Searching for Objects using Structure in Indoor Scenes - Supplementary Material

Varun K. Nagaraja
varun@umiacs.umd.edu

Vlad I. Morariu
morariu@umiacs.umd.edu

Larry S. Davis
lsd@umiacs.umd.edu

University of Maryland
College Park, MD. USA.

1 Experiments and Results

Table 1 shows the operating point of the region classifier. The thresholds for the detectors are set based on the best F1 point on the validation curves. We work with 18 categories and do not include the box category as its performance values are very low with an average precision of 1.4%.

																			mean
AP (%)	22.9	52.9	27.0	16.7	29.1	3.8	13.0	14.6	16.2	9.7	10.4	11.6	25.2	20.2	33.3	11.5	18.5	29.4	20.3
Prec (%)	87.5	79.8	52.7	65.1	69.8	16.7	37.8	41.7	56.0	65.7	75.0	48.0	57.5	78.6	73.6	44.1	27.8	90.9	59.3
Rec (%)	23.3	54.5	34.8	20.7	35.3	11.7	19.5	17.5	18.7	11.7	12.4	15.2	32.5	21.2	37.6	18.0	28.9	29.4	24.6
F1 (%)	36.8	64.8	41.9	31.4	46.9	13.8	25.7	24.7	28.0	19.8	21.2	23.1	41.5	33.3	49.7	25.5	28.3	44.4	34.8

Table 1: **Operating characteristics of RCNN-depth.** The table shows the performance of the region classification module RCNN-depth on the testing set of NYU depth v2 dataset. We use the top 100 region proposals and the threshold for the detectors is chosen such that we obtain the best F1 score on the validation set. RCNN-depth is called by our search strategy after exploring a region.
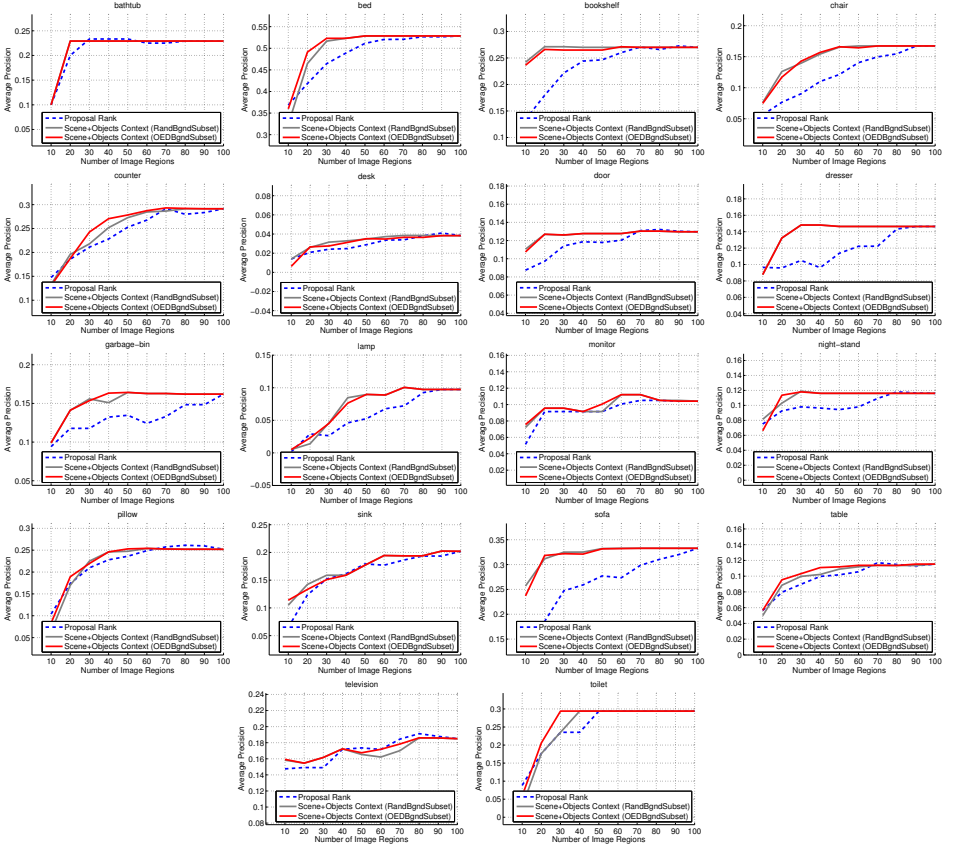
Figure 1: **Scene+Objects Context: Comparison of background selection techniques.** The search strategy trained with a background subset selected using determinant maximization performs better or equally well as the strategy trained with a background subset selected randomly. The main advantage of the determinant maximization based subset selection is the repeatability of experiments unlike the random subset selection.
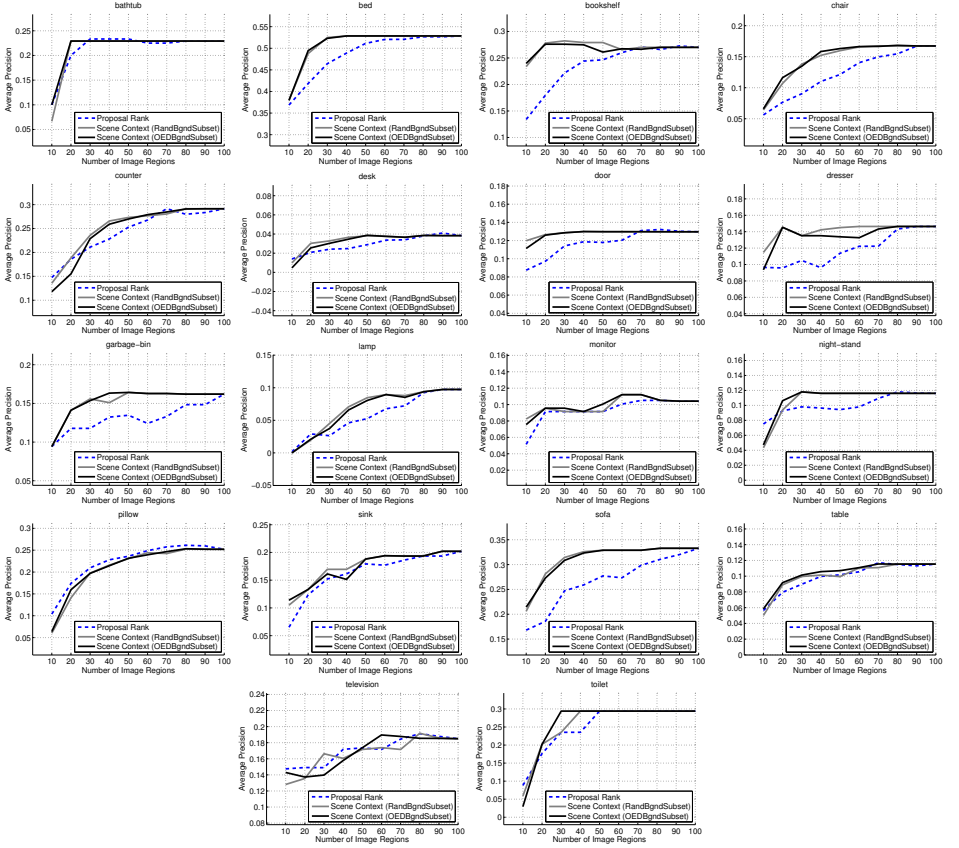
Figure 2: **Scene Context: Comparison of background selection techniques.** The performance of the classifier trained with a background subset selected using determinant maximization is comparable to that of the classifier trained with a random background subset. The main advantage of the determinant maximization based subset selection is the repeatability of experiments unlike the random subset selection.

