Data-Driven 3D Primitives for Single Image Understanding

Addendum

David F. Fouhey, Abhinav Gupta, Martial Hebert
The Robotics Institute, Carnegie Mellon University

To facilitiate comparison with other methods, we provide results on the NYU v2 dataset using the train-test split used in Silberman et al. in the table below. This split contains 795 images for training compared to the 1086 – 1089 images available for training in each split of our 4-fold cross validation ($\approx 30\%$ less data).

	Summary Stats. (°) (Lower Better)			% Good Pixels (Higher Better)		
	Mean	Median	RMSE	11.25°	22.5°	30°
	With Manhattan World Constraints					
Lee et al.	43.3	36.3	54.6	25.1	40.4	46.1
Hedau et al.	40.0	23.5	54.1	34.2	49.3	54.4
3D Primitives.	**36.0**	**20.5**	**49.4**	**35.9**	**52.0**	**57.8**
	Without Manhattan World Constraints					
Karsch et al.	40.7	37.8	46.9	8.1	25.9	38.2
Hoiem et al.	36.0	33.4	41.7	11.4	31.3	44.5
Saxena et al.	48.0	43.1	57.0	10.7	27.0	36.3
RF + SIFT	36.0	33.4	41.7	11.4	31.4	44.5
SVR + SIFT	36.6	33.6	42.5	10.6	30.6	44.0
3D Primitives	**34.2**	**30.0**	**41.4**	**18.6**	**38.6**	**49.9**