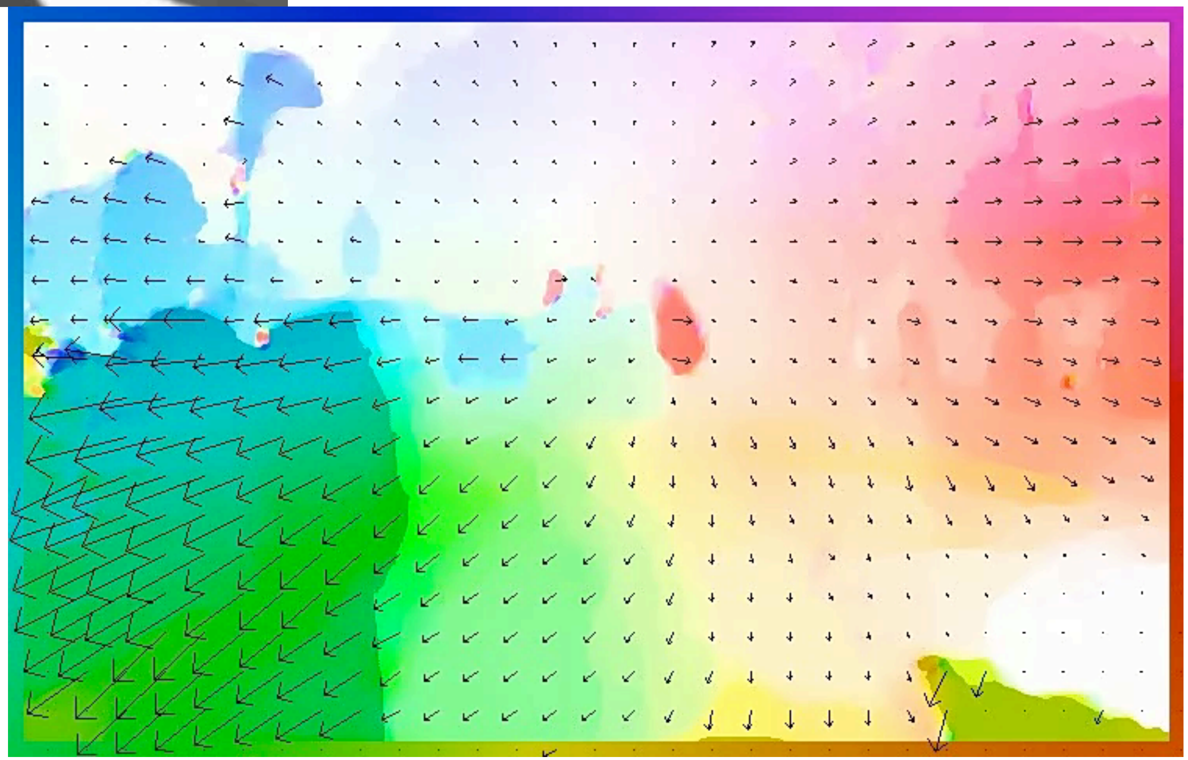


optic flow – aims at subpixel accuracy

Situation
default conditions

Optic flow technique
TV L^1

[C.Zach,T.Pock, H.Bischof 2007]



Some early interaction between optic flow techniques (often TV) and stereo matching

[N. Slesareva, A. Bruhn, J. Weickert DAGM 2005], ...

Combining various techniques for

Lane Detection & Tracking

Example of a special modul of vision-based DAS

For intelligent cruise control or road modeling
Lane departure warning in cars since 1990s

[McCall, 2006: A complete review on lane detection methods]

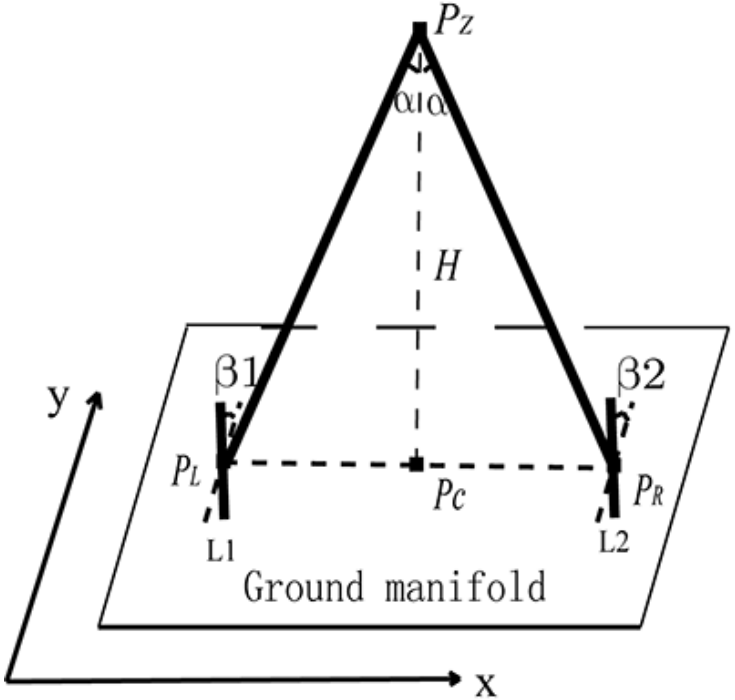
Various road conditions
Plenty of road models
(parabolic, hyperbolic, linear, spline, clothoid, …)

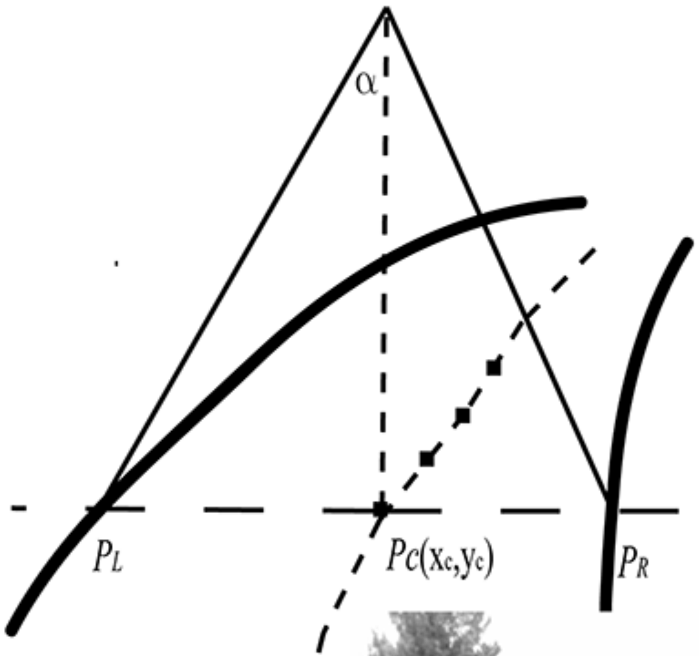
Sensor fusion
(camera, internal vehicle state, GPS, laser, radar)
Tracking methods (Kalman filter, particle filter)

Ongoing challenges:
robustness and generalization (all kinds of situations)

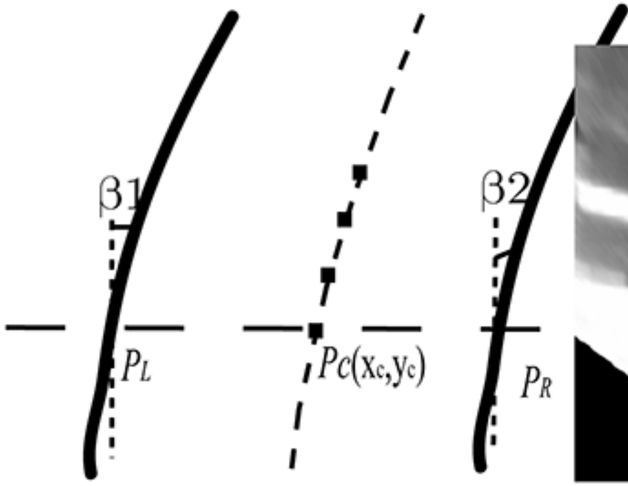
Curved & (sometimes) unmarked roads

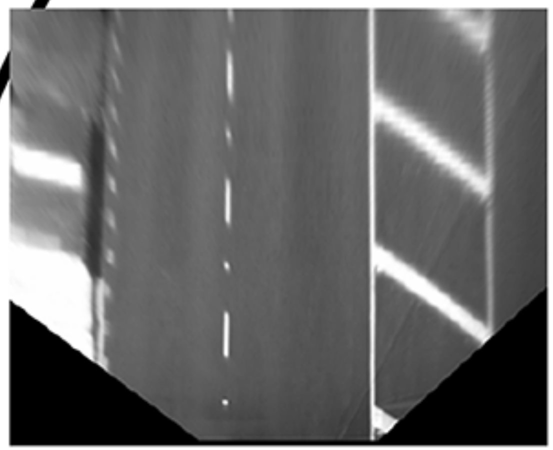
New (weak) lane model $(\alpha, \beta_1, \beta_2, x_c, y_c)$





