ICCV 2021

Fast Light-field Disparity Estimation with Multi-disparity-scale Cost Aggregation

Zhicong Huang[1,2], Xuemei Hu[1], Zhou Xue[2], Weizhu Xu[1], Tao Yue[1]
Nanjing University[1], ByteDance Inc.[2]

Challenge

- Sub-pixel Accuracy
- High Precision
- Heavy Computation
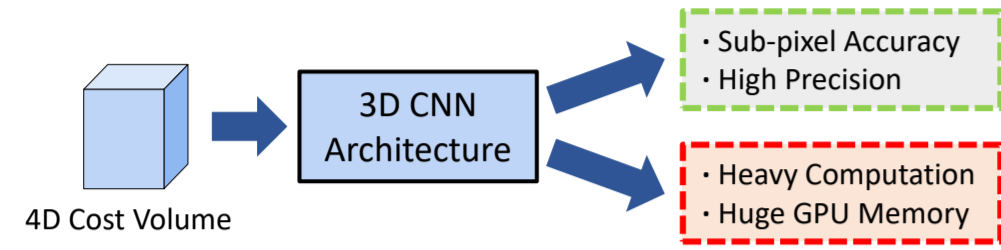- Huge GPU Memory

4D Cost Volume → 3D CNN Architecture

Problem statement

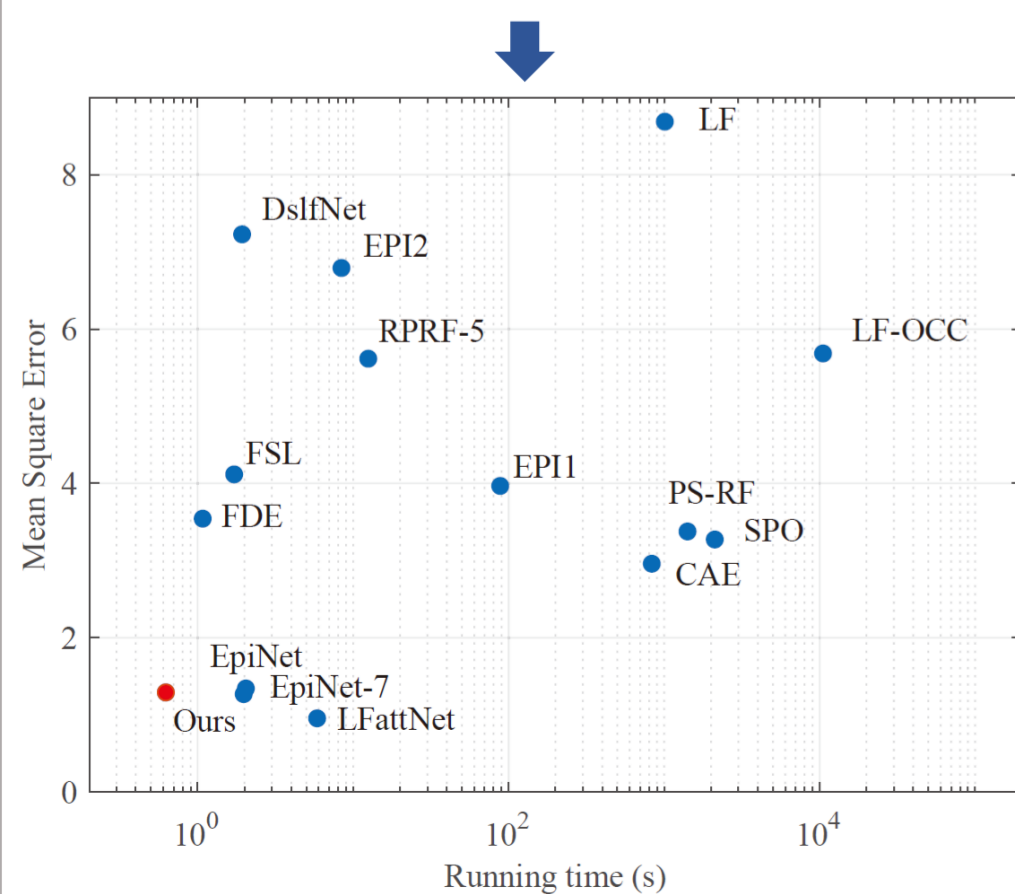Can we have the best of both worlds and reduce computation cost while achieving high performance?

