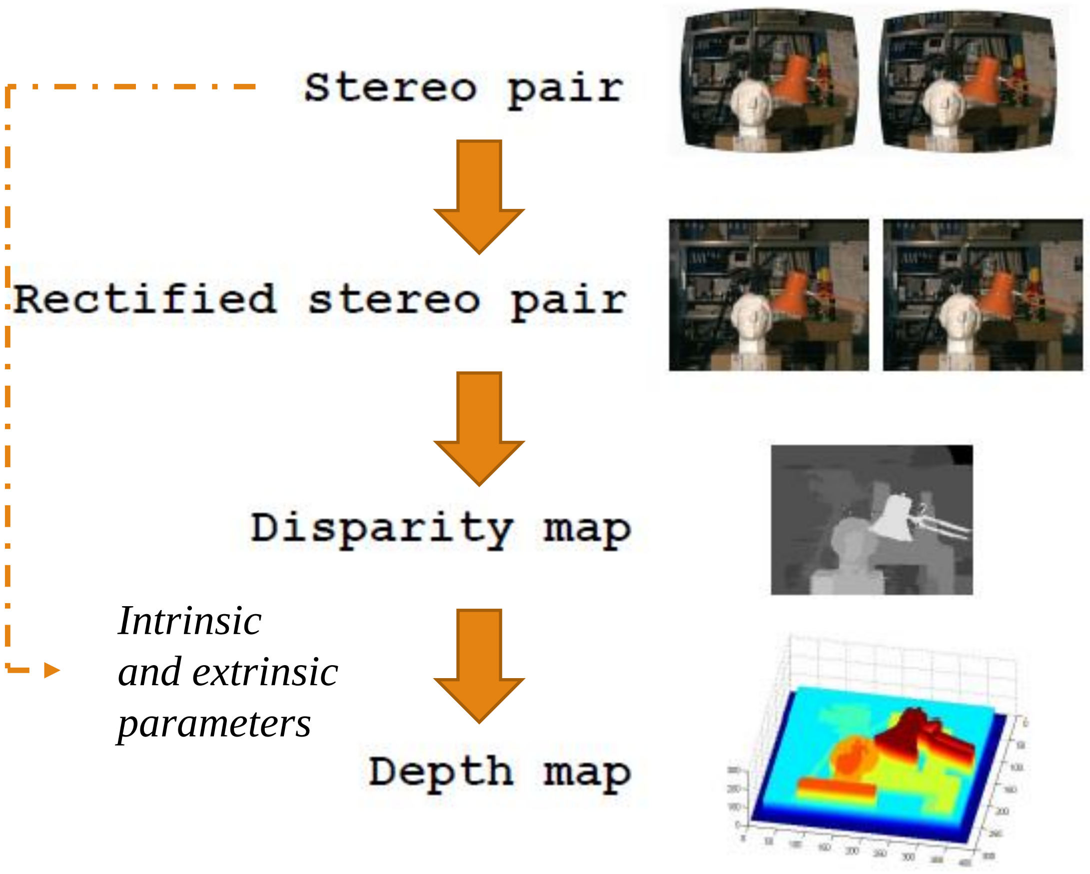
Stereo matching

ZEHUA FU

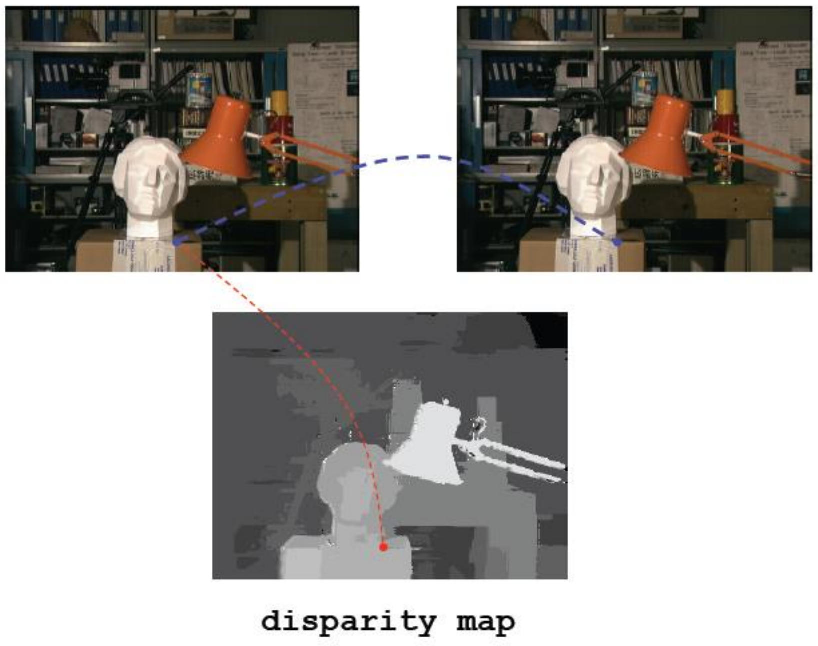
DIRECTEUR DE THÈSE： MOHSEN ARDABILIAN

Stereo Matching



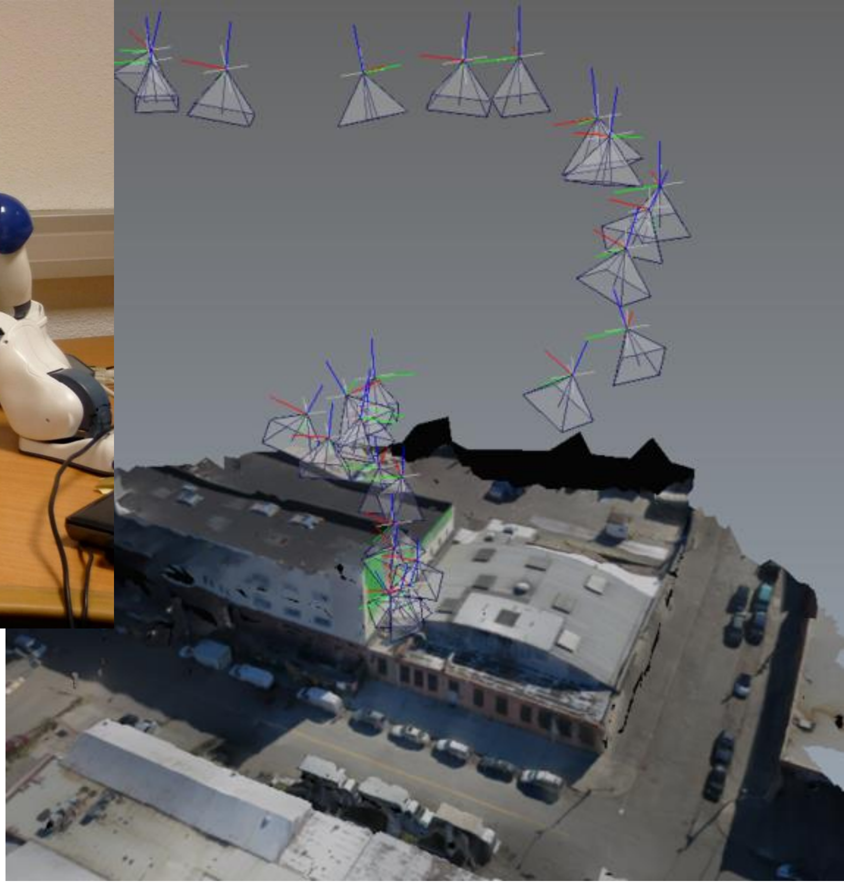
- **Problem Description:**

  Find the stereo correspondence between two given image pair.



