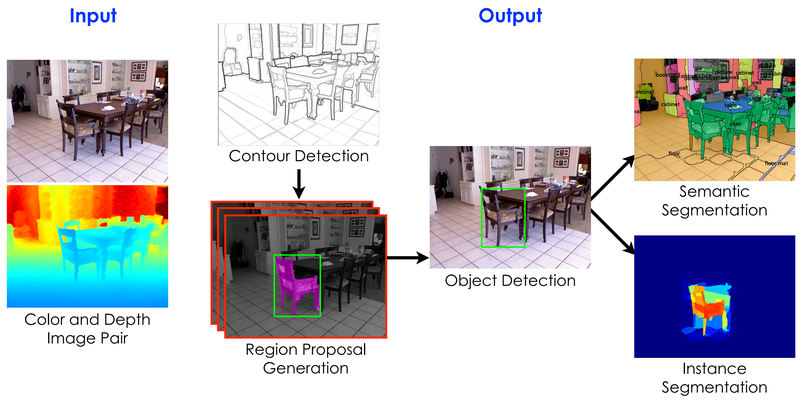
Learning Rich Features from RGB-D Images for Object Detection and Segmentation

Saurabh Gupta[1] Pablo Arbeláez[1,2] Ross Girshick[1,3] Jitendra Malik[1]

[1]UC Berkeley [2]Universidad de los Andes, Colombia [3]Microsoft Research

Overview

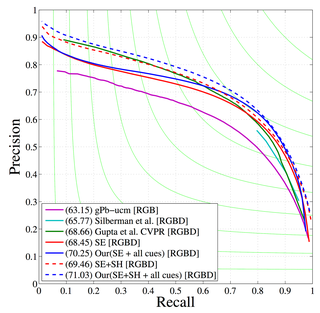


Contour Detection and Region Proposals

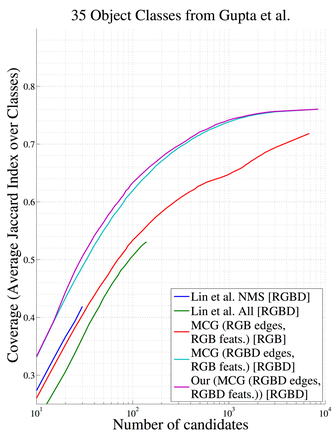
- Better contours by enriching SE [Dollár et al. (2013)] with surface normal gradients
- Better candidates by generalizing MCG [Arbeláez et al. (2014)] to RGB-D images

Improved **RGB-D** Contours → Multi-scale Combinatorial Grouping → Features from **RGB-D** image → *k* candidates

Contour Performance



Region Performance



Acknowledgements: This work was sponsored by ONR SMARTS MURI N00014-09-1-1051, ONR MURI N00014-10-1-0933 and a Berkeley Fellowship. The GPUs used in this research were generously donated by the NVIDIA Corporation. We are also thankful to Bharath Hariharan, for all the useful discussions. We also thank Piotr Dollár for helping us with their contour detection code.

[Arbeláez et al.] P. Arbeláez, J. Pont-Tuset, J. Barron, F. Marques, J. Malik **Multiscale Combinatorial Grouping**, CVPR 2014

[Dollar et al.] P. Dollár and L. Zitnick **Structured Forests for fast edge detection**, ICCV 2013

[Lin et al.] D. Lin, S. Fidler and R. Urtasun. **Holistic Scene Understanding for 3D Object Detection with RGBD cameras**, ICCV 2013

[Gupta et al.] S. Gupta, P Arbeláez, J. Malik **Perceptual Organization and Recognition in Indoor RGB-D Images**, CVPR 2013

[Hariharan et al.] B. Hariharan, P. Arbeláez, R. Girshick, J. Malik **Simultaneous Detection and Segmentation**, ECCV, 2014

[Girshick et al.] R. Girshick,J. Donahue, T. Darell, J. Malik **Rich feature hierarchies for accurate object detection and semantic segmentation**, CVPR 2014

[Silberman et al.] N. Silberman, D. Hoiem, P. Kohli, R. Fergus **Indoor segmentation and support inference from RGBD images**, ECCV 2012

