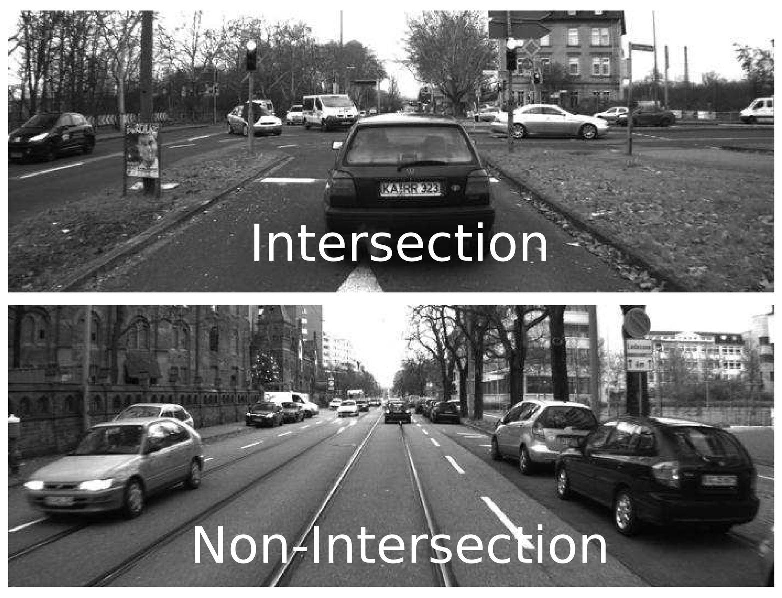
ObjectFlow: A Descriptor for Classifying Traffic Motion

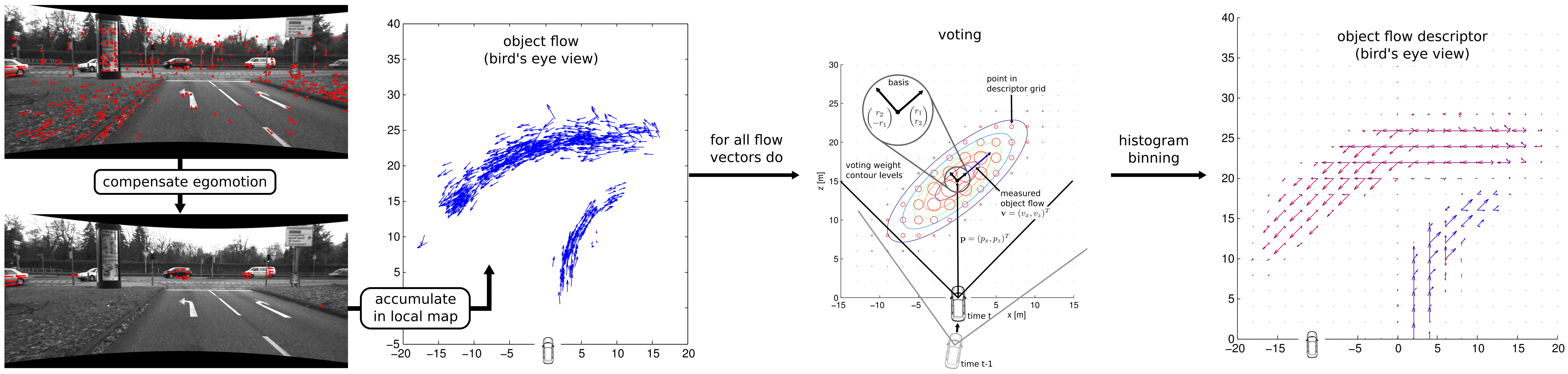
Andreas Geiger   Bernd Kitt   Karlsruhe Institute of Technology   {geiger,bernd.kitt}@kit.edu

Introduction: How can you visually detect an upcoming intersection?