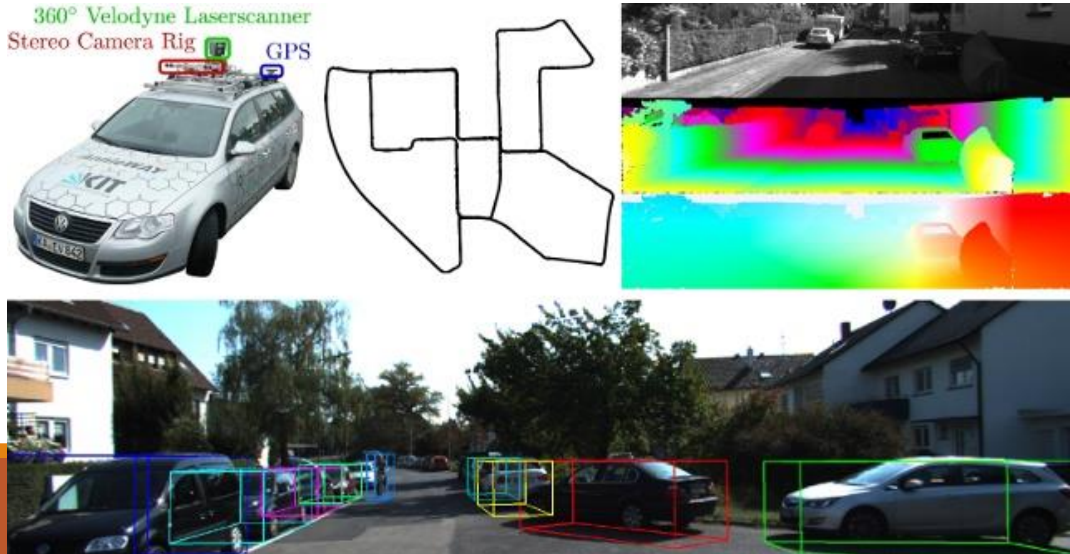
Applications

- Autonomous driving
- Robotics
- Intermediate view generation



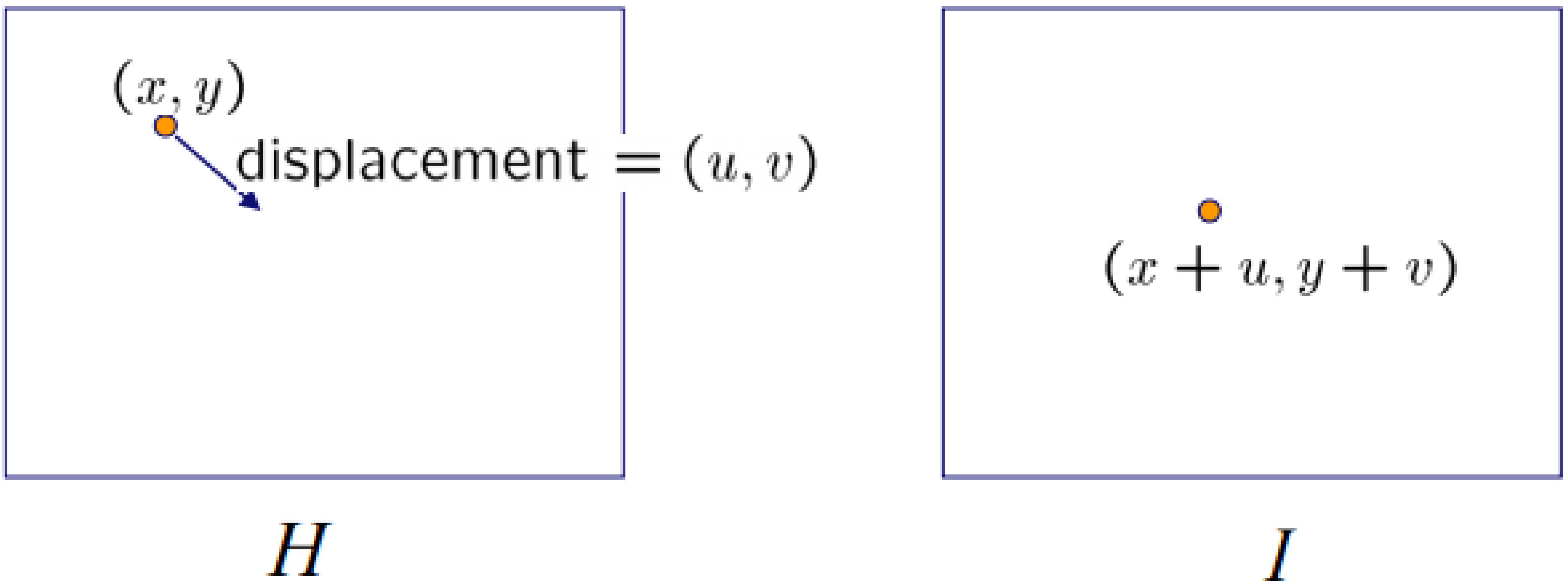
Spatial information

Temporal information

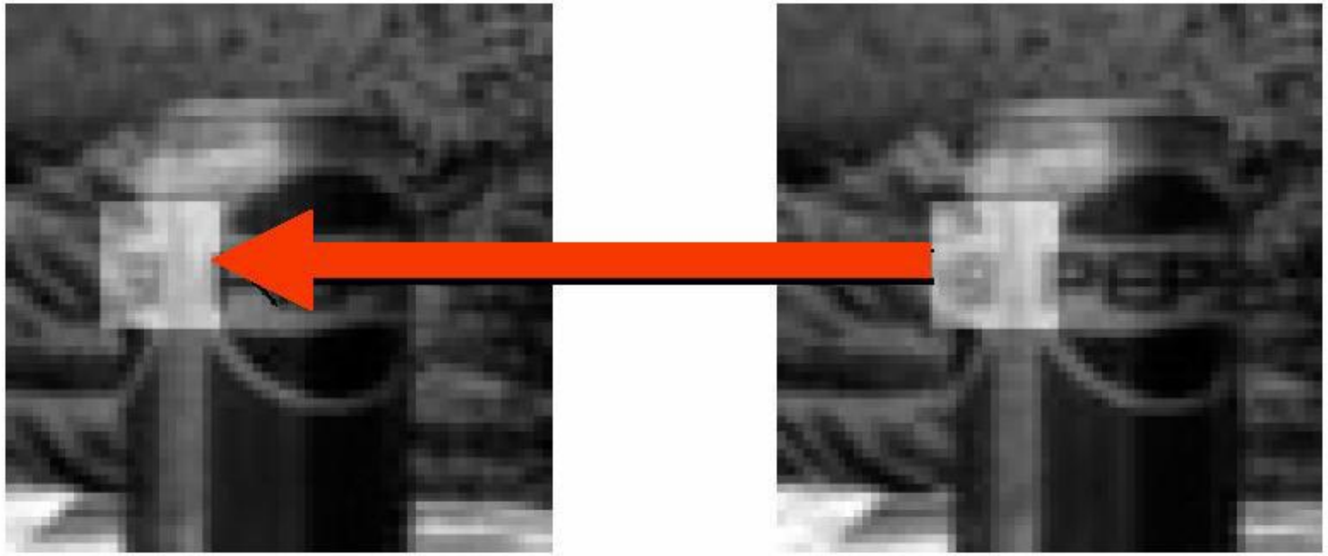


[1] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for autonomous driving? the KITTI vision benchmark suite,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3354–3361, 2012.
[2] https://www.wikiwand.com/en/Nao_(robot)
[3] https://github.com/OpenDroneMap/OpenDroneMap/wiki

