Image Segmentation using Contour, Surface, and Depth Cues

Xiang Fu, Chen Chen, Jian Li, C.-C. Jay Kuo
University of Southern California

Changhu Wang
Toutiao AI Lab

Automatic Image Segmentation

- Challenges:
  - Leakage problem: separate object and environment that are too involved to find the boundaries for two adjacent regions using surface properties (e.g. in the shadow)
  - Over-segmentation problem: combine textured regions, which have high contrast inside and no clear contour outside the region
  - Diverse ground truths: different semantic understandings on image segmentation

Automatic Image Segmentation

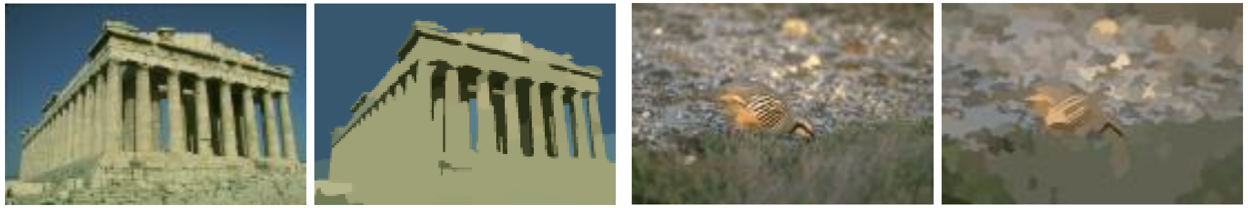
- Contributions:
  - Fully utilize 1D contour, 2D surface, and 3D depth cues for image segmentation
  - Proposed a 3D depth cue to segment different textured regions with similar color, and merge similar textured areas
  - Proposed a content-dependent spectral (CDS) graph for layered affinity models