Contribution

- We propose a fast and lightweight end-to-end network for light field disparity estimation.
- We present a physical-based multi-disparity-scale network for fast and high-performance cost volume regularization.
- We design an edge guidance sub-network to guide the disparity estimation with edge cues for better performance on challenging regions.
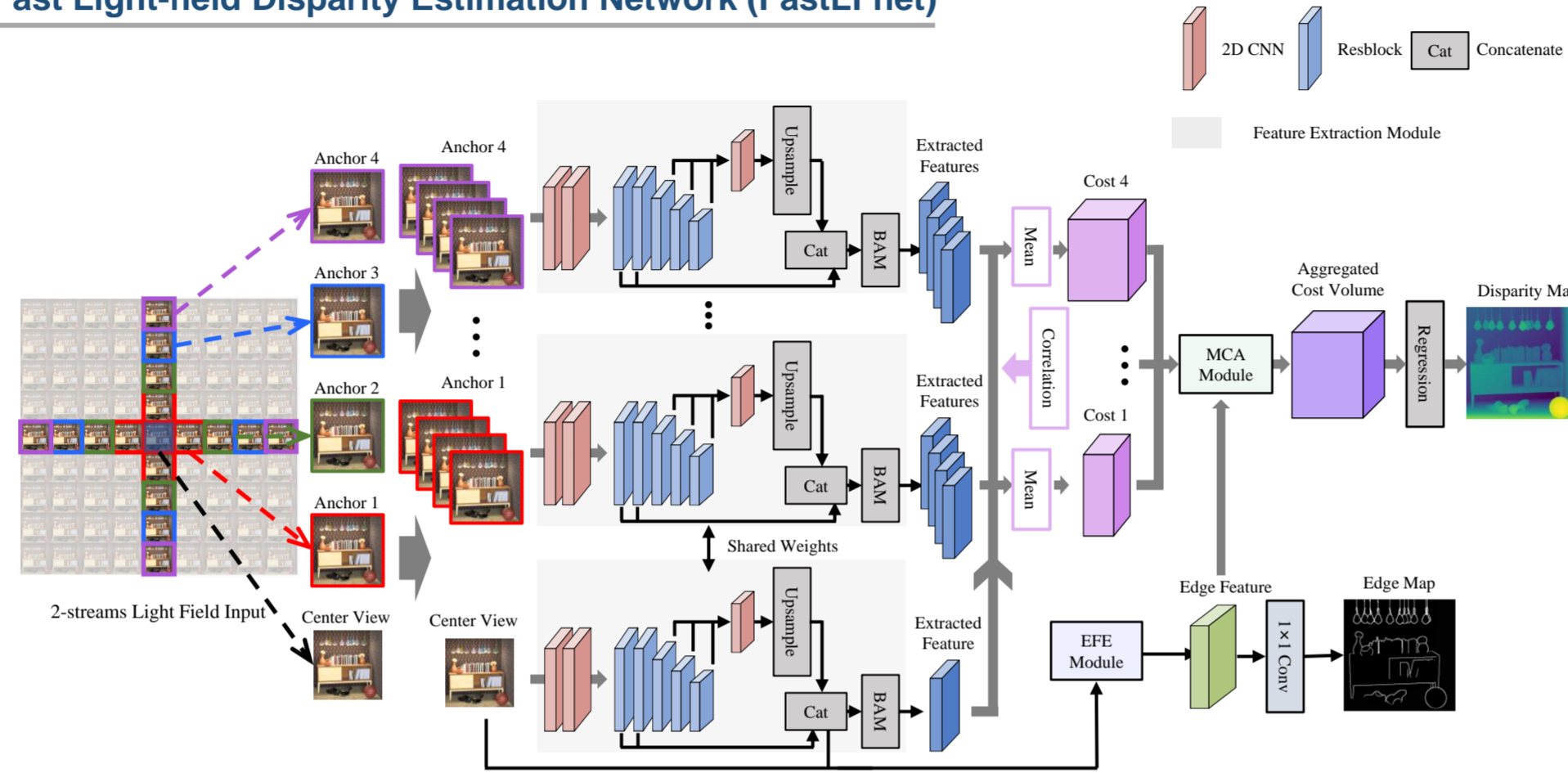
$$C(d,h,w) = \frac{1}{N}\left\langle F_c(c,h,w), f_{warp}[F_s(c,h,w),d]\right\rangle$$

