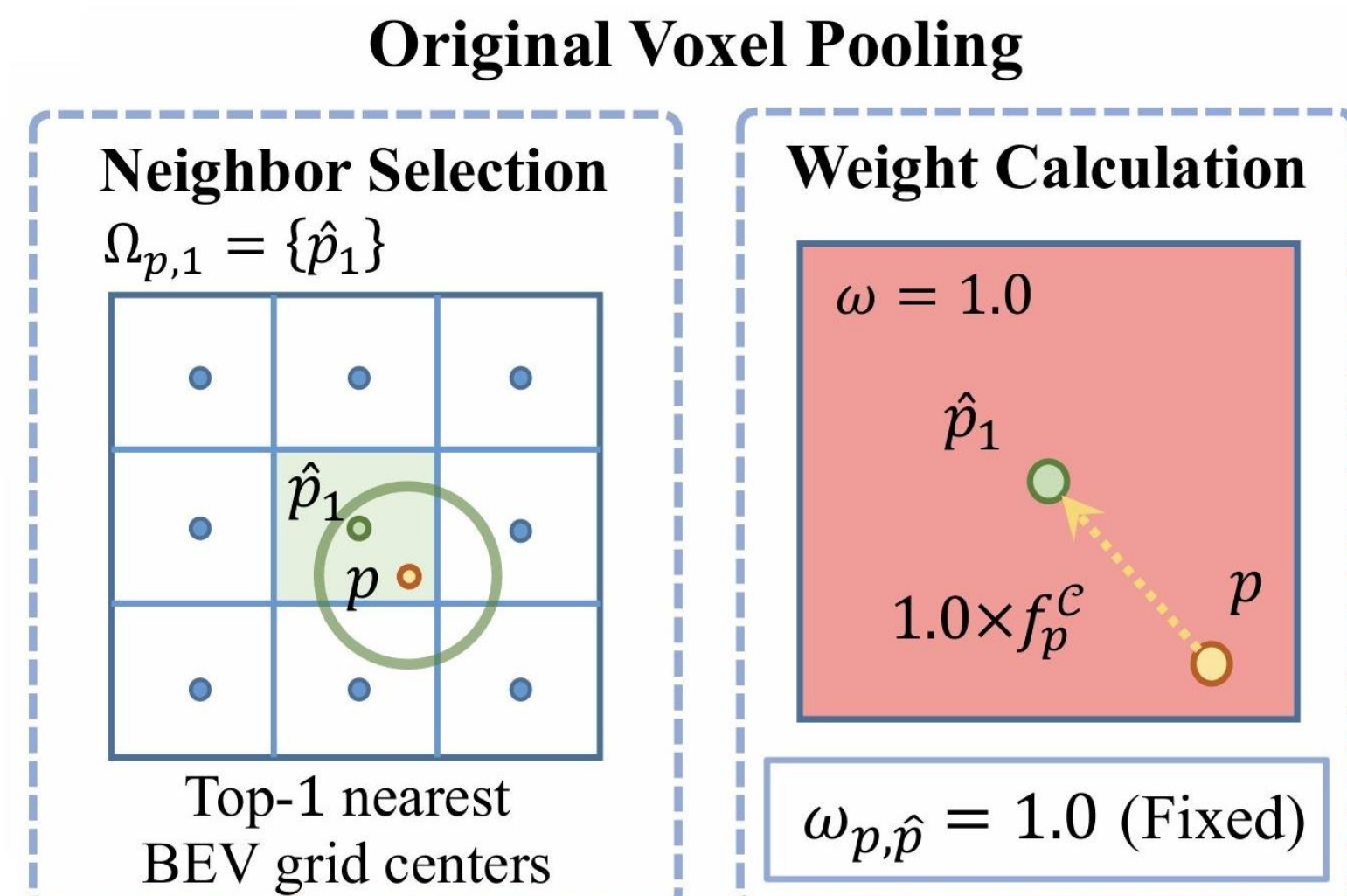


Wenjie Wang^{1*}, Yehao Lu^{1*}, Guangcong Zheng¹, Shuigen Zhan², Xiaoqing Ye³, ZichangTan³, Jingdong Wang³, GaoangWang¹, and Xi Li^{1,2,4*}

