the threshold on Pg used to stop merging is trained. This variant, too, performs less well than ISCRA.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	0.65	0.59^{23}	1.54	1.66^{10}	0.85	$\mathbf{0.81}^{50}$
Hoiem[11]	0.59	0.55^{17}	1.70	1.83^{10}	0.82	0.79^{17}
ISCRA	**0.66**	$\mathbf{0.60}^{12}$	**1.40**	$\mathbf{1.61}^{8}$	**0.86**	$\mathbf{0.81}^{16}$

Table 1. Results for large segment regime, **BSDS300**. OIS: optimal scale per (test) image, ODS: optimal scale for entire test data set. Superscripts: the average number of segments at the optimal scale for each method/measure. Best results are shown in bold.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	0.64	$\mathbf{0.59}^{18}$	1.49	1.69^{13}	**0.85**	$\mathbf{0.83}^{59}$
Hoiem[11]	0.60	0.56^{18}	1.66	1.78^{11}	0.84	0.81^{18}
ISCRA	**0.66**	$\mathbf{0.59}^{24}$	**1.42**	$\mathbf{1.60}^{10}$	**0.85**	0.82^{24}

Table 2. Large segment regime, **BSDS500**. See caption of Table 1.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	0.74	0.64^{6}	1.05	1.30^{3}	**0.85**	0.78^{12}
Hoiem[11]	0.67	0.65^{7}	1.34	1.37^{7}	0.80	0.77^{8}
ISCRA	**0.75**	$\mathbf{0.67}^{4}$	**1.02**	$\mathbf{1.18}^{3}$	**0.85**	0.77^{14}

Table 3. Large segment regime, **MSRC**. See caption of Table 1.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	**0.58**	0.55^{7}	1.70	1.75^{7}	**0.63**	$\mathbf{0.60}^{7}$
Hoiem[11]	0.56	0.54^{7}	1.72	1.73^{7}	0.61	0.59^{7}
ISCRA	0.57	$\mathbf{0.56}^{5}$	**1.64**	$\mathbf{1.65}^{5}$	0.62	$\mathbf{0.60}^{5}$

Table 4. Large segment regime, **VOC2012**. See caption of Table 1.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	0.64	0.58^{24}	1.63	1.88^{11}	0.89	0.86^{41}
[11]	0.67	$\mathbf{0.62}^{20}$	1.52	1.64^{12}	**0.90**	$\mathbf{0.87}^{20}$
ISCRA	**0.68**	$\mathbf{0.62}^{21}$	**1.50**	$\mathbf{1.73}^{15}$	**0.90**	$\mathbf{0.87}^{36}$

Table 5. Large segment regime, **SBD**. See caption of Table 1.

Method	Covering		VOI		PRI	
	OIS	ODS	OIS	ODS	OIS	ODS
UCM[2]	0.51	0.46^{141}	2.33	2.52^{34}	**0.90**	0.88^{330}
Hoiem[11]	0.51	0.48^{87}	2.36	2.50^{28}	0.89	0.88^{194}
ISCRA	**0.54**	$\mathbf{0.50}^{62}$	**2.21**	$\mathbf{2.34}^{20}$	**0.90**	$\mathbf{0.89}^{97}$

Table 6. Large segment regime, **NYU**. See caption of Table 1.

6. Conclusions

We present ISCRA: Image Segmentation by Cascaded Region Agglomeration. ISCRA consists of a cascade of probabilistic models that predict the probability of grouping neighboring regions. It is trained in sequence, allowing adaptation of feature weights to increasing segmentation scale; when applied on an image it produces a hierarchical segmentation, allowing the user to directly control the scale and the number of resulting regions. In experimental comparison on six data sets, ISCRA is a clear winner in region-based measures. It also is competitive in boundary-based measures in the superpixel regime, obtaining best results for part of the range for some data sets. ISCRA tends to achieve these results with fewer segments per image than other methods, making it potentially appealing for use as preprocessing step for semantic segmentation and other high-level perception tasks.

As the experiments show, a major limitation on ISCRA is its dependence on the initial segmentation. For instance, in our experiments here, it can not obtain ASA or boundary recall values above those of the finest scale of OWT-UCM. We plan to investigate initialization methods that combine fine-scale segmentations from multiple algorithms. We also are working on designing additional features that will improve accuracy of Pg estimated in later stages of ISCRA.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 2012.

[2] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE TPAMI*, 33(5), 2011.

[3] J. Carreira, R. Caseiro, J. Batista, and C. Sminchisescu. Semantic segmentation with second-order pooling. In *ECCV*, 2012.

[4] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE TPAMI*, 24(5), 2002.

[5] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2012 (VOC2012). http://www.pascal-network.org/challenges/VOC/voc2012.

[6] P. Felzenszwalb and D. Huttenlocher. Efficient graph-based image segmentation. *IJCV*, 59(2), 2004.

[7] B. Fulkerson, A. Vedaldi, and S. Soatto. Class segmentation and object localization with superpixel neighborhoods. In *ICCV*, 2009.

[8] M. Galun, E. Sharon, R. Basri, and A. Brandt. Texture segmentation by multiscale aggregation of filter responses and shape elements. In *ICCV*, 2003.

[9] S. Gould, R. Fulton, and D. Koller. Decomposing a scene into geometric and semantically consistent regions. In *ICCV*, 2009.

[10] D. Hoiem, A. Efros, and M. Hebert. Geometric context from a single image. In *ICCV*. IEEE, 2005.

[11] D. Hoiem, A. Efros, and M. Hebert. Recovering occlusion boundaries from an image. *IJCV*, 91(3), 2011.

[12] S. Kim, S. Nowozin, P. Kohli, and C. D. Yoo. Higher-order correlation clustering for image segmentation. In *NIPS*, 2011.

[13] A. Levinshtein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson, and K. Siddiqi. Turbopixels: Fast superpixels using geometric flows. *IEEE TPAMI*, 31(12), 2009.