Optical Flow



Brightness Constancy

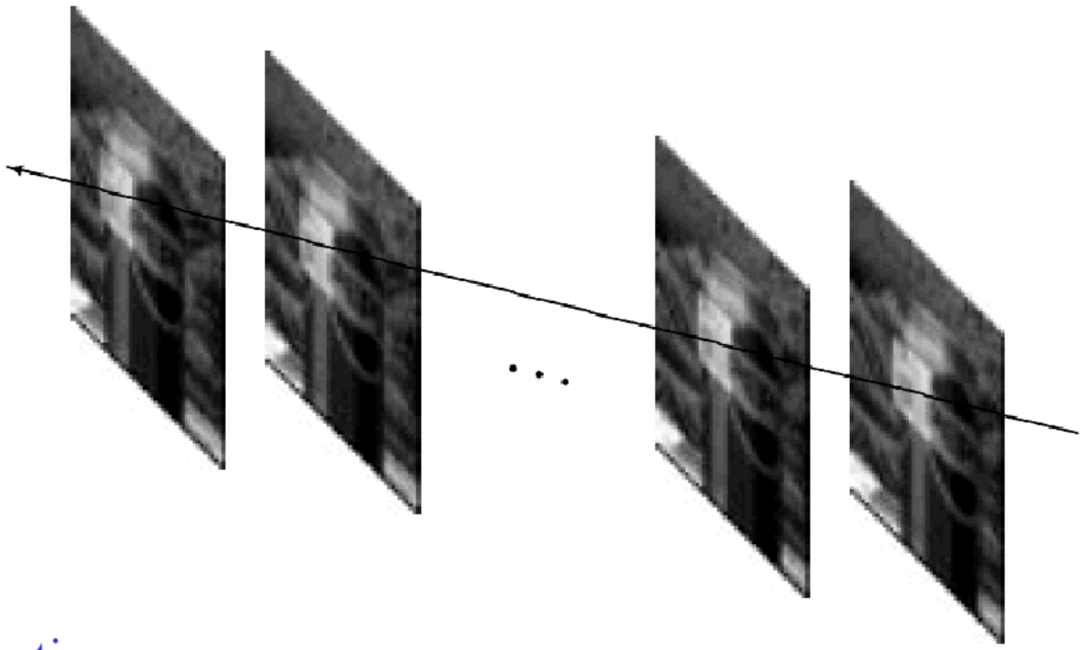


Assumption

Image measurements (e.g. brightness) in a small region remain the same although their location may change.

$$I(x+u, y+v, t+1) = I(x, y, t)$$

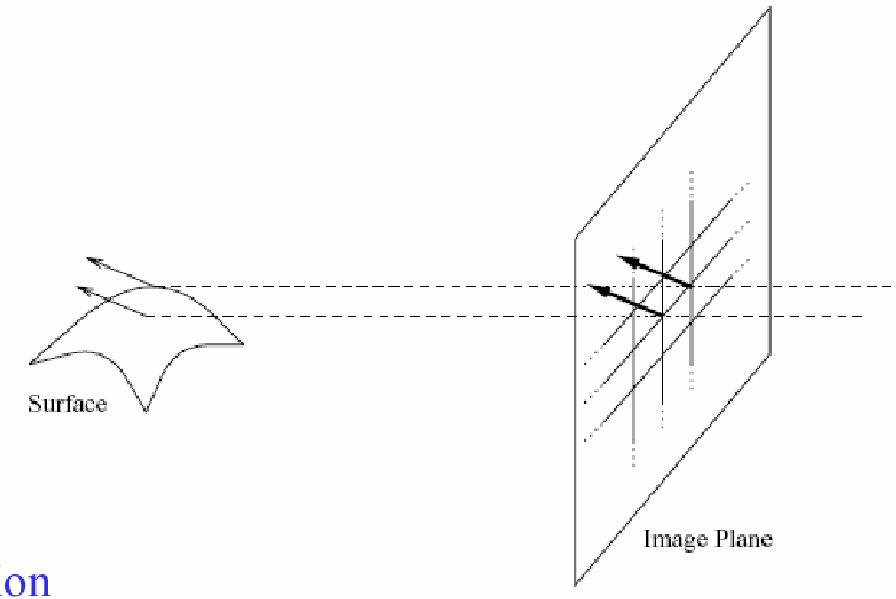
Temporal Persistence



Assumption:

The image motion of a surface patch changes gradually over time.

Spatial Coherence



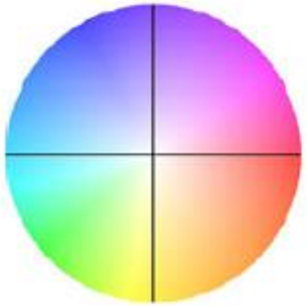
Assumption

* Neighboring points in the scene typically belong to the same surface and hence typically have similar motions.
* Since they also project to nearby points in the image, we expect spatial coherence in image flow.