perspective view



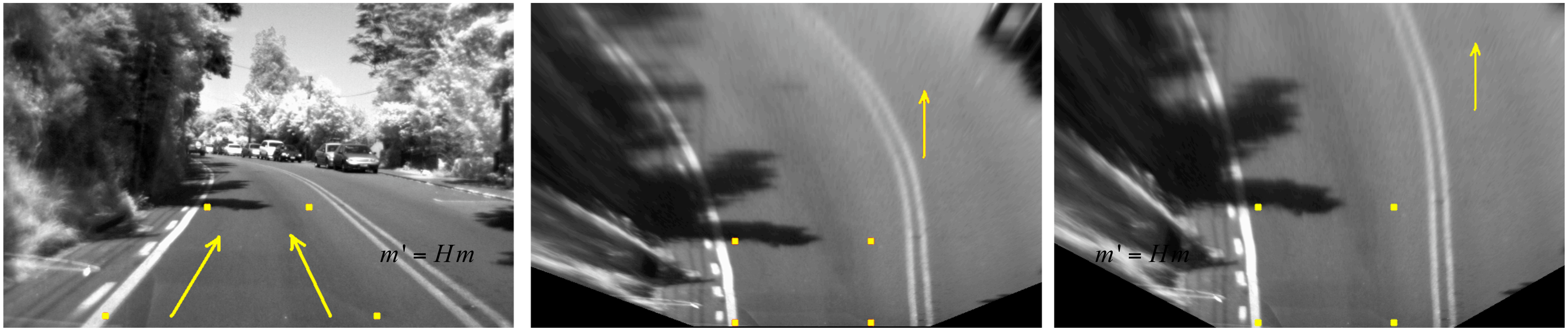


bird's eye view

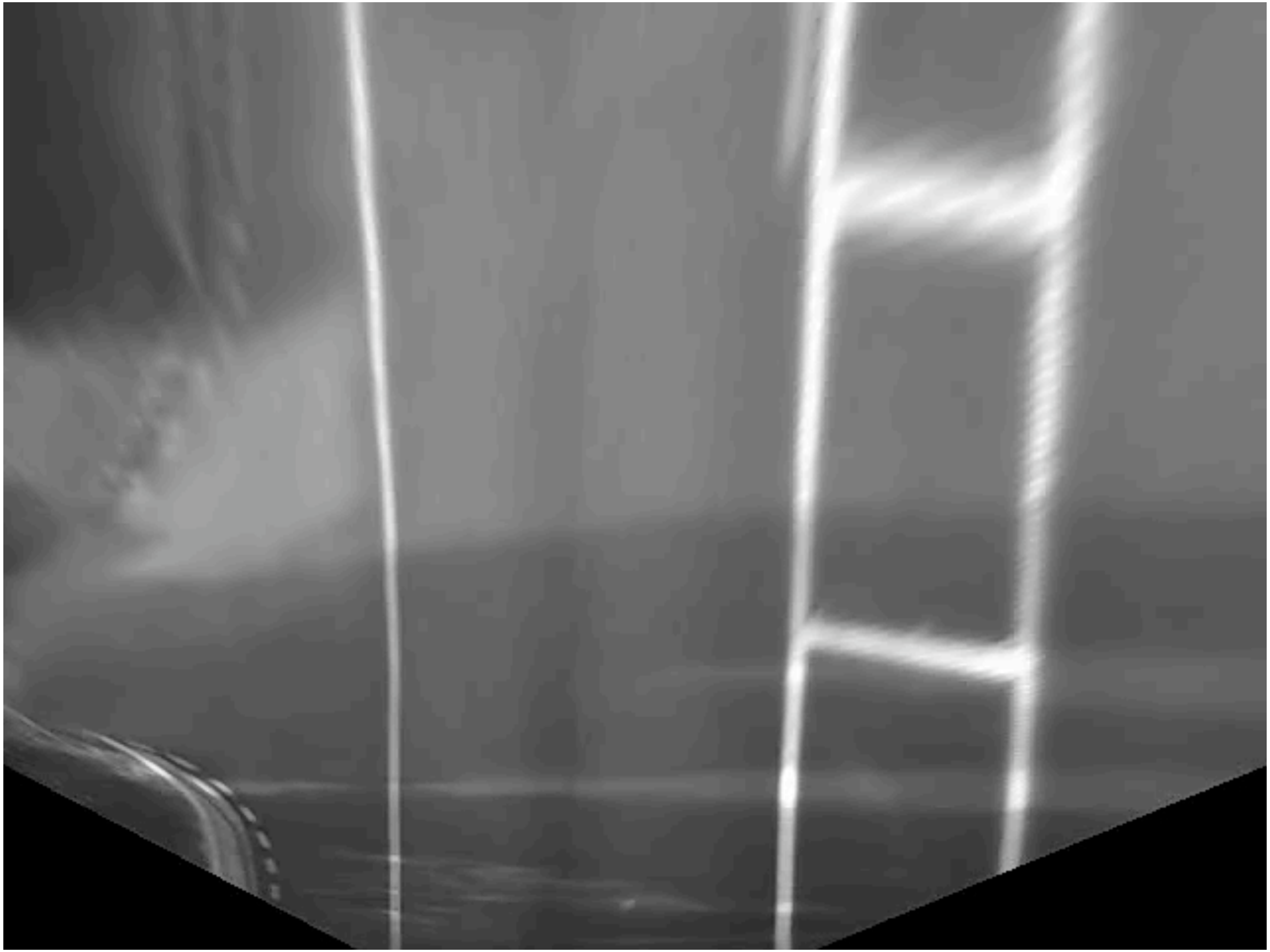
Bird’s-eye view mapping

Homography described by a 4x4 matrix
(use of homogeneous coordinates)

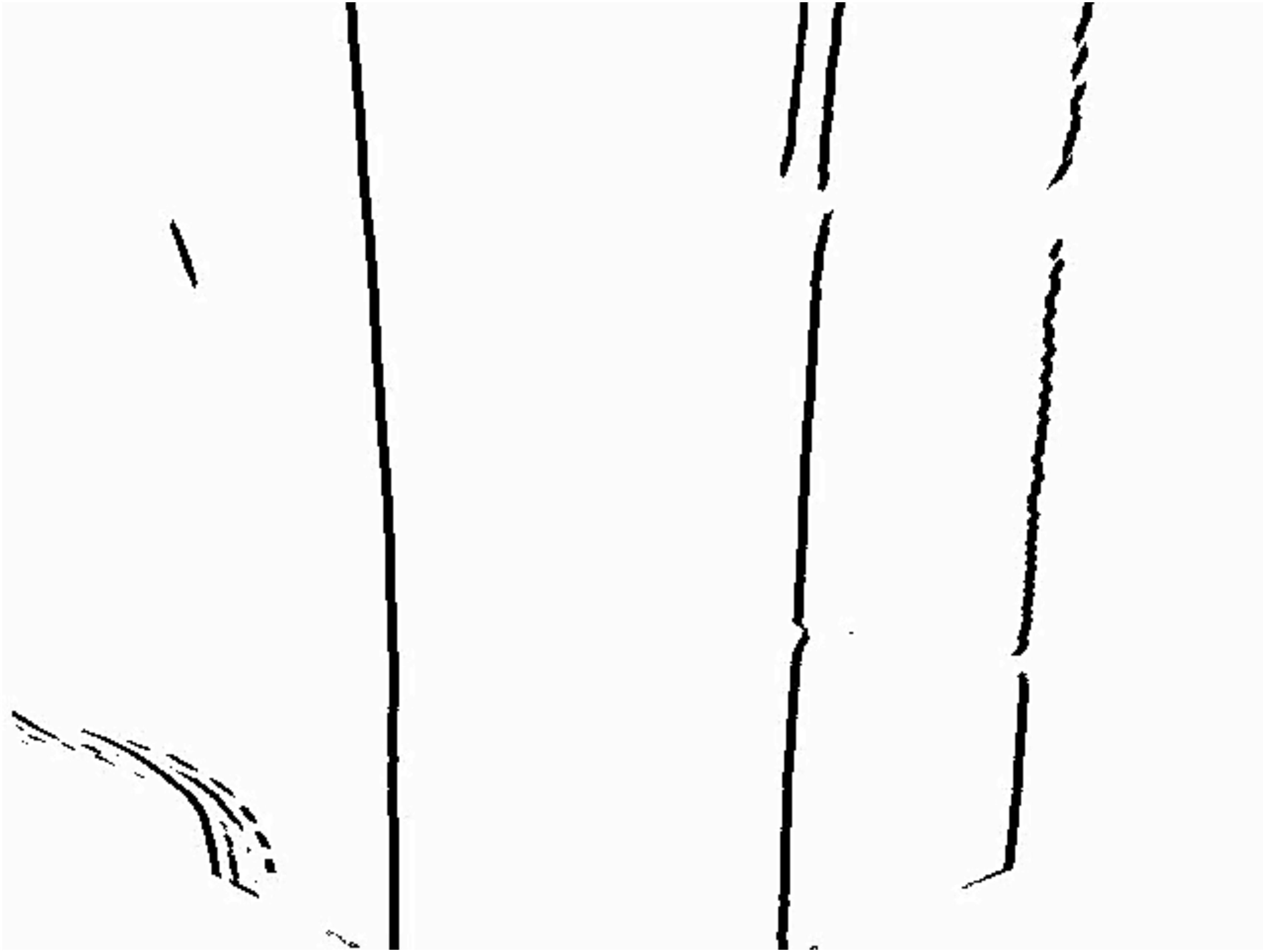
Four pairs of corresponding points determine
a homography.