Only use 2-streams input images   Abandon 3D CNN architecture
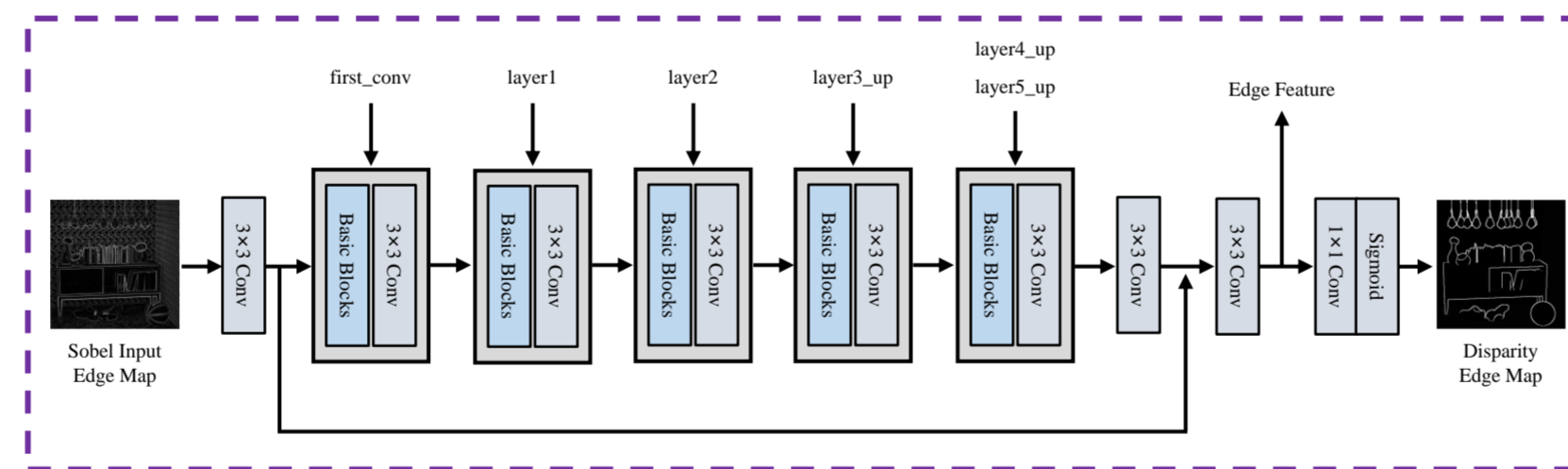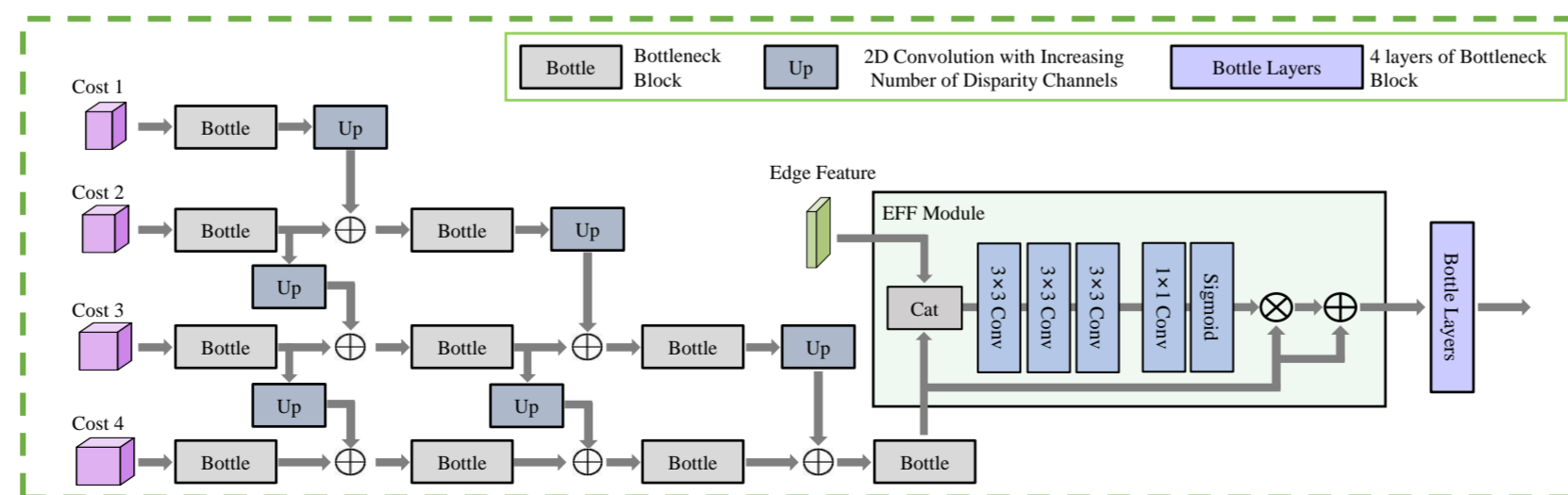

Method

Fast Light-field Disparity Estimation Network (FastLFnet)