Slide from Michael Black, CS143

Optical Flow

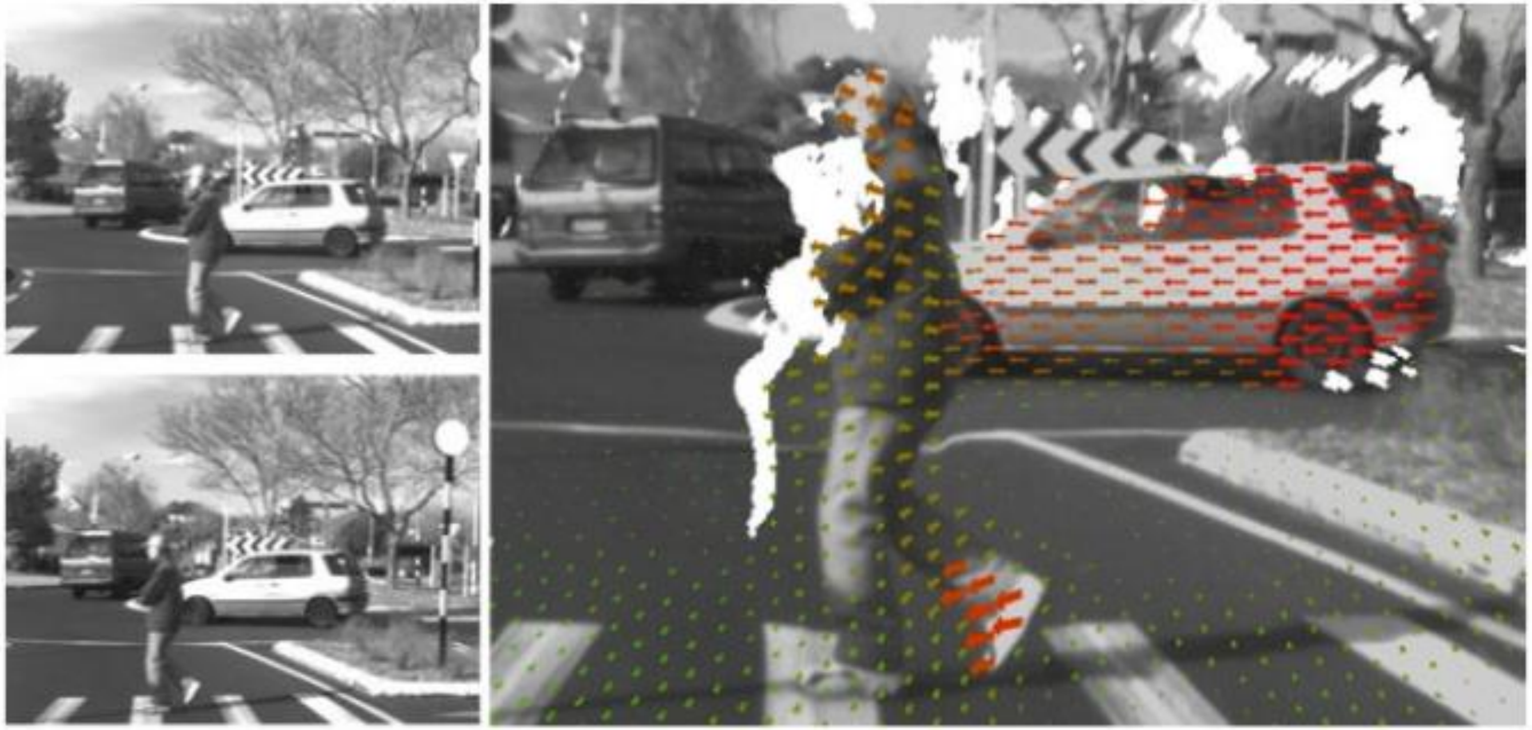
Fig. 2.1 Color coding of flow vectors: Direction is coded by hue, length by saturation. The example on the *right* shows the expanding flow field of a forward motion. Flow vectors above 20 px are saturated and appear in *darker* colors



- Which pixel went where?