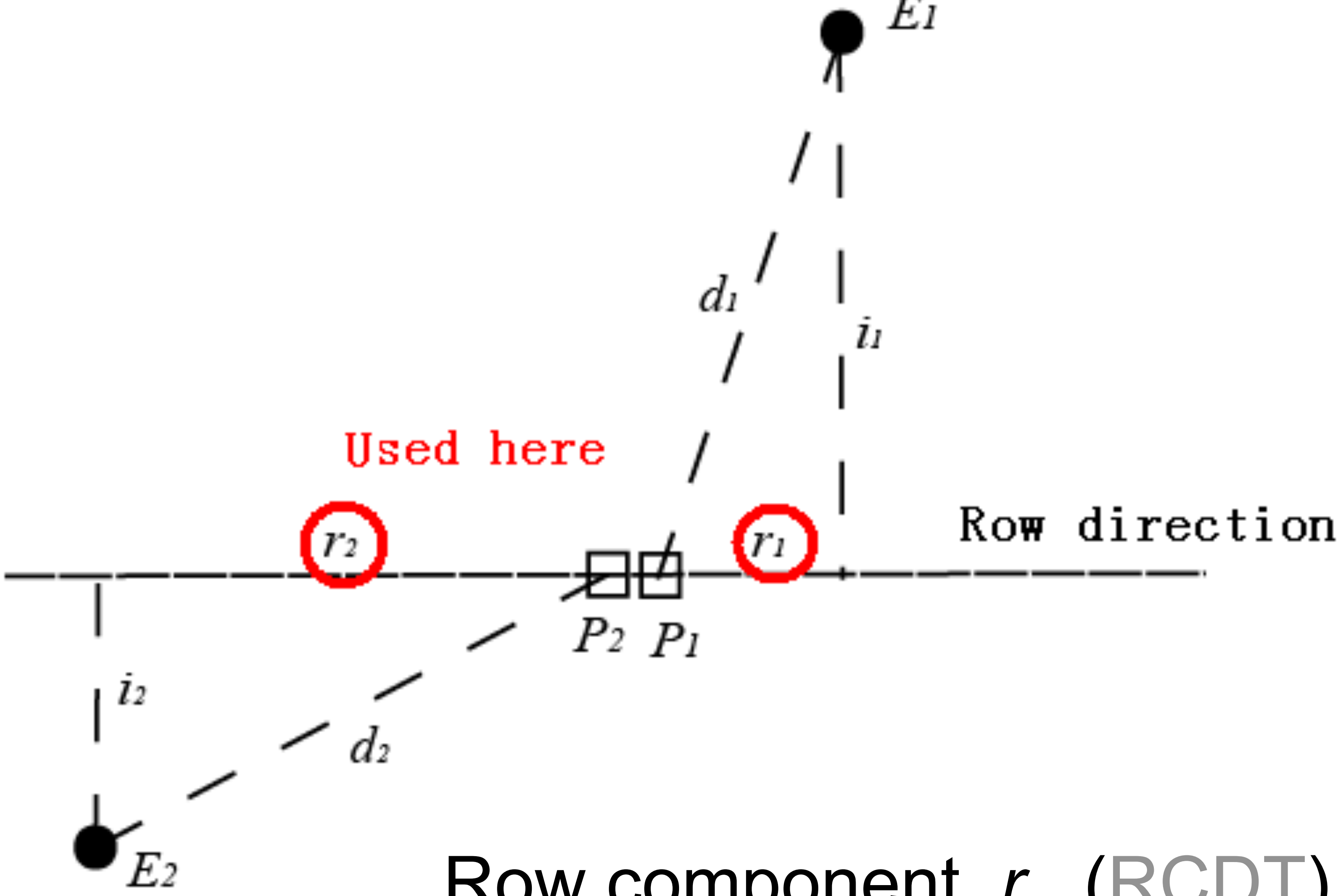
Bird’s-eye view sequence



Edge detection (dominant vertical edges)

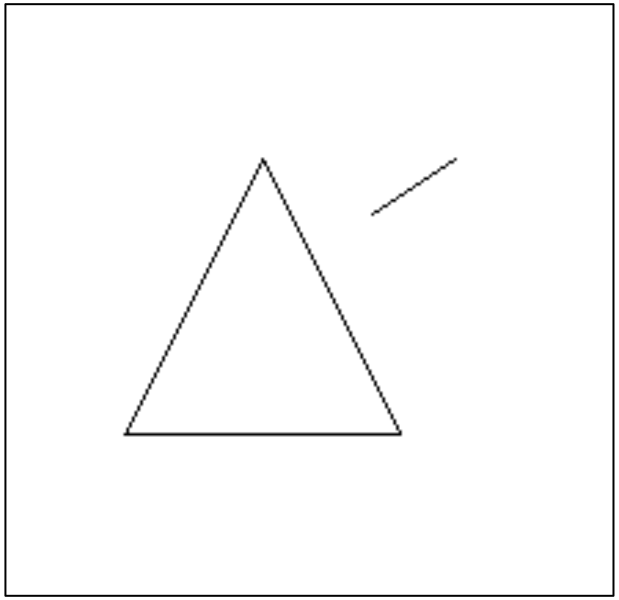


Distance transform (EDT)



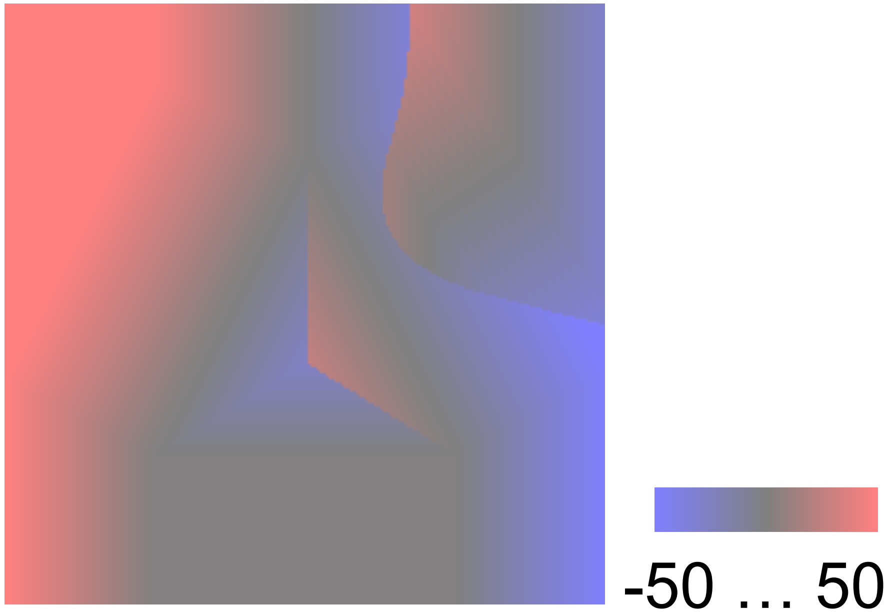
Row component *r* (RCDT)

Column component *i* (CCDT)

