Paper ID #851

Asymmetric Event-Image Stereo with Temporal Feature Gating and Iterative Structure-Detail Refinement

Hao Zhuang[1], Yan Yang[2], Liyuan Pan[1†]

[1]Beijing Institute of Technology, [2]Australian National University

Introduction

Our task

Estimate disparity from event–frame stereo pairs.

Key words

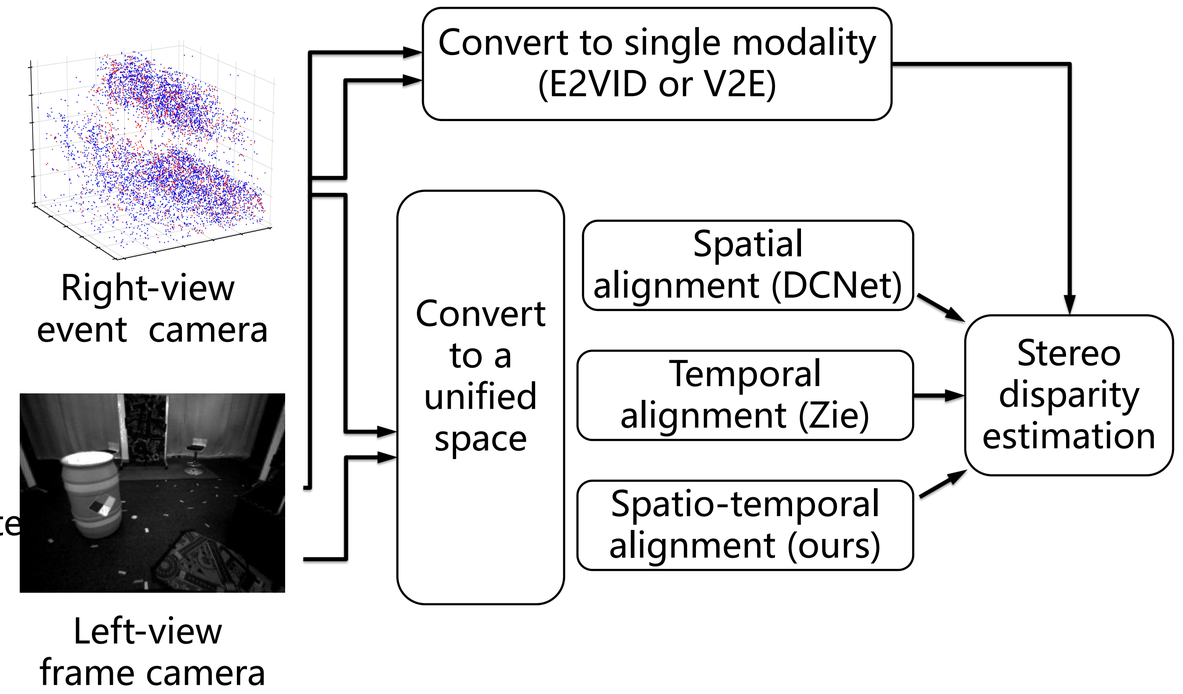
Depth estimation, Event camera, Cross-modality correspondence.

Motivation

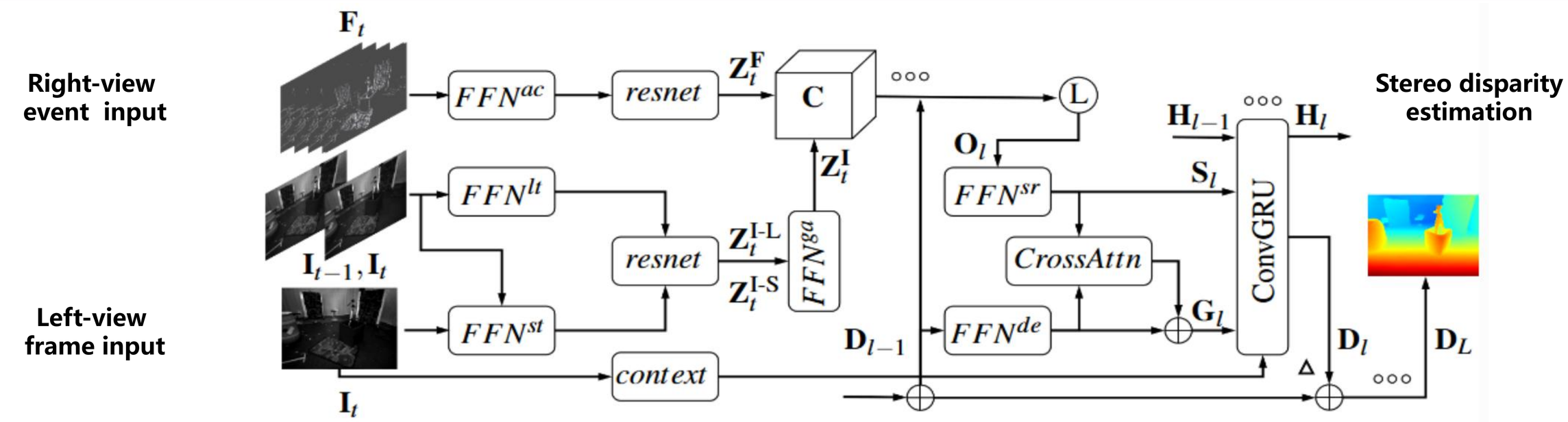
(1) How to bridge the domain gap between events and frames?
(2) How to exploit structural cues from events and fine details from frames ?