Three Elementary Cues

- 1D Contour: discontinuities
- 2D Surface: similarities
- 3D Depth: layout

- Each cue might be sufficient to segment some images, but fails for others

(a) Contour Cue



(b) Surface Cue

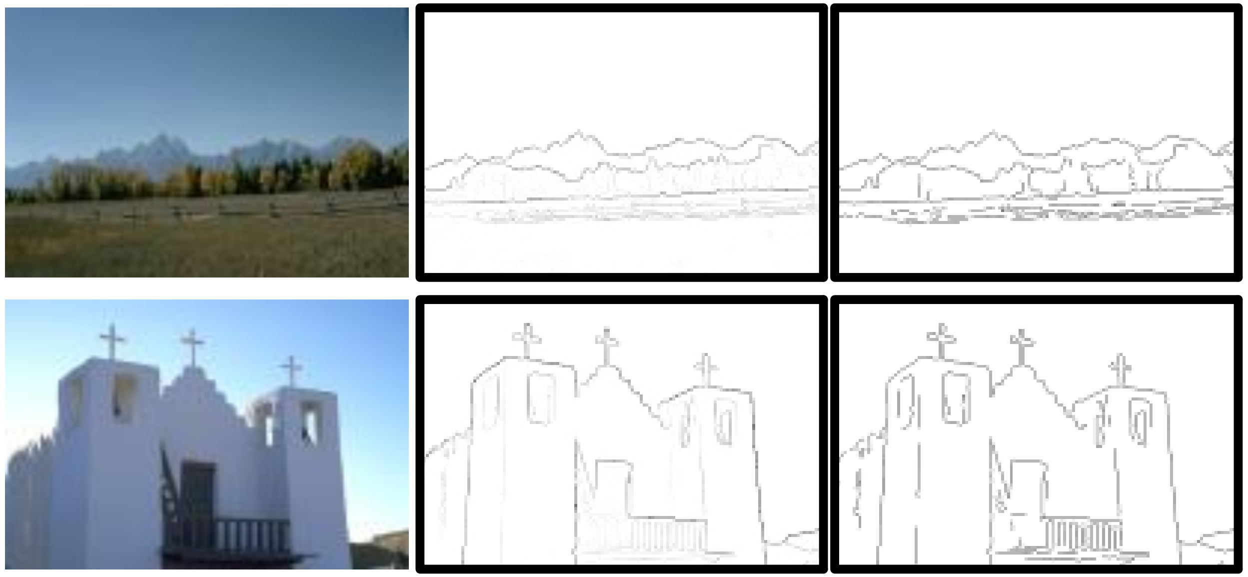
(c) Depth Cue

Examples for positive impact of cue

Examples for negative impact of cue

1D Contour Cue

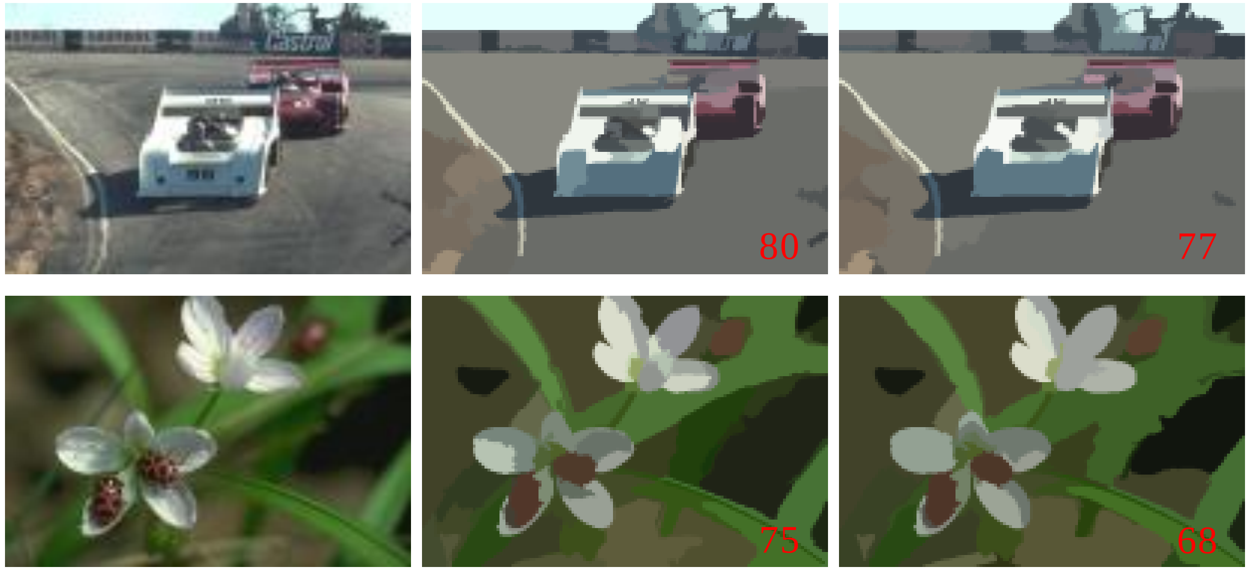
- More reliable if the contour is longer and more closed
- Fails if the boundary is blurred, in low contrast, or in smooth transition (leakage)



(a) Original (b) Structured Edge Detection [12] (c) 1D Contour Cue

2D Surface Cue

- Proved to be successful for region-based segmentation
- Unable to simplify textured regions with high variance (over-segmentation)



(a) Original

(b) CCP with h_r=5

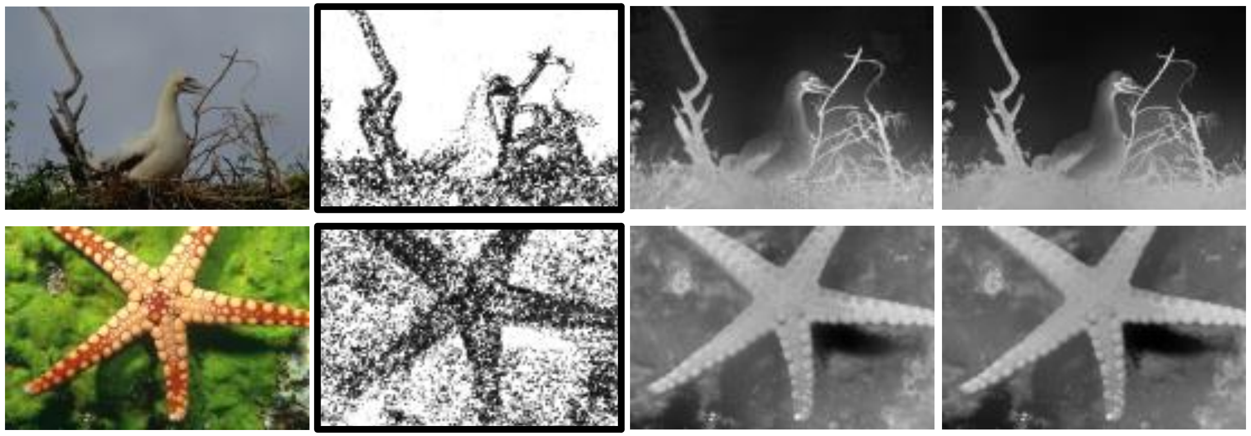
(c) CCP with h_r=6

2D Surface Cues

https://github.com/fuxiang87/MCL_CCP.