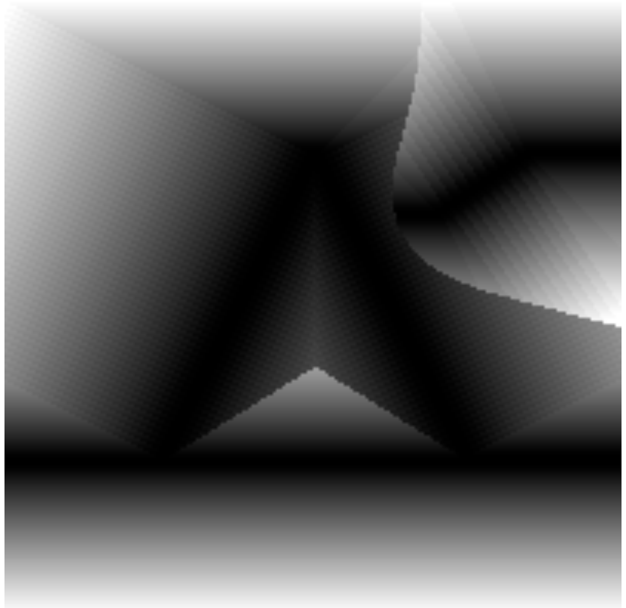


Original edge map

EDT



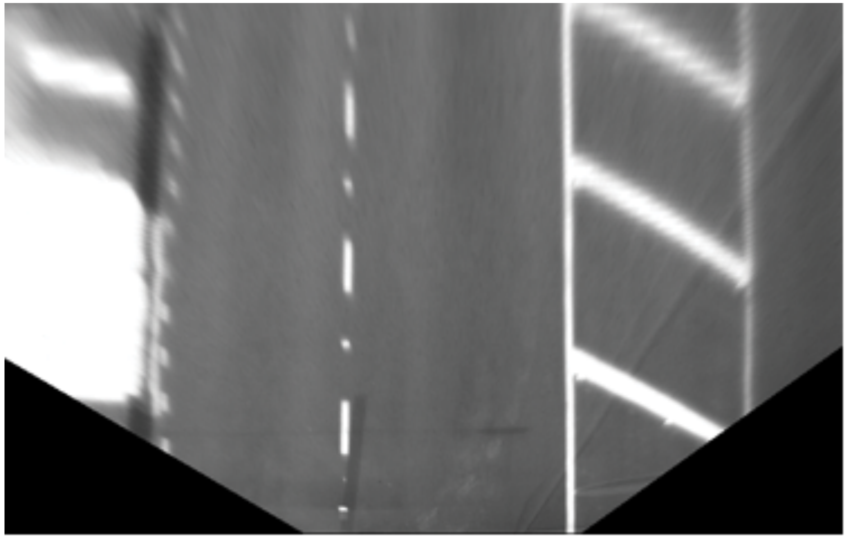
RCDT



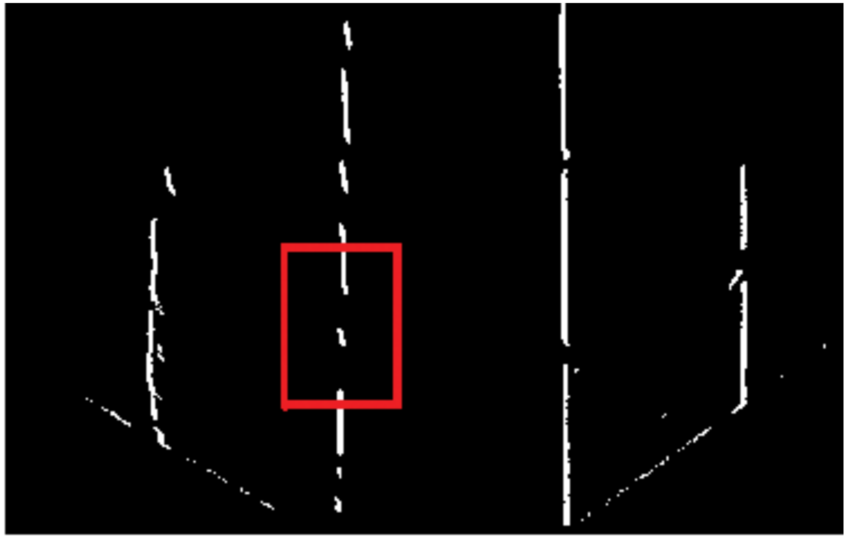
CCDT

Information from RCDT

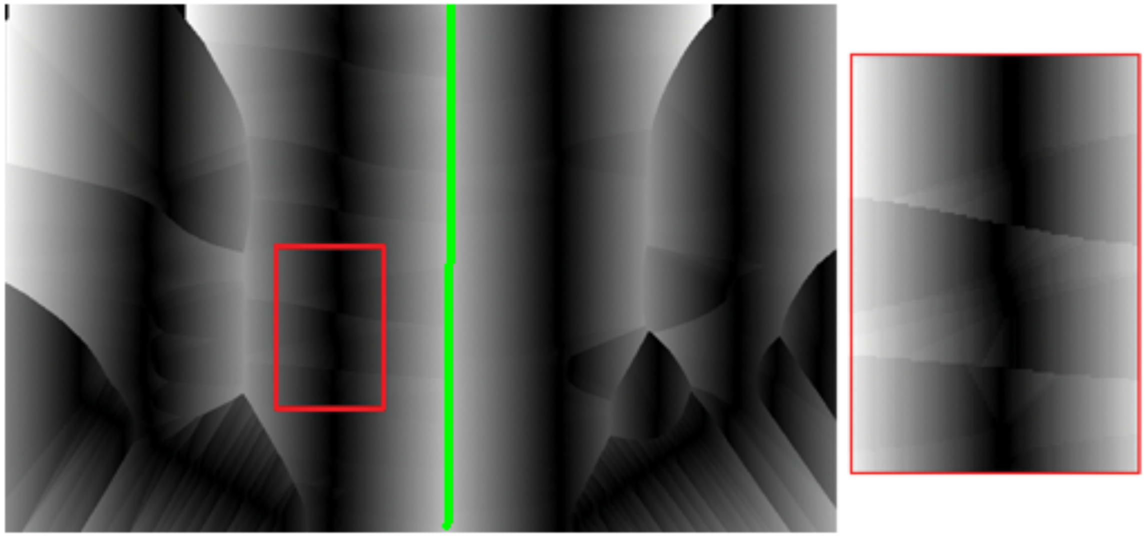
centerline, broken lane mark



Bird's-eye view



Edge detection



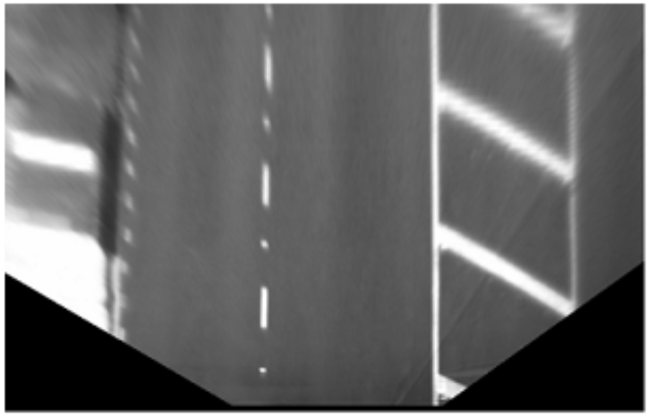
RCDT

Euclidean distance transform

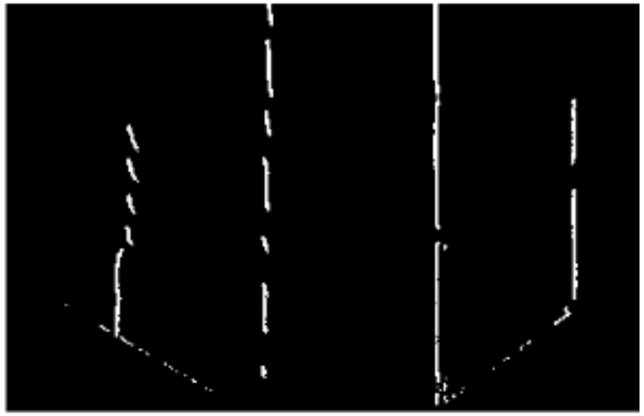
RCDT (negative or positive)

Workflow of lane detection (10 Hz currently in HAKA1, 640x480)