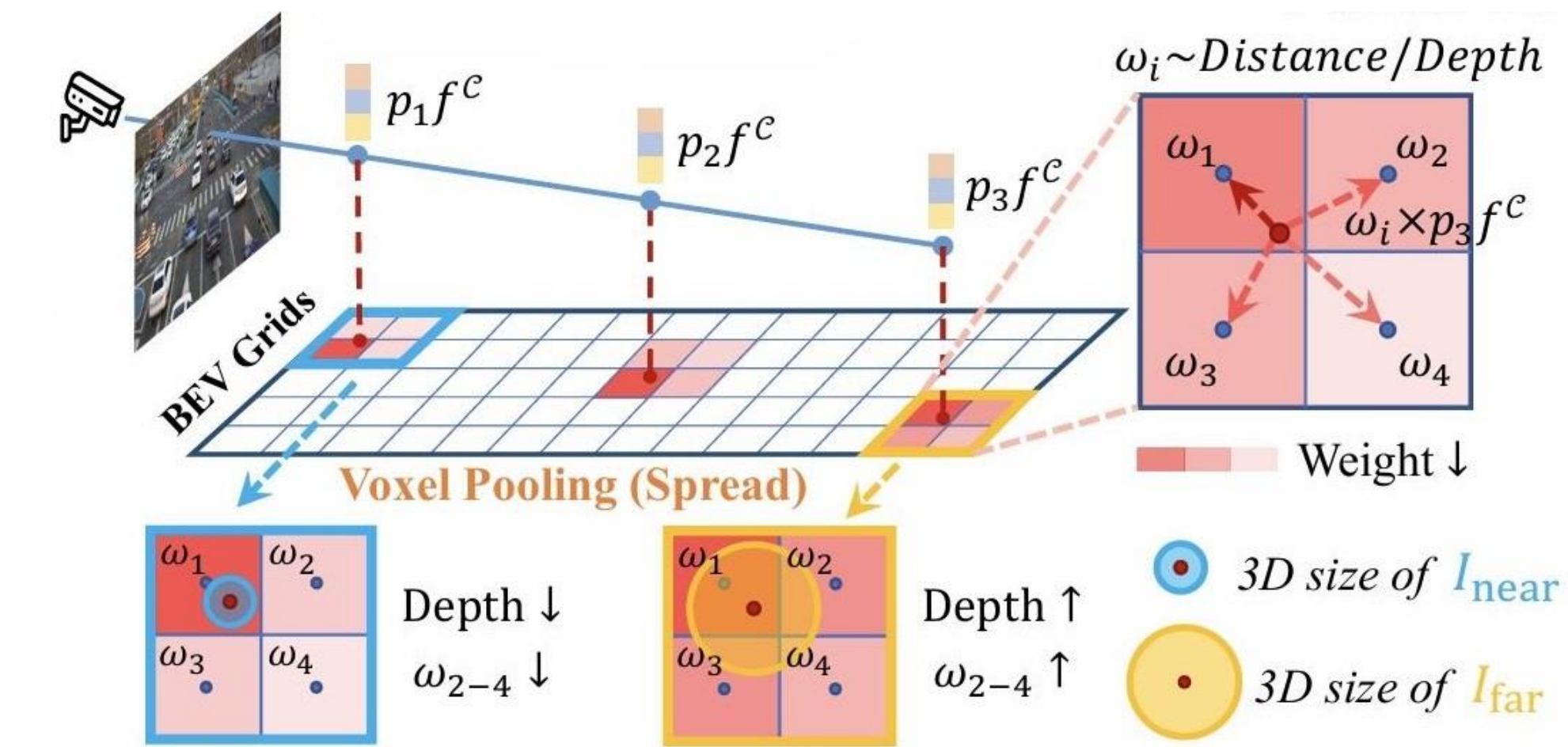
¹College of Computer Science & Technology, Zhejiang University ²Polytechnic Institute, Zhejiang University ³Baidu ⁴Zhejiang – Singapore Innovation and AI Joint Research Lab

Problems

- ◆ **Keywords**
 - Autonomous Driving Vehicle to Infrastructure
 - Roadside 3D Object Detection
- ◆ **Motivation**
 - Approximation Error in Voxel Pooling
 - Limited Computing Resources in Roadside Unit



