



REVERIE: Remote Embodied Visual Referring Expressions in Real Indoor Environments

Yuankai Qi^{1,2}, Qi Wu¹, Peter Anderson³, Xin Wang⁴, William Yang Wang⁴,
Chunhua Shen¹, Anton van den Hengel¹

¹*Australian Centre for Robotic Vision, The University of Adelaide, Australia*

²*Harbin Institute of Technology, Weihai, China*

³*Georgia Institute of Technology, USA* ⁴*University of California, Santa Barbara, USA*

The REVERIE Task



Results

Results on the REVERIE Task Using MAttNet as Pointer

Methods	Val Seen					Val UnSeen					Test (Unseen)				
	Navigation Acc.				RGS	Navigation Acc.				RGS	Navigation Acc.				RGS
	Succ.	OSucc.	SPL	Length		Succ.	OSucc.	SPL	Length		Succ.	OSucc.	SPL	Length	
Random	2.74	8.92	1.91	11.99	1.97	1.76	11.93	1.01	10.76	0.96	2.30	8.88	1.44	10.34	1.18
Shortest	100	100	100	10.46	68.45	100	100	100	9.47	56.63	100	100	100	9.39	48.98
R2R-TF [1]	7.38	10.75	6.40	11.19	4.22	3.21	4.94	2.80	11.22	2.02	3.94	6.40	3.30	10.07	2.32
R2R-SF [1]	29.59	35.70	24.01	12.88	18.97	4.20	8.07	2.84	11.07	2.16	3.99	6.88	3.09	10.89	2.00
RCM [28]	23.33	29.44	21.82	10.70	16.23	9.29	14.23	6.97	11.98	4.89	7.84	11.68	6.67	10.60	3.67
SelfMonitor [19]	41.25	43.29	39.61	7.54	30.07	8.15	11.28	6.44	9.07	4.54	5.80	8.39	4.53	9.23	3.10
FAST-Short [14]	45.12	49.68	40.18	13.22	31.41	10.08	20.48	6.17	29.70	6.24	14.18	23.36	8.74	30.69	7.07
FAST-Lan-Only	8.36	23.61	3.67	49.43	5.97	9.37	29.76	3.65	45.03	5.00	8.15	28.45	2.88	46.19	4.34
Ours	50.53	55.17	45.50	16.35	31.97	14.40	28.20	7.19	45.28	7.84	19.88	30.63	11.61	39.05	11.28
Human	—	—	—	—	—	—	—	—	—	—	81.51	86.83	53.66	21.18	77.84

Take Home Message

Code and Dataset