Input image

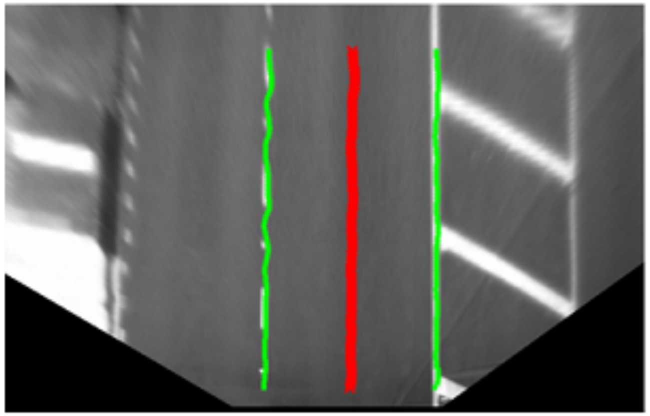


Bird's-eye View

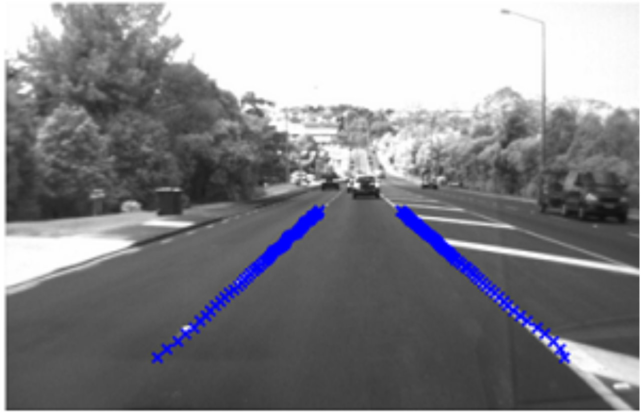


Edge detection

Row component of
Euclidean distance

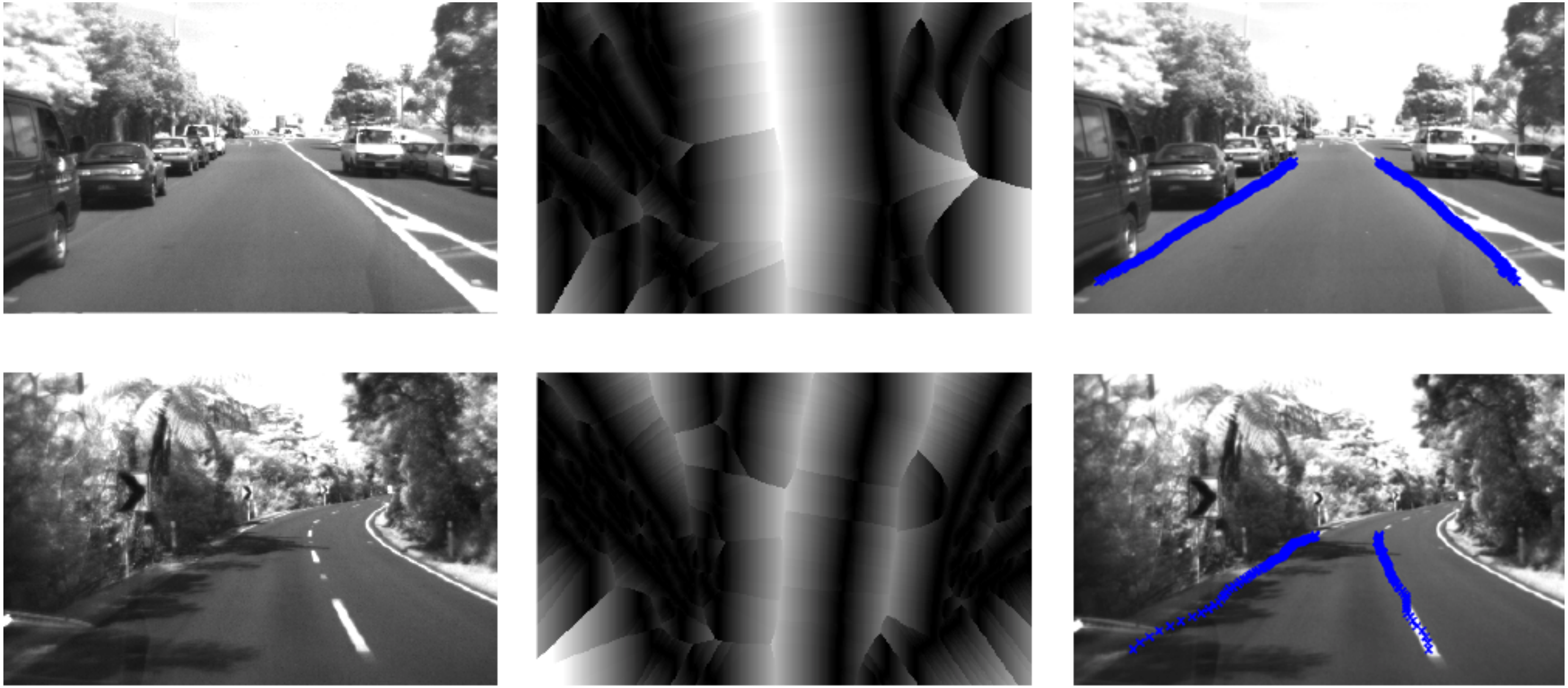


Lane detected in
bird's-eye view



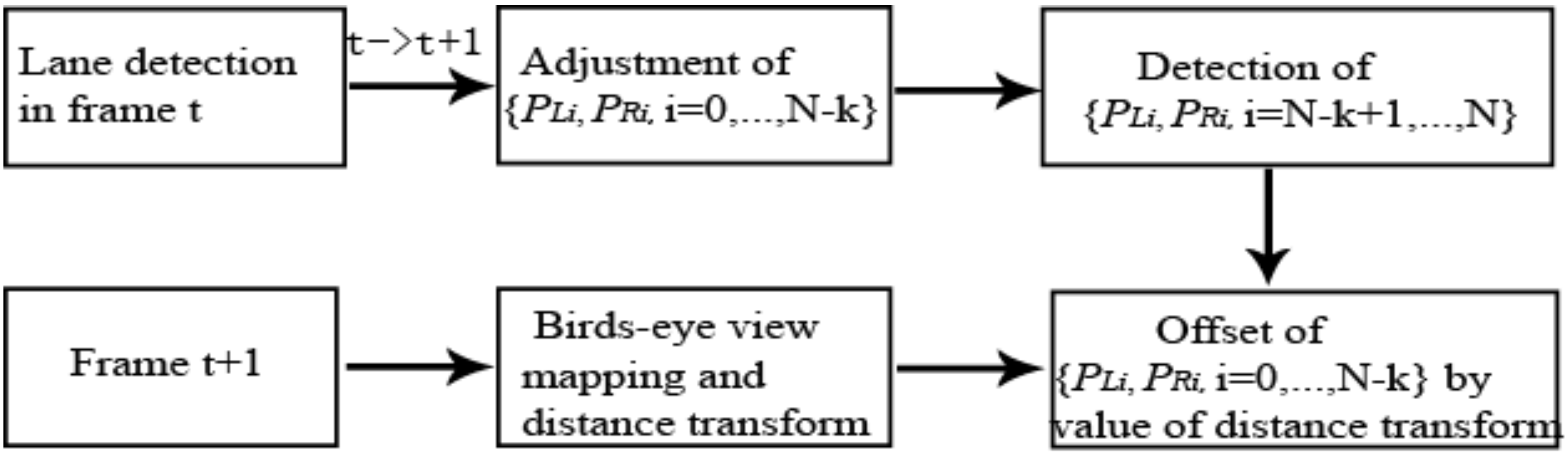
Lane detected

Examples: no border or no marks on the left

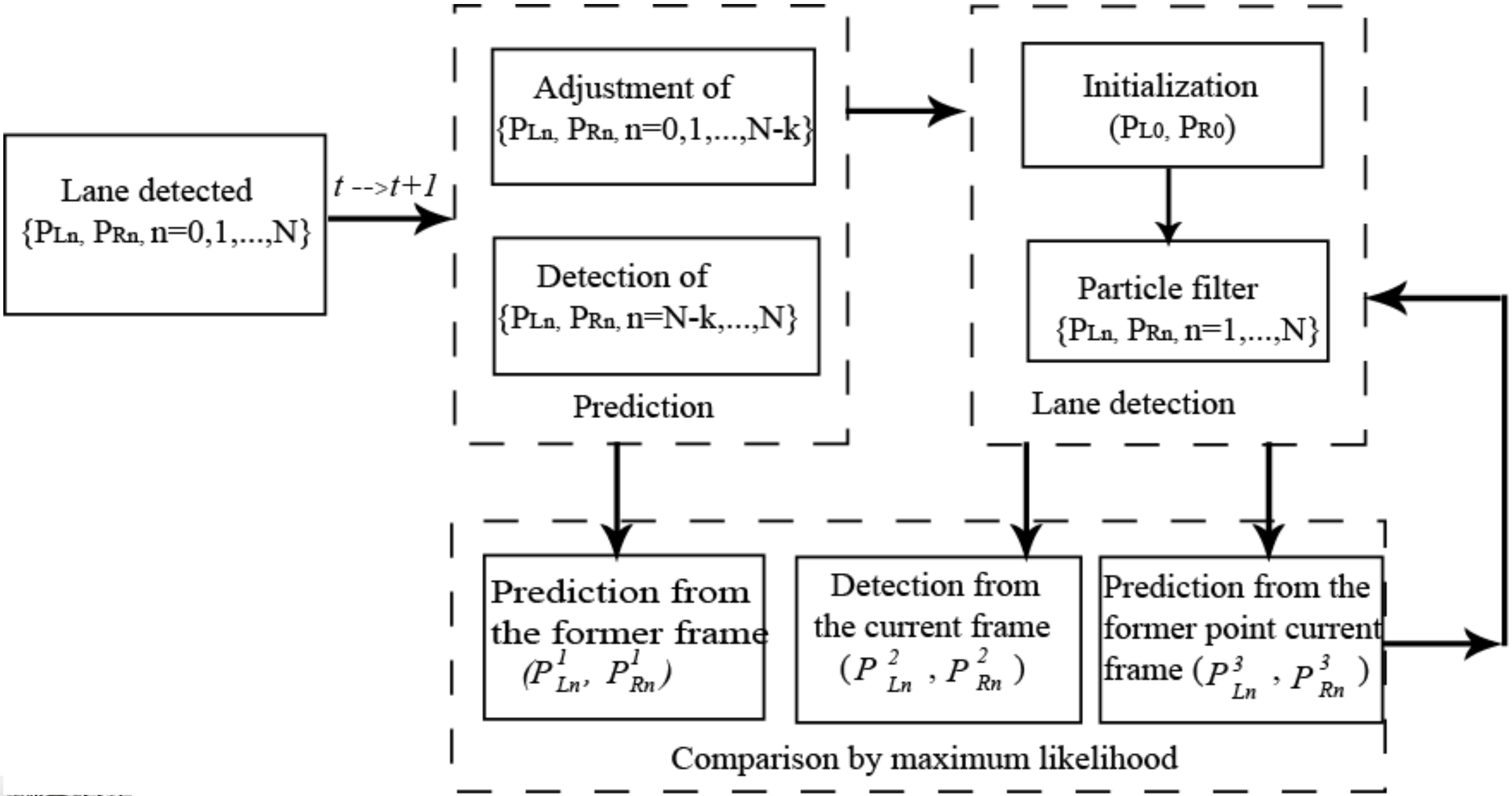