Contributions

(1) We introduce temporal-gated frame enhancement for more accurate cross-modality correspondence.
(2) We design an iterative structure-detail disparity refinement scheme for accurate depth estimation.



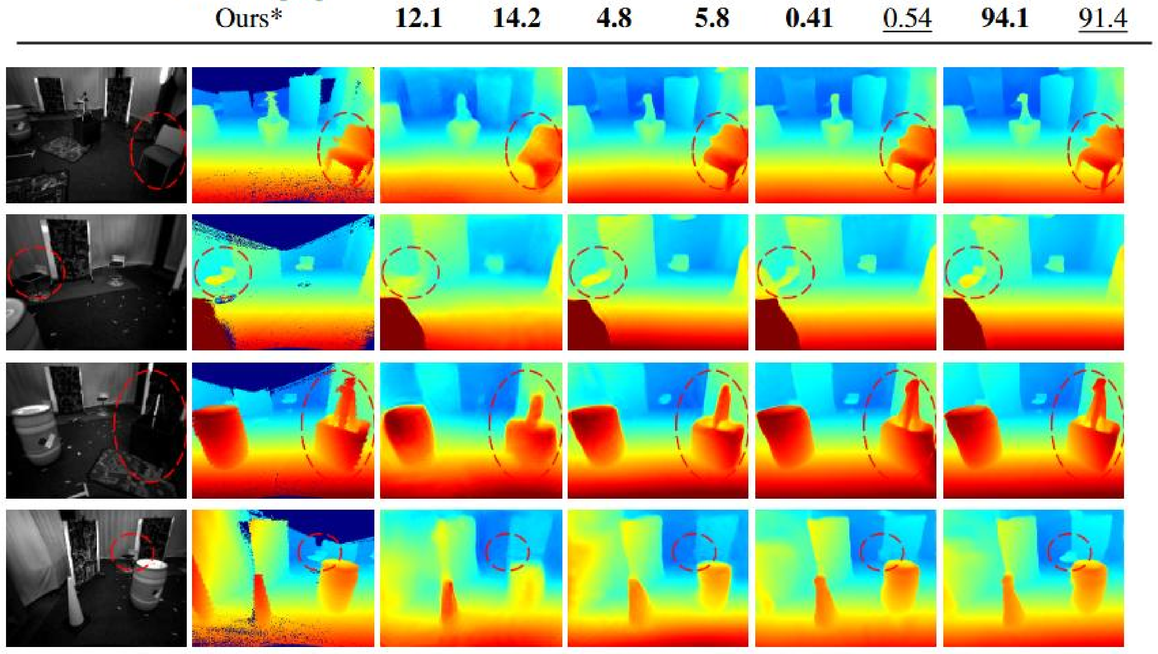
Method



Results

MVSEC dataset

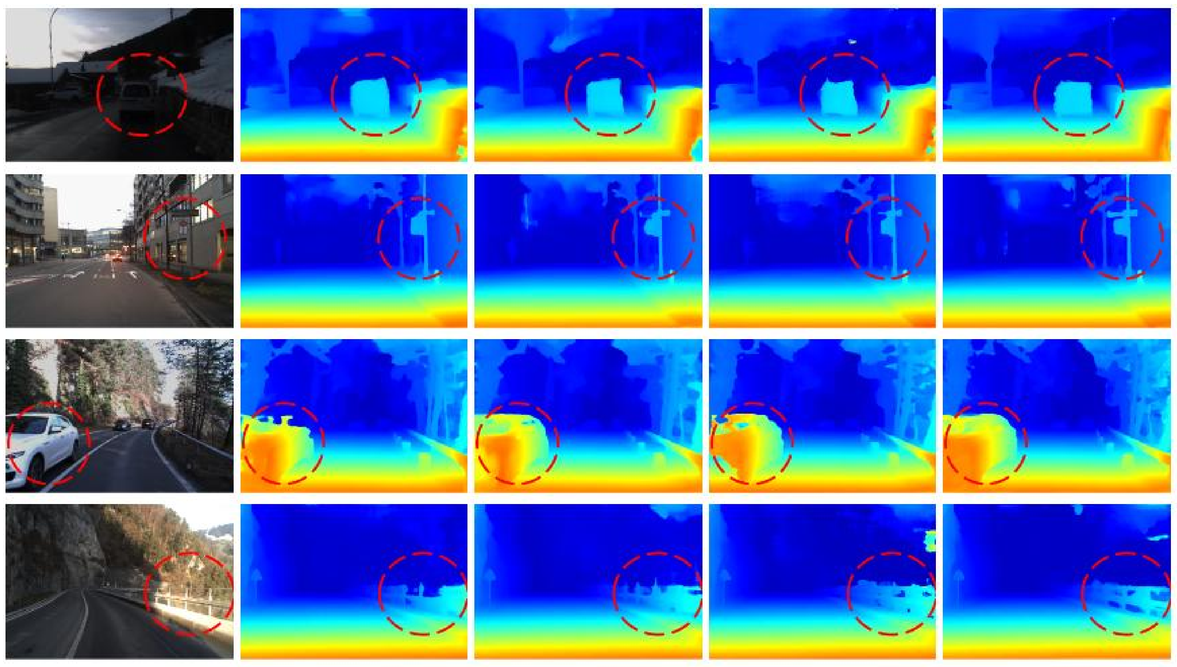
Modality	Method	MDE [cm] ↓		MdDE [cm] ↓		MDispE [pix] ↓		1PA [%] ↑	
		Split 1	Split 3	Split 1	Split 3	Split 1	Split 3	Split 1	Split 3
2E+2F	EIS [22]	13.7	22.4	–	–	–	–	89.0	88.1
	SMC-Net [5]	11.2	14.5	4.5	6.3	0.37	0.52	94.3	92.0
	EGEI-Net [38]	9.7	11.1	3.9	4.7	0.35	0.38	95.1	93.4
	Ours	10.1	11.3	3.9	4.5	0.35	0.38	94.9	93.4
2F	PSMNet [3]	15.9	18.3	8.0	10.2	0.57	0.68	88.6	83.1
	GWC-Net [14]	15.0	17.4	7.5	9.3	0.53	0.64	89.9	85.8
	SStereo [29]	14.0	16.4	6.4	7.8	0.48	0.56	91.5	88.5
2E	DDES [27]	16.7	27.8	6.8	14.7	0.59	0.94	89.8	74.8
	EITNet [1]	14.2	19.4	5.9	10.4	0.55	0.75	92.1	89.6
	DTC-SPADE [36]	13.5	17.1	5.2	7.9	0.46	0.60	93.0	89.7