- Most intersection detection methods heavily rely on lane markings
- Lane markings are often occluded or damaged
- We present a novel stereo-vision-based motion descriptor
- Our goal is to classify traffic scenes by their object motion
- We register flow vectors over time using visual odometry
- Votes cast by each flow vector are accumulated in histograms
- We classify the histograms into 'intersection' and 'non-intersection'



System Overview



Iterated Sigma Point Kalman Filter

State Model: $y_{k+1} = f(y_k) + w_k$ (constant velocity; state vector)

Observation Model: $z_{k+1} = h(y_{k+1}) + v_{k+1}.$ (trifocal tensor; state vector)

State Vector: $y = (V_X, V_Y, V_Z, \omega_X, \omega_Y, \omega_Z)^T$

For more details come to our talk here at IV'10: "Visual Odometry based on Stereo Image Sequences with RANSAC-based Outlier Rejection Scheme"

Weight of individual Vote (squared exponential):

$$w(\mathbf{q}, \xi^{2D}) = \exp\left\{-(\mathbf{q}-\mathbf{p})^{\mathbf{T}}\mathbf{W}(\mathbf{q}-\mathbf{p})\right\} \in [0..1]$$

grid point q; flow vector (p,v); precision matrix $\mathbf{W} = \mathbf{R}\mathbf{\Lambda}\mathbf{R}^{\mathbf{T}}$

Eigenvalue Decomposition of W:

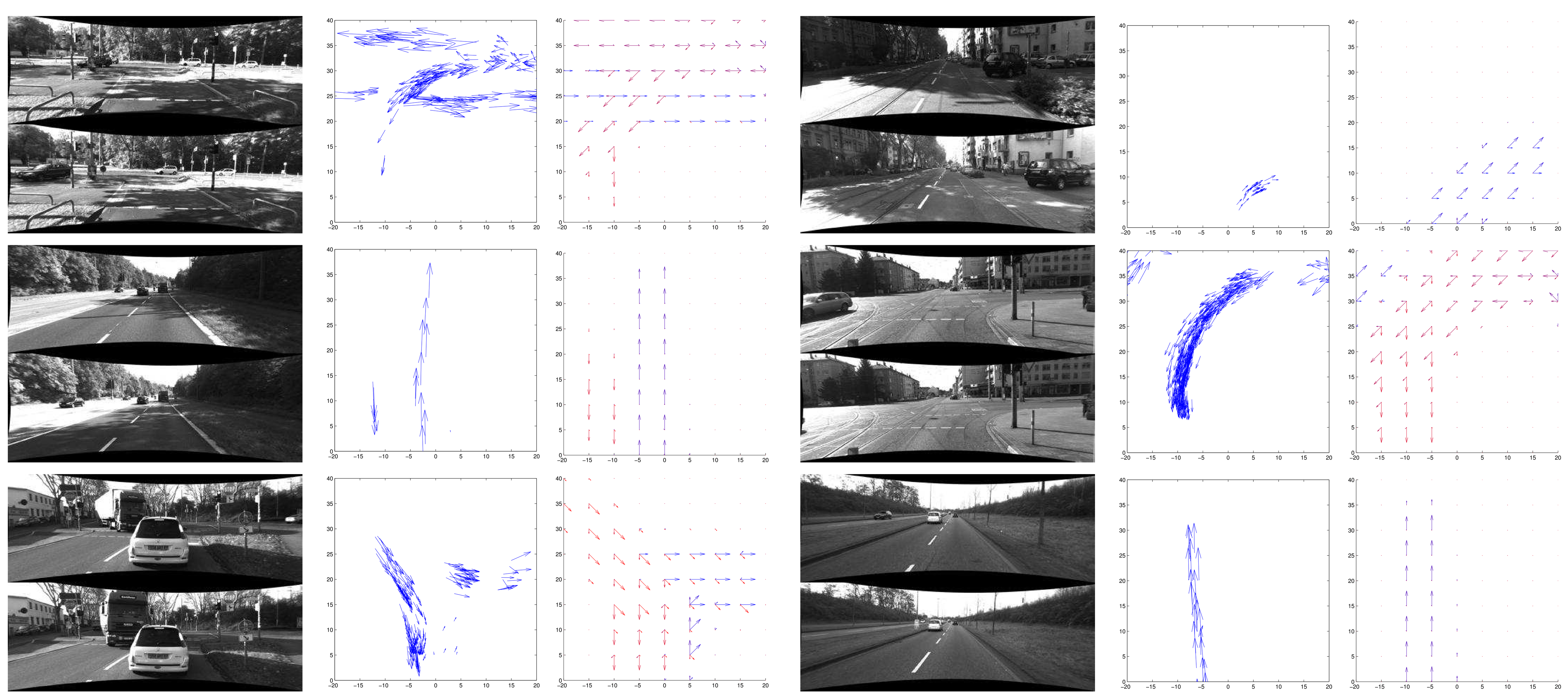
$$\mathbf{R} = \begin{pmatrix} r_1 & r_2 \\ r_2 & -r_1 \end{pmatrix} \qquad \mathbf{\Lambda} = \begin{pmatrix} 1/\lambda_1^2 & 0 \\ 0 & 1/\lambda_2^2 \end{pmatrix}$$