Input image RCDT Lane detected

Efficient tracking based on RCDT



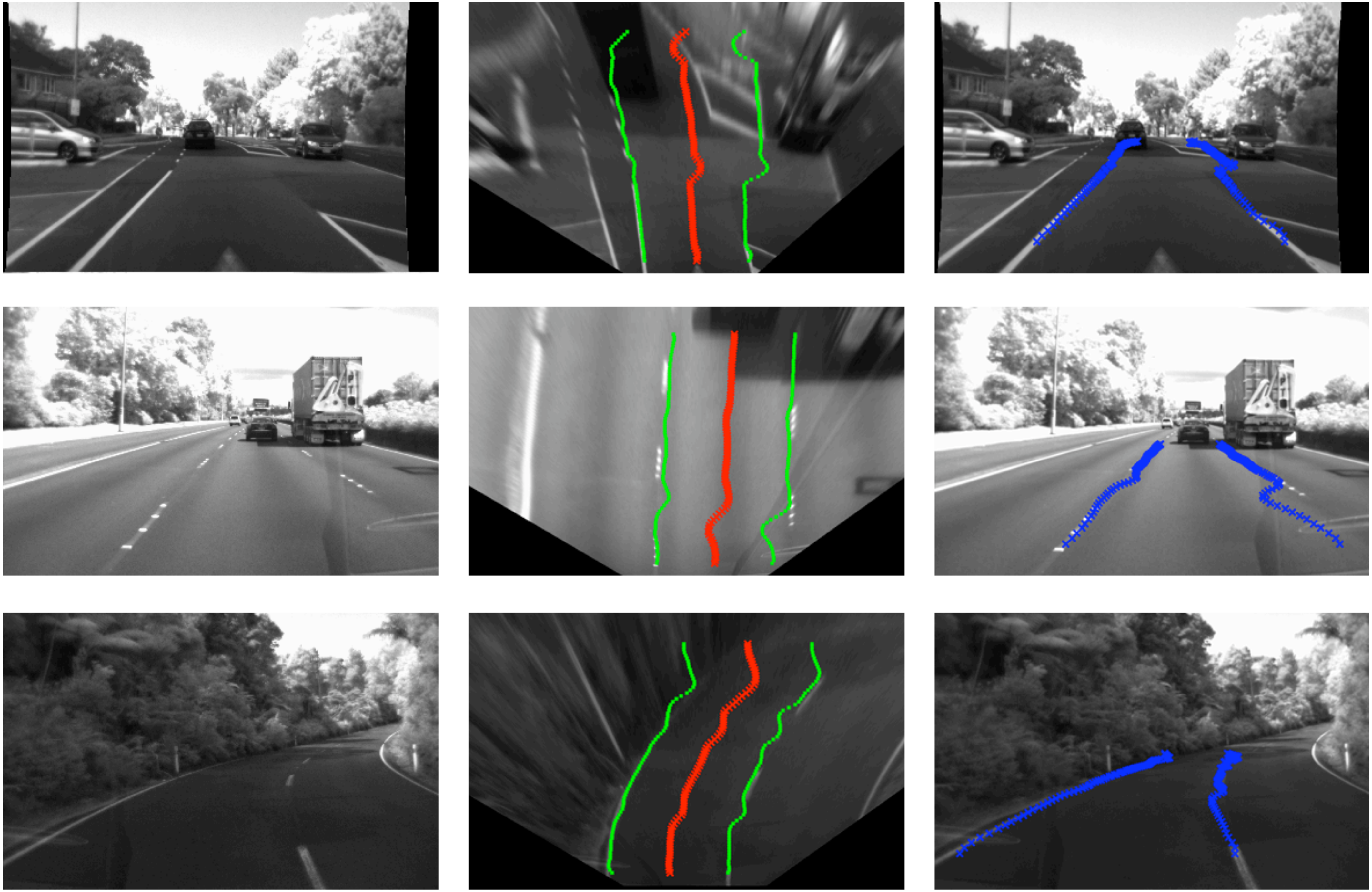
Robust tracking from RCDT



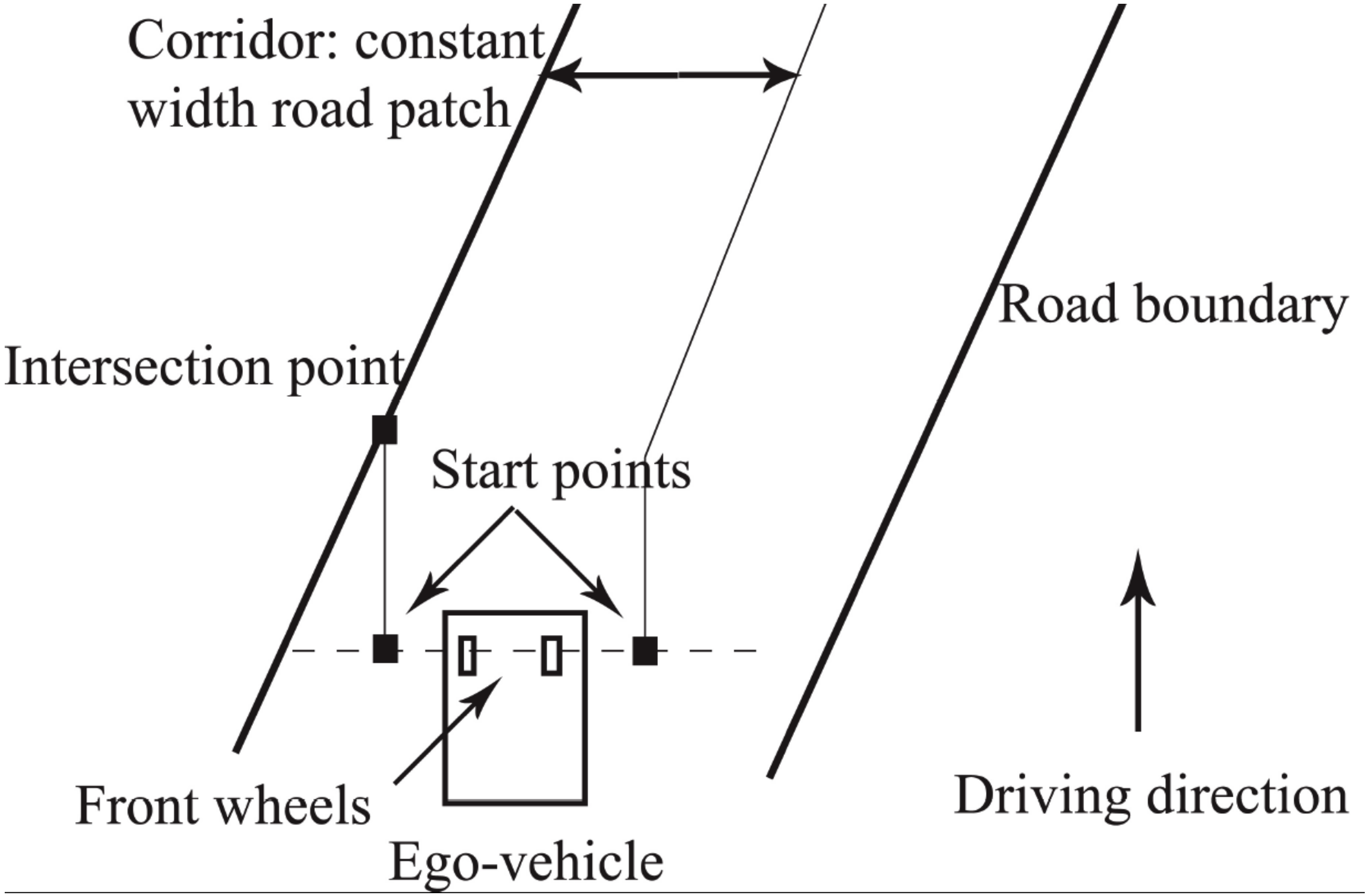
Efficient lane tracking

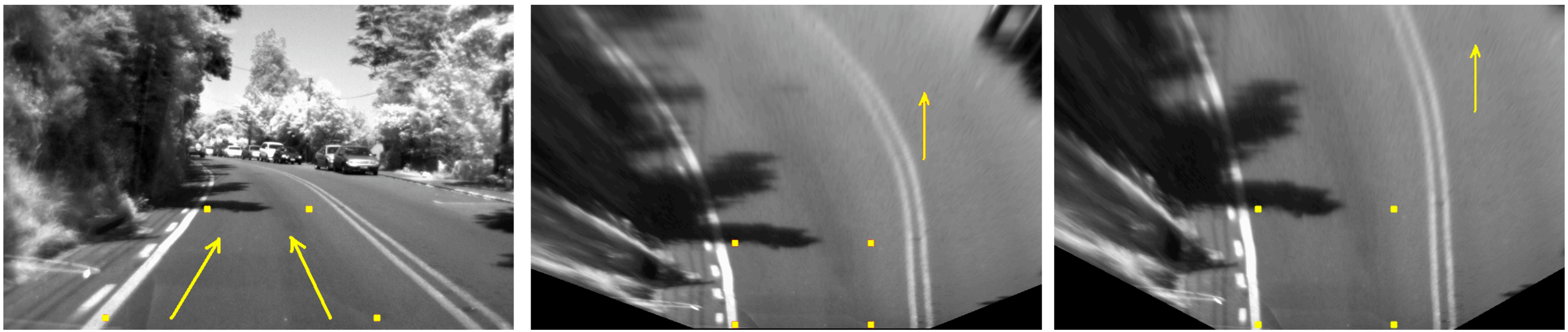
Robust lane tracking

Difficult situations > Corridor also based on trajectory



Corridor Detection and Tracking


Corridor: constant
width road patch