Spread Voxel Pooling vs. Original Voxel Pooling



Objectives

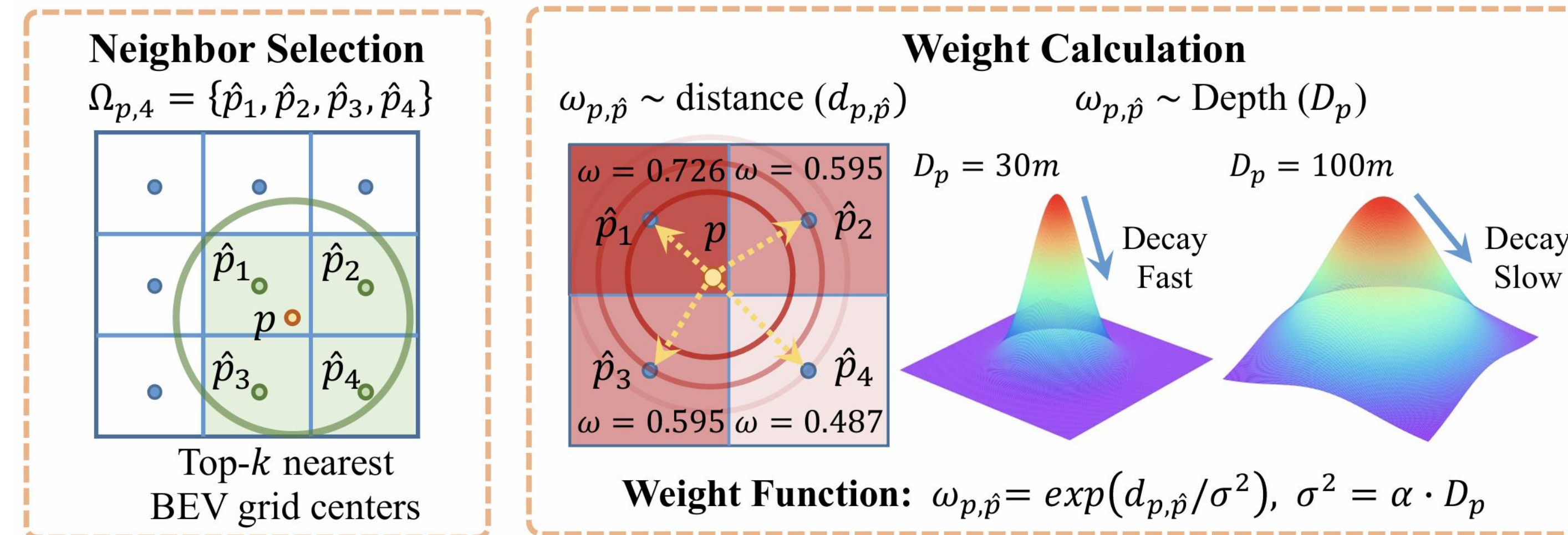
- Reduction of the Approximation Error
- Scale up Features of Distant Objects

Methods

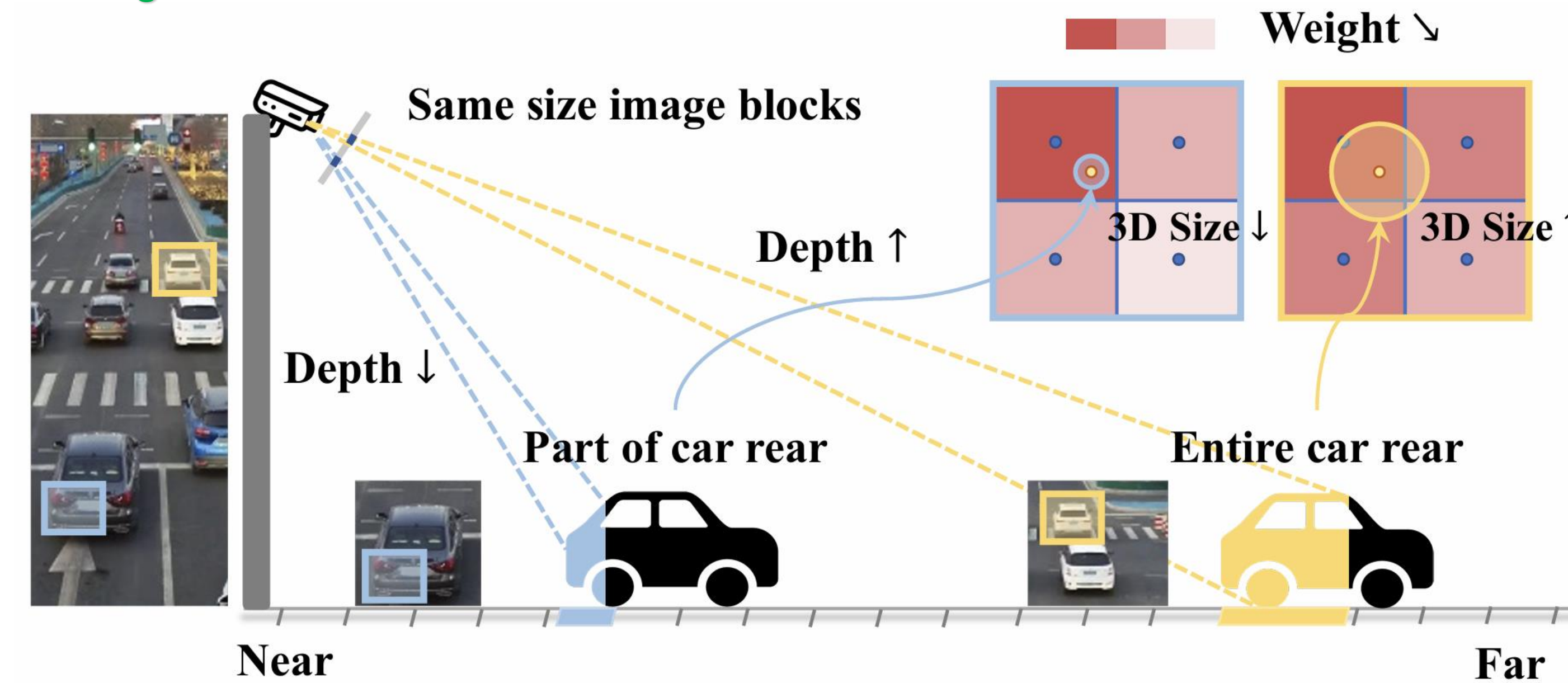
Spread Strategy in Voxel Pooling Process