Edge Feature Extraction (EFE) Module

Multi-disparity-scale Cost Aggregation (MCA) Module
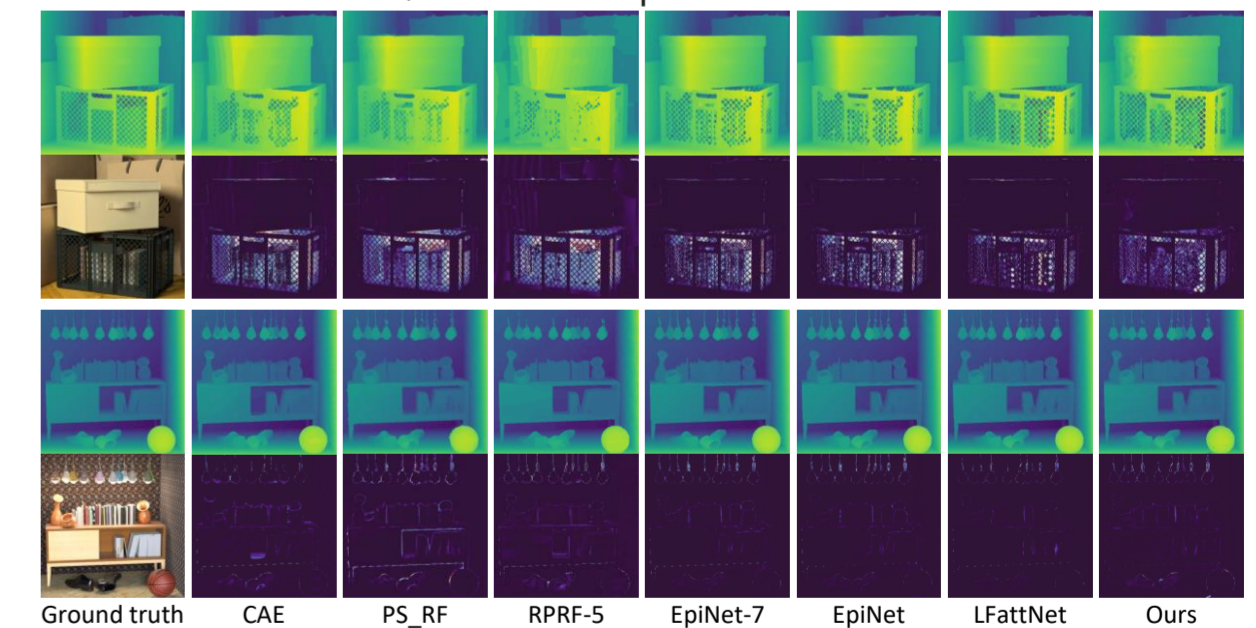

Experiments

4D light field dataset

- Comparison of the contributions of each component

Processing methods	One Scale	w/o BAM	w/o Edge	FastLFnet
MSE_x100	1.650	1.492	1.844	1.218
Running time / s	0.725	0.582	0.576	0.593
GPU Memory / GB	2.323	2.103	2.189	2.107
Parameters / M	1.281	1.361	0.982	1.366

- Quantitative comparison with state-of-the-art methods

Methods	CAE [21]	PS_RF [11]	RPRF-5 [8]	EpiNet-7 [27]	EpiNet [27]	LFattNet [29]	w/o Edge	Fast-LFnet
Boxes	8.162	8.771	10.333	6.042	5.845	**3.869**	5.658	4.260
Cotton	1.704	1.227	0.949	**0.206**	0.235	0.220	0.318	0.339
Dino	0.376	0.730	0.603	0.162	0.147	**0.090**	0.350	0.184
Sideboard	0.860	1.899	1.224	0.814	0.794	0.518	1.070	0.742
Backgammon	4.762	5.559	3.024	1.500	1.893	1.762	2.658	**1.488**
Dots	4.589	7.881	20.114	1.155	1.549	**0.959**	4.508	3.070
Pyramids	0.047	0.043	0.042	0.008	0.007	**0.004**	0.010	0.018
Stripes	3.171	0.905	8.643	0.265	0.264	**0.220**	0.854	0.231
Average	2.959	3.377	5.616	1.269	1.342	**0.955**	1.928	1.291
Fattening	7.614	6.597	5.262	4.702	4.990	**3.810**	5.752	4.300
Thinning	**1.153**	2.237	2.568	1.548	1.430	2.230	3.499	2.427
Running time / s	832.081	1412.623	12.498	1.976	2.041	5.862	**0.611**	0.624
GPU Memory / GB	-	-	-	4.319	5.103	10.953	2.189	**2.107**
Parameters / M	-	-	-	5.116	5.118	5.058	**0.982**	1.366

- Qualitative comparison results

Ground truth | CAE | PS_RF | RPRF-5 | EpiNet-7 | EpiNet | LFattNet | Ours

Sparse light field dataset

- Results of the performance comparison

Light fields	MSE					
	EBSM [9]	OHLF [13]	SflfNet [6]	EpiNet [27]	DslfNet [26]	FastLFnet
Furniture	0.37	1.94	9.18	1.73	0.42	**0.17**
Lion	0.10	0.87	1.59	3.41	0.09	**0.05**
Toy_bricks	0.22	1.10	3.70	0.36	0.57	**0.16**
Elec_dev	0.20	0.63	7.82	0.74	0.20	**0.09**
Average	0.22	1.14	5.57	1.56	0.32	**0.12**