Time: t

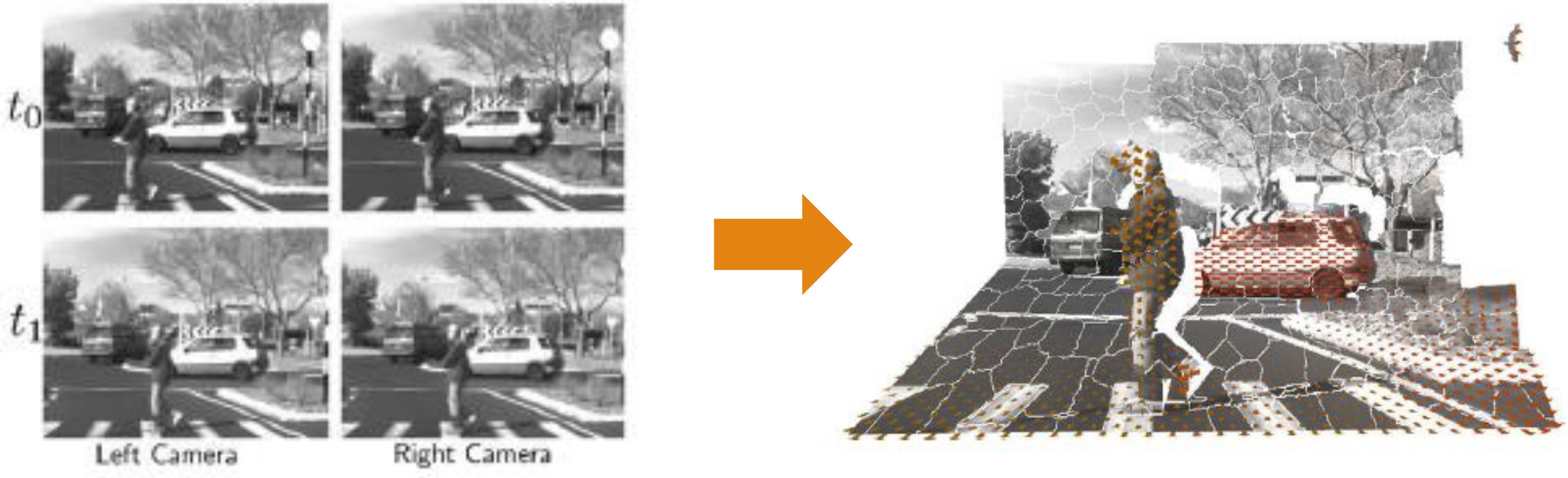
Time: t + dt



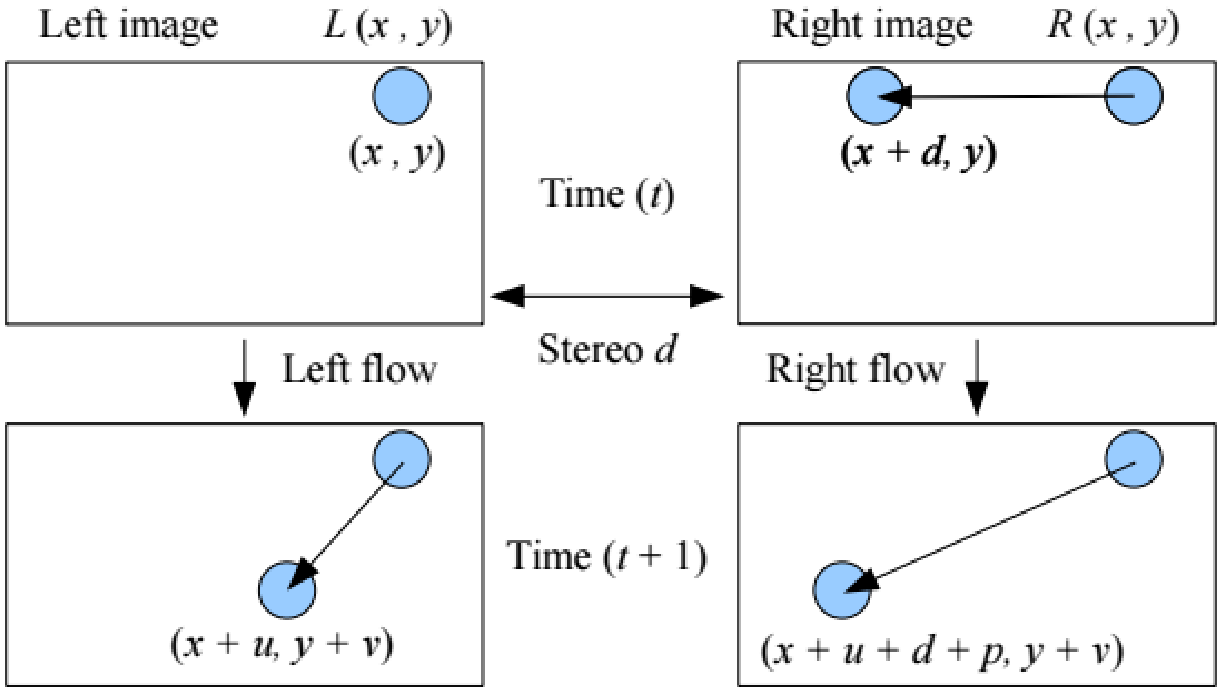
[1] Wedel A, Cremers D. Stereo scene flow for 3D motion analysis[M]. Springer Science & Business Media, 2011.