3D Depth Cue

- ➕ Helpful to clean the regions, especially for the textures
- ➕ Alleviate the limitations of contour and surface cues
- ➖ Unreliable if there is no edge details within these regions



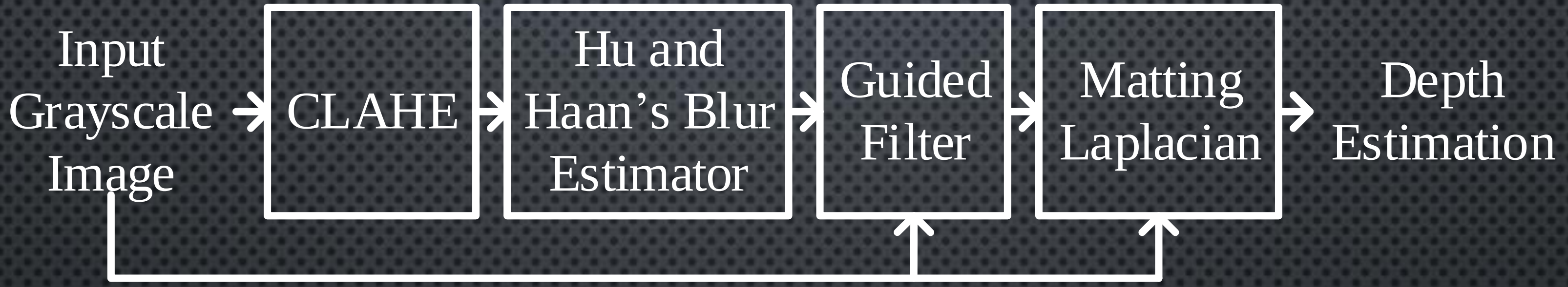
(a) Original

(b) Sparse defocus map by Hu and Haan [13]

(c) After Guided Filter [18]

(d) 3D Depth Cue

Proposed Depth Estimation



- Assumption: when taking a photo, the objects close to the focal plane are in focus, while the objects far from the focal plane are out of focus

- CLAHE [17] (contrast limited adaptive histogram equalization): preprocessing to improve the local contrast, and bring out more details for dark areas
- Hu and Haan’s Blur Estimator [13]: generate a sparse defocus map around the object boundaries
- Guided Filter [18]: attenuate the depth variance inside the regions to group the object regions in the same depth layer
- Matting Laplacian [19]: propagate the sparse defocus map to the entire image

Depth Estimation via CLAHE

- CLAHE [17]
(contrast limited adaptive histogram equalization): preprocessing to improve the local contrast, and bring out more details for dark areas