Road boundary
Intersection point
Start points
Front wheels
Ego-vehicle
Driving direction



Corridor

Two start points

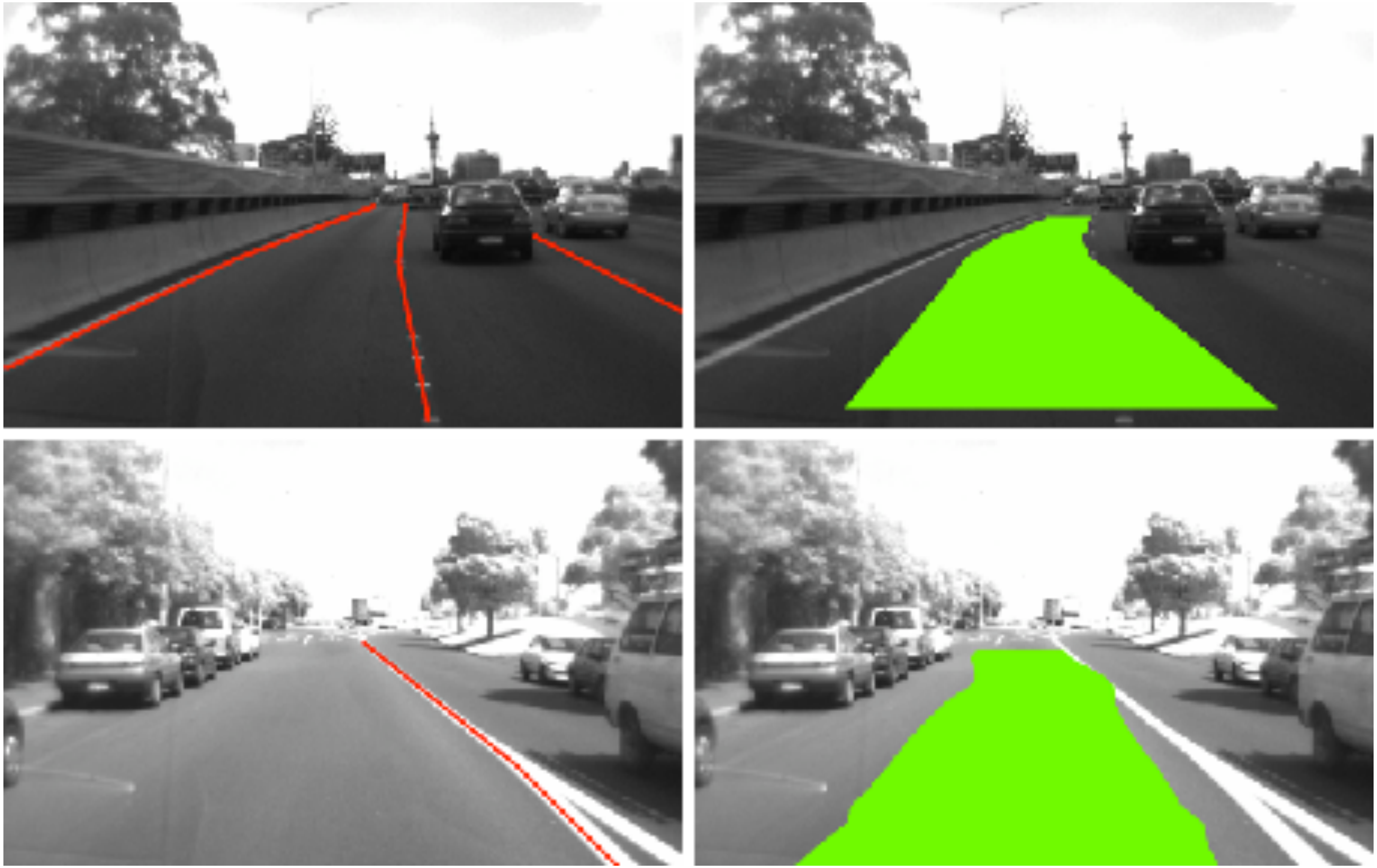
Constant width (slightly larger than that of the ego-vehicle)

Smooth borders

Proceeds in driving direction if not curved due to lane borders

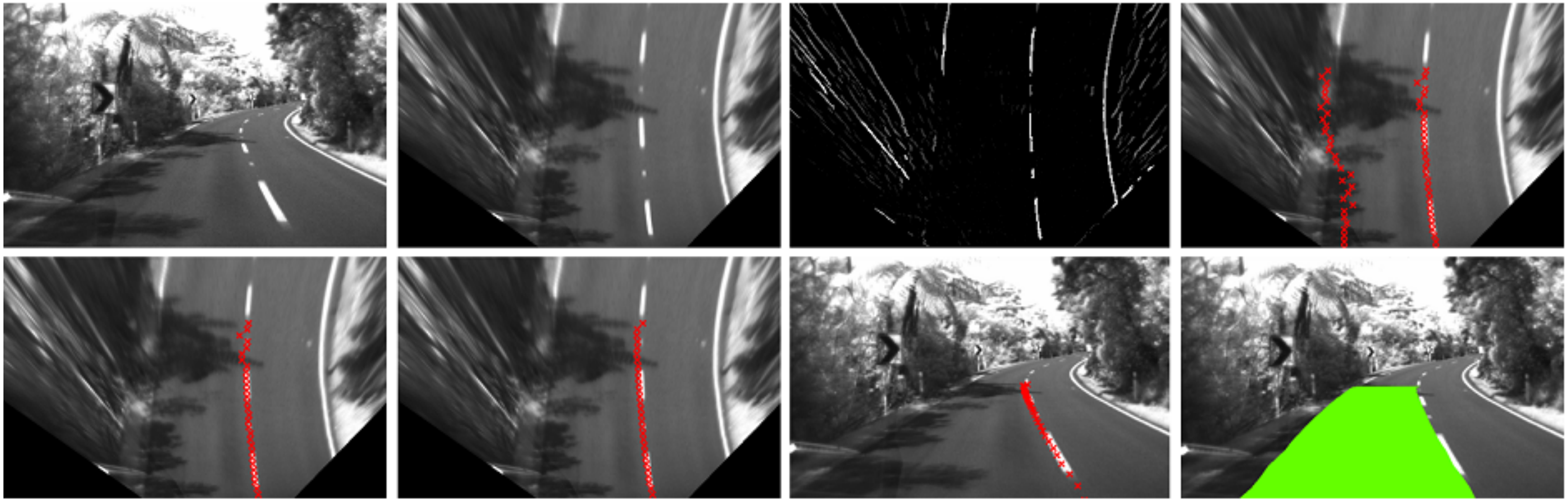
Constrained by lane borders (if possible)