3D Scene Flow

- "Joint stereo and optical flow"
- Given ⩾2 video frames from ⩾2 different viewpoints
- Estimate dense 3D shape and 3D motion field



Scene Flow



Left image
L (x , y)
(x , y)
Right image
R (x , y)
(x + d, y)
Time (t)
Stereo d
Left flow
Right flow
(x + u, y + v)
Time (t + 1)
(x + u + d + p, y + v)

Solution: use scene flow to improve stereo matching

stereo matching VS scene flow

- dense correspondence
- spatial principal difficulties

  matching ambiguities, occlusion, illumination…
- data information

  spatial ***VS*** spatial & temporal

  still ***VS*** still & 3D motion

  (Traditional stereo matching don’t share information between frames)

KITTI: optical flow 2015

	Method	Data	Code	D1-bg	D1-fg	D1-all	D2-bg	D2-fg	D2-all	Fl-bg	Fl-fg	Fl-all	SF-bg	SF-fg	SF-all	Density	Time	Environment	C
1	PRSM		code	3.02	10.52	4.27	5.13	15.11	6.79	5.33	17.02	7.28	6.61	23.60	9.44	99.99	300 s	1 core @ 2.5 Ghz (C/C++)	
C. Vogel, K. Schindler and S. Roth: 3D Scene Flow Estimation with a Piecewise Rigid Scene Model. ijcv 2015.																			
2	OSF		code	4.54	12.03	5.79	5.45	19.41	7.77	5.62	22.17	8.37	7.01	28.76	10.63	100.00	50 min	1 core @ 2.5 Ghz (C/C++)	
M. Menze and A. Geiger: Object Scene Flow for Autonomous Vehicles. Conference on Computer Vision and Pattern Recognition (CVPR) 2015.																			
3	SFFG			4.57	13.04	5.98	7.92	20.76	10.06	10.40	30.33	13.71	12.21	36.97	16.33	99.99	80 s	1 core @ 2.5 Ghz (C/C++)	
Anonymous submission																			
4	PR-Sceneflow		code	4.74	13.74	6.24	11.14	20.47	12.69	11.73	27.73	14.39	13.49	33.72	16.85	100.00	150 s	4 core @ 3.0 Ghz (Matlab + C/C++)	
C. Vogel, K. Schindler and S. Roth: Piecewise Rigid Scene Flow. ICCV 2013.																			
5	SGM+SF			5.15	15.29	6.84	14.10	23.13	15.60	20.91	28.90	22.24	23.09	37.12	25.43	100.00	45 min	16 core @ 3.2 Ghz (C/C++)	
H. Hirschmüller: Stereo Processing by Semiglobal Matching and Mutual Information. PAMI 2008. M. Hornacek, A. Fitzgibbon and C. Rother: SphereFlow: 6 DoF Scene Flow from RGB-D Pairs. CVPR 2014.																			
6	PCSF			6.31	19.24	8.46	19.15	36.27	22.00	14.89	62.42	22.80	25.77	69.35	33.02	100.00	0.09 s	GPU @ 2.0 Ghz (C/C++)	
Anonymous submission																			
7	SGM+C+NL		code	5.15	15.29	6.84	28.77	25.65	28.25	34.24	45.40	36.10	38.21	53.04	40.68	100.00	4.5 min	1 core @ 2.5 Ghz (C/C++)	
H. Hirschmüller: Stereo Processing by Semiglobal Matching and Mutual Information. PAMI 2008. D. Sun, S. Roth and M. Black: A Quantitative Analysis of Current Practices in Optical Flow Estimation and the Principles Behind Them. IJCV 2013.																			
8	SGM+LDOF		code	5.15	15.29	6.84	29.58	23.48	28.56	40.81	35.42	39.91	43.99	44.79	44.12	100.00	86 s	1 core @ 2.5 Ghz (C/C++)	
H. Hirschmüller: Stereo Processing by Semiglobal Matching and Mutual Information. PAMI 2008. T. Brox and J. Malik: Large Displacement Optical Flow: Descriptor Matching in Variational Motion Estimation. PAMI 2011.																			
9	HWBSF			19.61	22.69	20.12	35.72	28.15	34.46	40.74	35.53	39.87	46.42	43.99	46.02	100.00	7 min	4 cores @ 3.5 Ghz (C/C++)	
Anonymous submission																			
10	GCSF		code	11.64	27.11	14.21	32.94	35.77	33.41	47.38	45.08	47.00	52.92	59.11	53.95	100.00	2.4 s	1 core @ 2.5 Ghz (C/C++)	
J. Cech, J. Sanchez-Riera and R. Horaud: Scene Flow Estimation by growing Correspondence Seeds. CVPR 2011.																			
11	VSF		code	27.31	21.72	26.38	59.51	44.93	57.08	50.06	47.57	49.64	67.69	64.03	67.08	100.00	125 min	1 core @ 2.5 Ghz (C/C++)	
F. Huguet and F. Devernay: A Variational Method for Scene Flow Estimation from Stereo Sequences. ICCV 2007.																			