Groundtruth	Proposal Rank	Scene Context	Scene+Objects Context

(a) Searching for lamp. Number of regions processed = 35

(b) Searching for table. Number of regions processed = 15

(c) Searching for chair. Number of regions processed = 45

(d) Searching for counter. Number of regions processed = 20
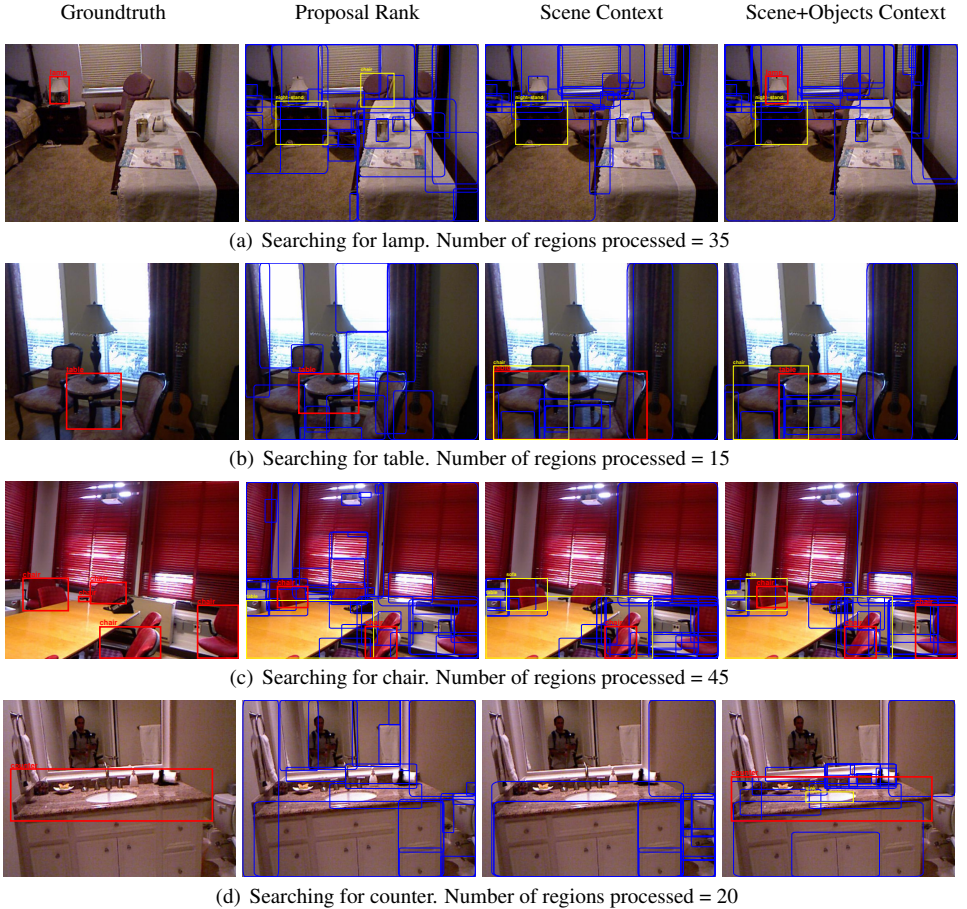
Figure 3: **Search results for different queries.** We compare three strategies - ranked sequence obtained from the region proposal technique (unaware of query class), ranked sequence obtained from a classifier trained for a query class using scene context features alone and sequence produced by a search strategy trained for a query class using both scene context and object-object context features. Red boxes indicate regions labeled as query class, yellow boxes indicate regions other than the query class and blue boxes indicate regions labeled as background. The images show a state in the search sequence of different methods at a certain number of regions processed. Our strategy which uses both scene context and object-object context can locate an object of the query class earlier than the other methods.