Differences between lane and corridor



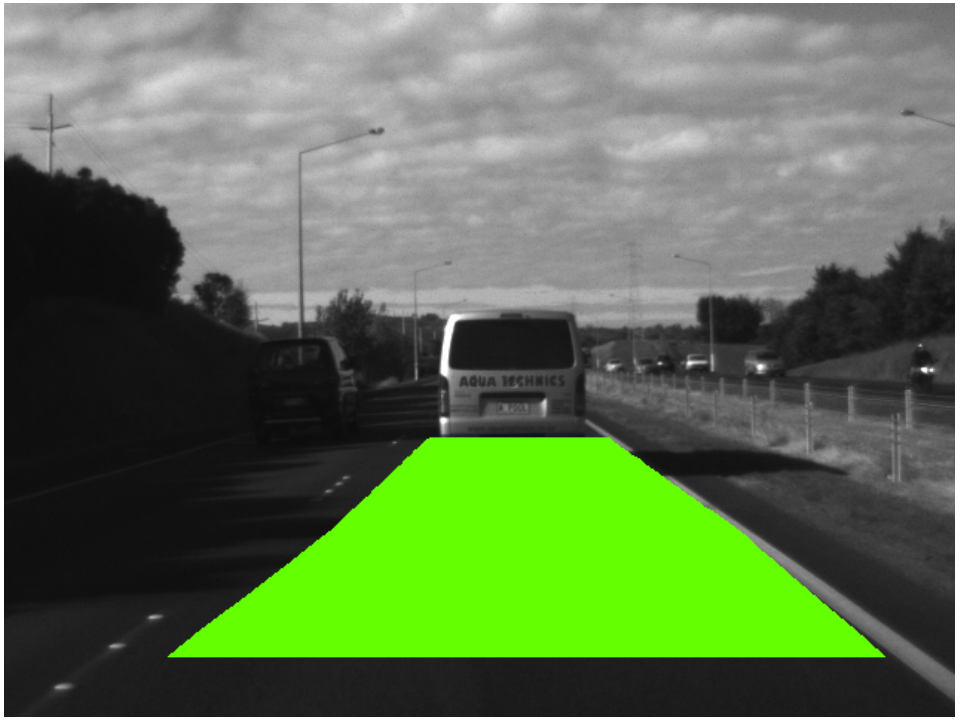
Workflow of corridor detection

Input image | Bird's eye view | Edge detected | Boundaries detected



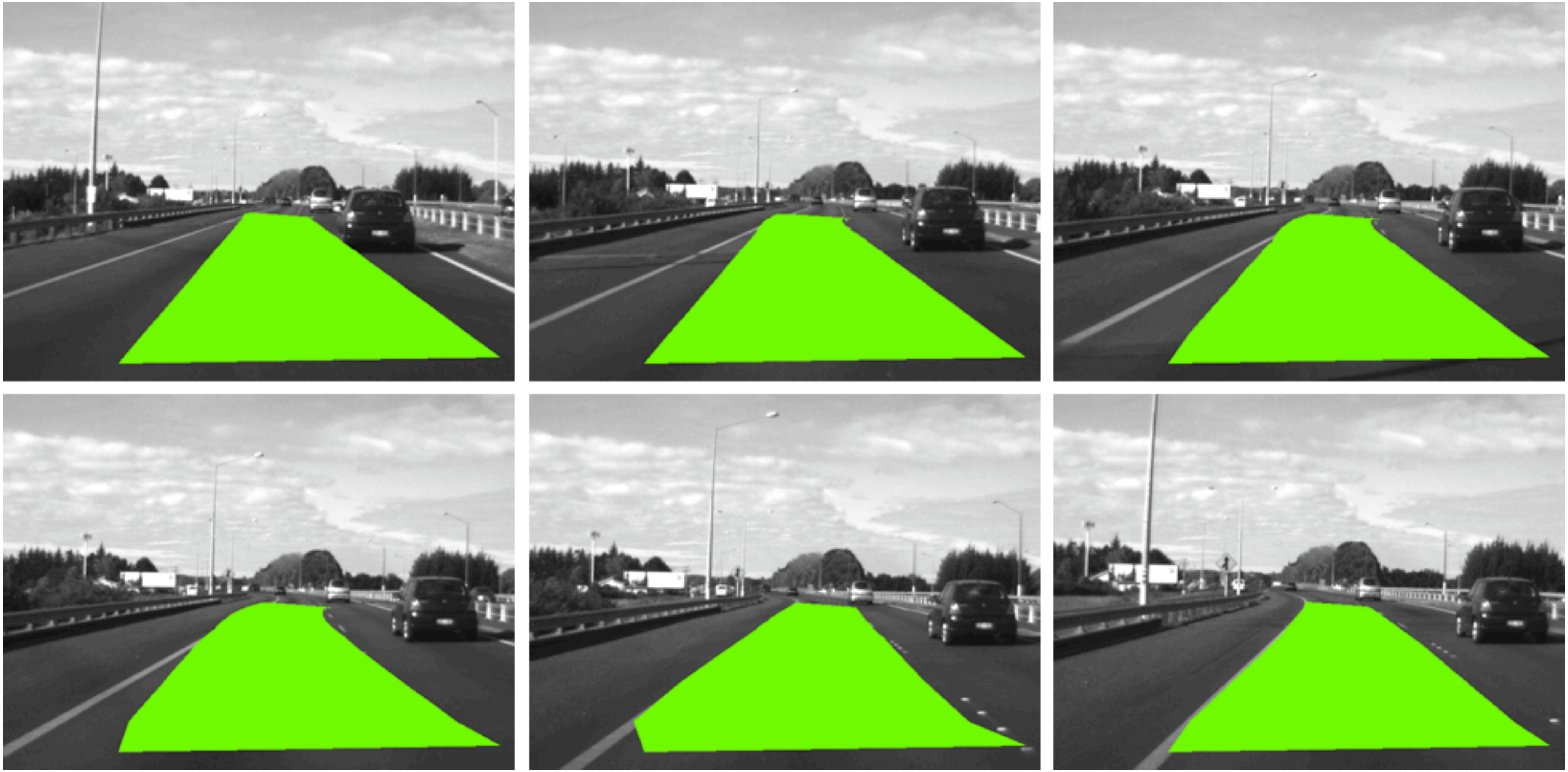
Boundary selection | Boundary smoothing | Back mapping | Corridor detected

Corridor detection with obstacle



Next: improved ground manifold calculation

Corridor detection during lane crossing



Concluding comments

The *.enpeda..* point of view

identify situations of traffic scenes

calculate winner (mean) and steadiness (variance)

for highly ranked methods and situations

calculate robustness by mean and variance across

identified situations

adaptation while driving:

(1) real-time situation recognition

(2) select method for the given situation

Stereo

Further (good) stereo matching paradigms besides scanline optimization, BP and GC?

Better ideas for temporal propagation/filtering?

Motion

Is there any competitive `discrete motion analysis' ?
For a start, see

[W. Trobin, T. Pock, D. Cremers, H. Bischof, ECCV 2008]
(an extension of graph-cut towards the continuous case)

More advanced modules for specific DAS tasks, certainly also with more interesting interactions with common areas of discrete mathematics, for

- ground manifold modeling
- ego-motion estimation
- object tracking (pedestrians, cars, ..)
- obstacle detection
- free-space detection (the space possibly to drive in)
- traffic sign recognition (not just speed or stop signs)

...

Vision-based DAS is the future.

The end.