(a)	(b)	(c)	(d)
Original	Depth Cue w/o CLAHE	CLAHE	Depth Cue w/ CLAHE

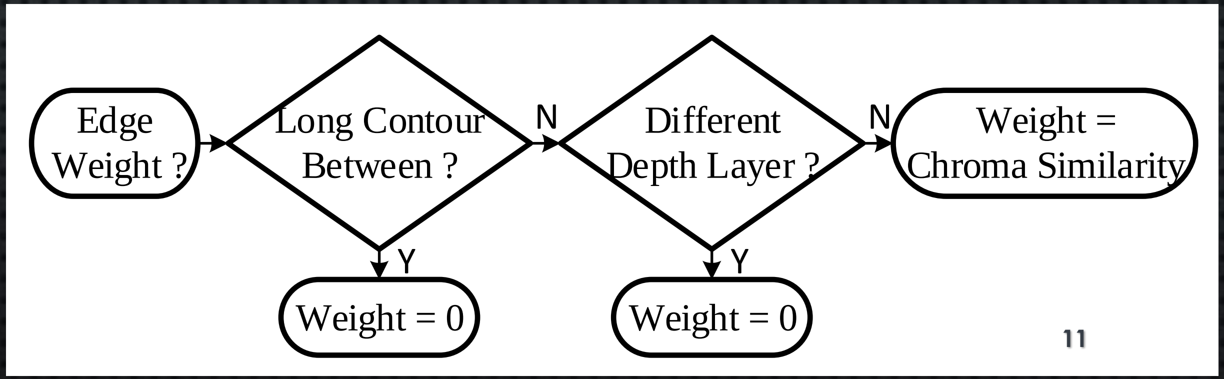
Depth Estimation via Guided Filter

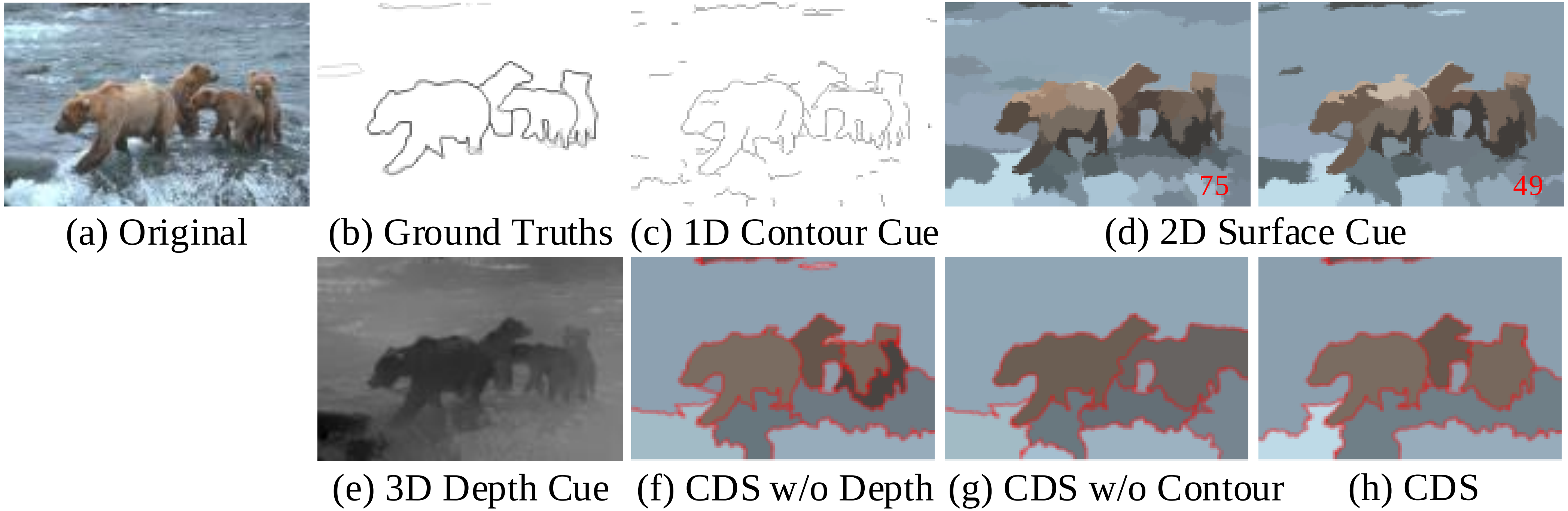
- Guided Filter [18]: attenuate the depth variance inside the regions to group the object regions in the same depth layer