	StereoFlow-Net* [7]	13.0	15.0	5.0	**5.8**	0.46	**0.49**	92.9	**92.6**
1E+1F	HDES [41]	16.0	18.0	–	–	–	–	86.4	80.1
	DCNet [30]	16.9	17.8	7.0	8.6	0.66	0.75	85.0	83.3
	TDANet [35]	15.8	19.7	–	–	0.75	0.87	88.1	77.4
	Ours*	**12.1**	**14.2**	**4.8**	**5.8**	**0.41**	0.54	**94.1**	91.4



(a) Frame (b) Ground Truth (c) DCNet (d) DCT-SPADE (e) StereoFlow-Net (f) Ours

DSEC dataset

Modality	Method	MAE ↓	1PE ↓	2PE ↓	RMSE ↓
2E+2F	EIS [22]	0.396	5.814	1.055	0.906
	SCSNet [6]	0.387	5.673	0.992	0.852
2E	DDES [27]	0.576	10.915	2.905	1.381
	EIS-E [22]	0.529	9.958	2.645	1.222
	DTC-SPADE [36]	0.527	9.277	2.416	1.291
	Se-cff [23]	0.519	9.583	2.620	1.231
	StereoFlow-Net* [7]	0.493	**8.662**	2.259	1.172
	MGC-Net [16]	0.514	9.625	2.512	1.187
1E+1F	HDES [41]	0.698	19.020	4.970	1.307
	DCNet [30]	0.587	13.050	2.850	1.193
	AFES-dc* [4]	0.543	10.970	**2.230**	1.175
	Ours*	**0.488**	8.881	2.335	**1.168**



(a) Frame (b) Se-cff (c) MGC-Net (d) StereoFlow-Net (e) Ours