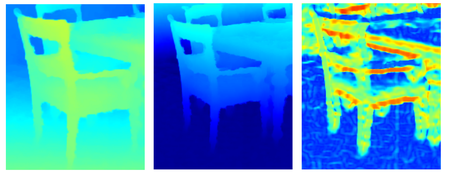
Object Detection

Learning representations from depth images for use in object detectors

Depth Images are **image-like enough** to use Convolutional Neural Network models

Geocentric embedding into *Horizontal Disparity, Height Above Ground,* and *Angle with Gravity* ***(HHA)*** works better than just raw disparity

Synthetic depth data can help



Horizontal Disparity **H** | Height Above Ground **H** | Angle with Gravity **A**

Geocentric Encoding of Depth

Performance (AP^b)

	mean	bath tub	bed	book shelf	box	chair	counter	desk	door	dresser	garbage bin	lamp	monitor	night stand	pillow	sink	sofa	table	television	toilet
RGB DPM	**9**	1	28	9	0	8	7	1	3	1	7	22	10	9	4	6	9	6	6	34
RGBD DPM	**24**	19	56	18	**1**	24	24	6	10	16	27	27	35	33	21	23	34	17	20	45
RGB RCNN	**22**	17	45	28	**1**	26	30	10	16	19	16	28	32	17	11	17	29	13	27	44
Our	**37**	**44**	**71**	**33**	**1**	**43**	**44**	**15**	**24**	**30**	**39**	**37**	**53**	**40**	**35**	**36**	**54**	**24**	**38**	**47**

Semantic Segmentation

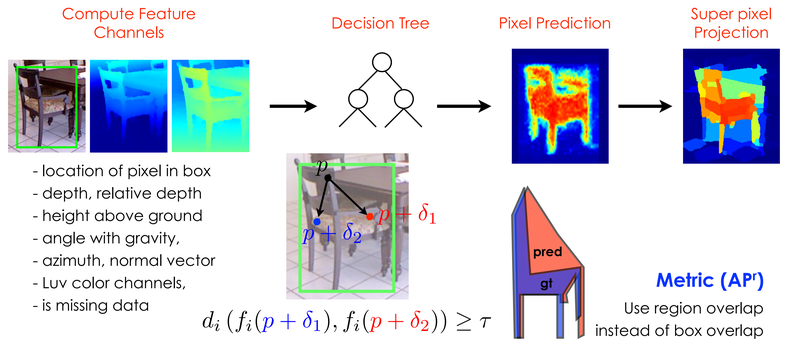
Gupta et al. (2013) + Additional superpixel features based on deep detectors

- Take detections with precision more than 50%
- Assign best scoring overlapping detection to each superpixel
- Compute features between superpixel and detection

score of detector, overlap between detector and superpixel, mean and median of depth in superpixel and detector

	Silberman et al. ECCV 12	Ren et al. CVPR 12	Gupta et al. CVPR 13	Gupta et al. + RGB-D DPM	Gupta et al. + Our
fwavacc	38.2	37.6	45.2	45.6	**47.0**
avacc	19	20.5	26.4	27.4	**28.6**
mean (maxIU)	-	21.4	29.1	30.5	**31.3**
pixacc	54.6	49.3	59.1	60.1	**60.3**
obj avg	18.4	21.1	28.4	31	**35.1**

Instance Segmentation



Metric (AP^r)

Use region overlap instead of box overlap in Pascal detection metric

	mean	bath tub	bed	book shelf	box	chair	counter	desk	door	dresser	garbage bin	lamp	monitor	night stand	pillow	sink	sofa	table	television	toilet
box	**14**	6	40	4	1	6	1	3	15	27	33	1	40	11	**6**	9	14	3	35	12
region	**28**	**32**	55	9	1	27	21	9	20	29	37	26	48	39	**33**	31	31	10	34	40
fg mask	**28**	15	60	9	1	29	5	7	**23**	33	**38**	31	**55**	39	32	32	36	11	37	38
Our	**32**	19	**66**	**10**	**2**	**36**	**33**	**10**	**23**	**34**	**38**	**36**	53	**43**	32	**34**	**41**	**14**	**37**	**51**

Results