(a) Original (b) Depth Cue w/o Guided Filter (c) Depth Cue w/ Guided Filter

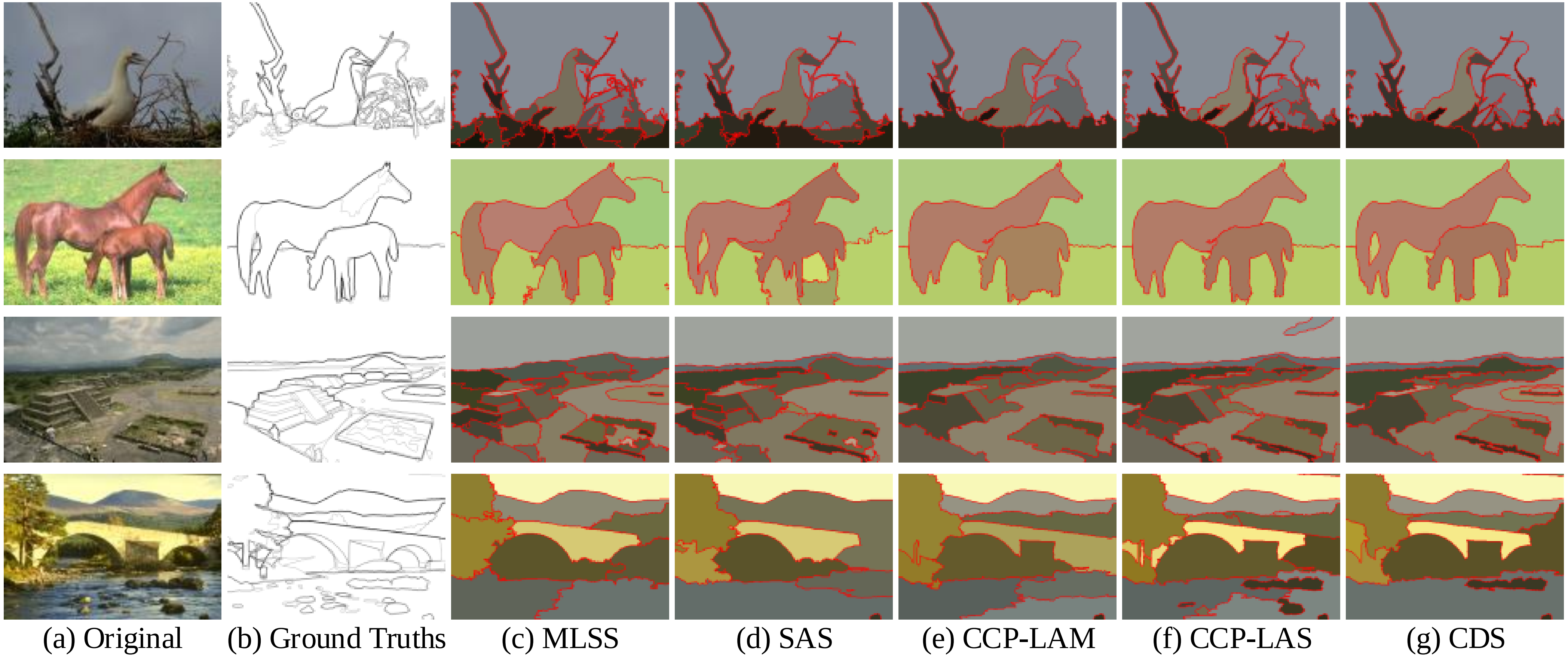
Content-Dependent Spectral Graph

- Layered affinity models use superpixel layers to connect pixels far from each other
  - Not performs well via old description of affinities between adjacent surface nodes
- CDS provides one solid solution
  - Reliable: long contours, large depth distance
  - Others: chroma similarity