- Selection of top-K nearest BEV grids
- Assignment of dynamic weights for these selected grids

Spread Voxel Pooling



Weight Function



$$\omega_{p,\hat{p}} = \exp\left(-\frac{d_{p,\hat{p}}^2}{\sigma^2}\right)$$

$$\sigma^2 = \alpha \cdot D_p$$

□ Code: DaTongjie/BEVSpread (github.com)



Results

Dataset	Method	Modality	Venue	Vehicle (IoU=0.5)			Pedestrian (IoU=0.25)			Cyclist (IoU=0.25)		
				Easy	Middle	Hard	Easy	Middle	Hard	Easy	Middle	Hard
DAIR-V2X-I [46]	PointPillars [12]	LiDAR	CVPR' 19	63.07	54.00	54.01	38.53	37.20	37.28	38.46	22.60	22.49
	SECOND [42]	LiDAR	Sensors	71.47	53.99	54.00	55.16	52.49	52.52	54.68	31.05	31.19
	MVXNet [32]	LiDAR & Camera	ICRA' 19	71.04	53.71	53.76	55.83	54.45	54.40	54.05	30.79	31.06
	ImVoxelNet [28]	Camera	WACV' 22	44.78	37.58	37.55	6.81	6.75	6.74	21.06	13.57	13.17
	BEVFormer [17]	Camera	ECCV' 22	61.37	50.73	50.73	16.89	15.82	15.95	22.16	22.13	22.06
	BEVDepth [16]	Camera	AAAI' 23	75.50	63.58	63.67	34.95	33.42	33.27	55.67	55.47	55.34
	BEVHeight [44]	Camera	CVPR' 23	77.78	65.77	65.85	41.22	39.29	39.46	60.23	60.08	60.54
BEVSpread (Ours)	Camera	-	79.07	66.82	66.88	46.54	44.51	44.71	62.64	63.50	63.75	
w.r.t. BEVHeight			+1.29	+1.05	+1.03	+5.32	+5.22	+5.25	+2.41	+3.42	+3.21	
DAIR-V2X-I* [46]	BEVHeight [44]	Camera	CVPR 23	81.62	75.90	75.94	40.89	38.98	39.18	60.29	60.60	61.13
	BEVSpread (Ours)	Camera	-	82.84	77.10	77.19	43.96	42.03	42.13	62.31	64.44	64.89
	w.r.t. BEVHeight			+1.22	+1.21	+1.25	+3.07	+3.05	+2.95	+2.02	+3.84	+3.76
Rope3D [45]	BEVDepth [16]	Camera	AAAI 23	76.90	66.91	66.89	30.42	28.08	28.11	55.34	53.53	53.51
	BEVHeight [44]	Camera	CVPR 23	77.93	67.50	67.49	36.26	30.35	30.30	61.49	56.98	56.90
	BEVSpread (Ours)	Camera	-	80.61	70.04	70.03	38.65	34.32	34.25	63.66	59.11	59.03
	w.r.t. BEVHeight			+2.69	+2.55	+2.54	+2.39	+3.97	+3.95	+2.17	+2.13	+2.13