KITTI: scene flow 2015

	Method	Data	Code	Fl-bg	Fl-fg	Fl-all	Density	Time	Environment	Compare
1	PRSM		code	**5.33 %**	**17.02 %**	**7.28 %**	100.00 %	300 s	1 core @ 2.5 Ghz (C/C++)	
C. Vogel, K. Schindler and S. Roth: 3D Scene Flow Estimation with a Piecewise Rigid Scene Model. ijcv 2015.										
2	OSF		code	5.62 %	22.17 %	8.37 %	100.00 %	50 min	1 core @ 2.5 Ghz (C/C++)	
M. Menze and A. Geiger: Object Scene Flow for Autonomous Vehicles. Conference on Computer Vision and Pattern Recognition (CVPR) 2015.										
3	SFFG			10.40 %	30.33 %	13.71 %	100.00 %	80 s	1 core @ 2.5 Ghz (C/C++)	
Anonymous submission										
4	PR-Sceneflow		code	11.73 %	27.73 %	14.39 %	100.00 %	150 s	4 core @ 3.0 Ghz (Matlab + C/C++)	
C. Vogel, K. Schindler and S. Roth: Piecewise Rigid Scene Flow. ICCV 2013.										
5	SOF			14.63 %	27.73 %	16.81 %	100.00 %	6 min	1 core @ 2.5 Ghz (Matlab)	
L. Sevilla-Lara, D. Sun, V. Jampani and M. Black: Optical Flow with Semantic Segmentation and Localized Layers. CVPR 2016.										
6	GV			14.84 %	28.50 %	17.11 %	100.00 %	x	4 core @ 2.5 ghz	
Anonymous submission										
7	MR-Flow			19.42 %	27.65 %	20.79 %	100.00 %	12 s	1 core @ 2.5 Ghz (Python + C/C++)	
Anonymous submission										
8	PatchBatch			19.98 %	30.24 %	21.69 %	100.00 %	50 s	GPU @ 2.5 Ghz (Python)	
Anonymous submission										
9	SGM+SF			20.91 %	28.90 %	22.24 %	100.00 %	45 min	16 core @ 3.2 Ghz (C/C++)	
H. Hirschmüller: Stereo Processing by Semiglobal Matching and Mutual Information. PAMI 2008. M. Hornacek, A. Fitzgibbon and C. Rother: SphereFlow: 6 DoF Scene Flow from RGB-D Pairs. CVPR 2014.										
10	DiscreteFlow		code	21.53 %	26.68 %	22.38 %	100.00 %	3 min	1 core @ 2.5 Ghz (Matlab + C/C++)	
M. Menze, C. Heipke and A. Geiger: Discrete Optimization for Optical Flow. German Conference on Pattern Recognition (GCPR) 2015.										

KITTI: stereo 2015

	Method	Data	Code	D1-bg	D1-fg	D1-all	Density	Time	Environment	Compare
1	Displets v2		code	3.00 %	5.56 %	3.43 %	100.00 %	265 s	>8 cores @ 3.0 Ghz (Matlab + C/C++)	
	F. Guney and A. Geiger: Displets: Resolving Stereo Ambiguities using Object Knowledge. Conference on Computer Vision and Pattern Recognition (CVPR) 2015.									
2	MC-CNN-acrt		code	2.89 %	8.88 %	3.89 %	100.00 %	67 s	Nvidia GTX Titan X (CUDA, Lua/Torch7)	
	J. Zbontar and Y. LeCun: Stereo Matching by Training a Convolutional Neural Network to Compare Image Patches. Submitted to JMLR .									
3	CNN-SPS			3.30 %	7.92 %	4.07 %	100.00 %	80 s	GPU @ 2.5 Ghz (C/C++)	
	Anonymous submission									
4	PRSM		code	3.02 %	10.52 %	4.27 %	99.99 %	300 s	1 core @ 2.5 Ghz (C/C++)	
	C. Vogel, K. Schindler and S. Roth: 3D Scene Flow Estimation with a Piecewise Rigid Scene Model. ijcv 2015.									
5	DispNetC			4.32 %	4.41 %	4.34 %	100.00 %	0.06 s	Nvidia GTX Titan X (Caffe)	
	N. Mayer, E. Ilg, P. Häusser, P. Fischer, D. Cremers, A. Dosovitskiy and T. Brox: A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. CVPR 2016.									
6	SGM+CNN			3.93 %	10.56 %	5.04 %	100.00 %	2 s	Nvidia GTX 970	
	Anonymous submission									
7	EEL			3.86 %	11.16 %	5.07 %	99.99 %	5 s	1 core @ 2.5 Ghz (C/C++)	
	Anonymous submission									
8	SPS-St		code	3.84 %	12.67 %	5.31 %	100.00 %	2 s	1 core @ 3.5 Ghz (C/C++)	
	K. Yamaguchi, D. McAllester and R. Urtasun: Efficient Joint Segmentation, Occlusion Labeling, Stereo and Flow Estimation. ECCV 2014.									
9	MDP			4.19 %	11.25 %	5.36 %	100.00 %	11.4 s	4 cores @ 3.5 Ghz (Matlab + C/C++)	
	Anonymous submission									
10	CNN-MS			3.89 %	13.28 %	5.45 %	100.00 %	3 min	GPU @ TITAN X (Lua/Torch)	
	Anonymous submission									

Solution: CNN & Scene Flow

Stereo Net:

Zbontar J, LeCun Y. Computing the stereo matching cost with a convolutional neural network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 1592-1599.

Optical flow Net:

Dosovitskiy A, Fischer P, Ilg E, et al. FlowNet: Learning Optical Flow With Convolutional Networks[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 2758-2766.

Dataset

Mayer N, Ilg E, Häusser P, et al. A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation[J]. arXiv preprint arXiv:1512.02134, 2015.(http://vision.in.tum.de/research/deeplearning)