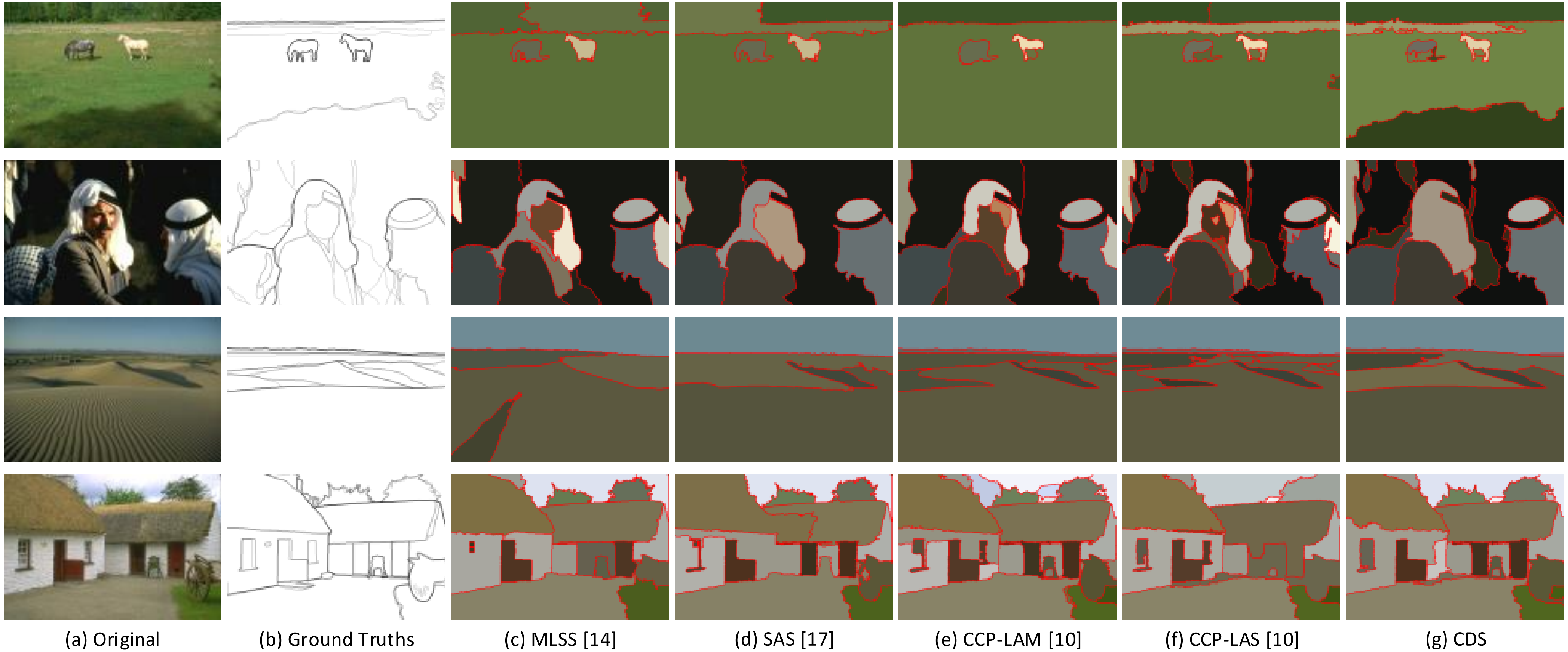




Three elementary cues and visual comparisons of segmentation results of CDS against CDS w/o depth and CDS w/o contour



Visual comparisons of segmentation results of CDS against four state-of-the-art methods: MLSS[5], SAS[6], CCP-LAM[9] and CCP-LAS[9].



Visual comparisons of segmentation results of CDS against four state-of-the-art methods: MLSS[5], SAS[6], CCP-LAM[9] and CCP-LAS[9].

Table 1. Performance comparison on the BSDS300 Dataset. The best result are highlighted in bold.

Algorithm	Cov↑	PRI↑	VoI↓	GCE↓	BDE↓
NCut [2]	0.44	0.7242	2.9061	0.2232	17.15
JSEG [24]	N/A	0.7756	2.3217	0.1989	14.40
MeanShift [3]	0.54	0.7958	1.9725	0.1888	14.41
FH [4]	0.51	0.7139	3.3949	0.1746	16.67
MNCut [25]	0.44	0.7559	2.4701	0.1925	15.10
NTP [26]	N/A	0.7521	2.4954	0.2373	16.30
MLSS [5]	0.53	0.8146	1.8545	0.1809	12.21
SAS [6]	0.62	0.8319	1.6849	0.1779	11.29
CCP-LAM [9]	**0.68**	0.8404	1.5715	0.1635	10.20
CCP-LAS [9]	**0.68**	0.8442	1.5871	0.1582	10.46
CDS	**0.68**	**0.8539**	**1.5712**	**0.1572**	**10.18**
CDS (w/o depth)	0.65	0.8449	1.6293	0.1580	10.48
CDS (w/o contour)	0.64	0.8426	1.6185	0.1597	10.51

***Note that we do not have the algorithm CDS w/o surface cue, for the whole CDS algorithm is designed based on the 2D surface cue.**

References

- [5] Tae Hoon Kim, Kyoung Mu Lee, and Sang Uk Lee, “Learning full pairwise affinities for spectral segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 7, pp. 1690–1703, July 2013.
- [6] ZhenguoLi,Xiao-MingWu,andShih-FuChang, “Segmentation Using Superpixels: A bipartite graph partitioning approach,” in CVPR’12, 2012, pp. 789–796.
- [9] XiangFu, Chien-YiWang, ChenChen, Changhu Wang, and C.-C. Jay Kuo, “Robust image segmentation using contour-guided color palettes,” in ICCV’15, 2015, pp. 1618 – 1625.
- [12] PiotrDollrandC.LawrenceZitnick, “Structuredforests for fast edge detection,” in ICCV’13, 2013, pp. 1841– 1848.
- [13] Hao Hu and Gerard de Haan, “Low cost robust blur estimator,” in ICIP’06, 2006, pp. 617–620.
- [17] Karel Zuiderveld, Graphics Gems IV, Academic Press Professional, Inc., San Diego, CA, USA, 1994.
- [18] Kaiming He, Jian Sun, and Xiaoou Tang, “Guided image filtering,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 6, pp. 1397–1409, June 2013.
- [19] AnatLevin,DaniLischinski,andYairWeiss, “Aclosedform solution to natural image matting,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 2, pp. 228–242, Feb. 2008.