Basis Vectors of Rotation Matrix:

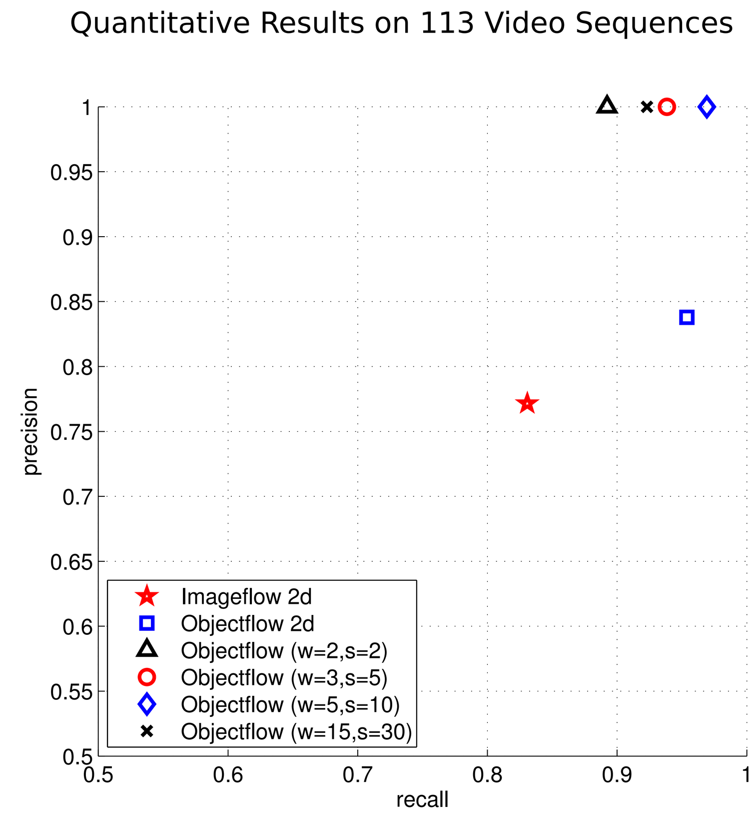
$$r_1 = \frac{v_x}{\sqrt{v_x^2 + v_z^2}} \qquad r_2 = \frac{v_z}{\sqrt{v_x^2 + v_z^2}}$$

flow velocity; flow velocity; influence in voting direction and perpendicular to voting direction

Experimental Evaluation: SVM-based leave-one-out classification of 113 Video Sequences



Quantitative Results on 113 Video Sequences



Method (w,s)	TP	TN	FP	FN	Precision	Recall
Img.Flow 2d	54	32	16	11	0.77	0.83
Obj.Flow 2d	62	36	12	3	0.84	0.95
Obj.Flow (2,2)	58	48	0	7	1.00	0.89
Obj.Flow (3,5)	61	48	0	4	1.00	0.94
Obj.Flow (5,10)	63	48	0	2	1.00	0.97
Obj.Flow (15,30)	60	48	0	5	1.00	0.92

Conclusion

- Robust feature extraction up to a distance of 50 m
- Complementary to lane-marking-based features
- Only available in the presence of moving traffic

Future Work

- GPU-based feature matching for real-time
- Multiple intersection classes
- Inference of most likely intersection geometry

Karlsruhe Institute of Technology: The cooperation of Forschungszentrum Karlsruhe GmbH and Universität Karlsruhe (TH)
The KSOP funding is provided by the German Research Foundation (DFG)