



Fast pseudo-semantic segmentation for joint region-based hierarchical and multiresolution representation

Rafik Sekkal, Clément Strauss, François Pasteau, Marie Babel, Olivier Déforges

► To cite this version:

Rafik Sekkal, Clément Strauss, François Pasteau, Marie Babel, Olivier Déforges. Fast pseudo-semantic segmentation for joint region-based hierarchical and multiresolution representation. SPIE Electronic Imaging - Visual Communications and Image Processing, Jan 2012, San Francisco, United States. pp.83050J, 10.1117/12.908829 . hal-00657805

HAL Id: hal-00657805

<https://hal.science/hal-00657805>

Submitted on 9 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fast pseudo-semantic segmentation for joint region-based hierarchical and multiresolution representation

Rafiq Sekkal¹, Clement Strauss¹, François Pasteau¹, Marie Babel², Olivier Deforges¹

¹ IETR - Image group Lab
CNRS UMR 6164
INSA de Rennes
20, avenue des Buttes de Coesmes
CS 70839
F 35708 Rennes Cedex 7

² Lagadic Team
Université Européenne de Bretagne
INSA de Rennes
IRISA, INRIA Rennes
Bretagne Atlantique
Rennes 35042, France

ABSTRACT

In this paper, we present a new scalable segmentation algorithm called JHMS (Joint Hierarchical and Multiresolution Segmentation) that is characterized by region-based hierarchy and resolution scalability. Most of the proposed algorithms either apply a multiresolution segmentation or a hierarchical segmentation. The proposed approach combines both multiresolution and hierarchical segmentation processes. Indeed, the image is considered as a set of images at different levels of resolution, where at each level a hierarchical segmentation is performed. Multiresolution implies that a segmentation of a given level is reused in further segmentation processes operated at next levels so that to insure contour consistency between different resolutions. Each level of resolution provides a Region Adjacency Graph (RAG) that describes the neighborhood relationships between regions within a given level of the multiresolution representation. Region label consistency is preserved thanks to a dedicated projection algorithm based on inter-level relationships. Moreover, a preprocess based on a quadtree partitioning reduces the amount of input data thus leading to a lower overall complexity of the segmentation framework. Experiments show that we obtain effective results when compared to the state of the art together with a lower complexity.

Keywords: Image segmentation, spatial scalability, hierarchical segmentation, RAG

1. INTRODUCTION

Image segmentation is a fundamental process in many image, video and computer vision applications. Image segmentation is used to partition an image into separate regions, according to a dedicated homogeneous criterion (texture, colors, gradient...) so that to reach a given level of granularity (regions, objects).¹

Segmentation techniques generally suffer from high computational complexity as they typically process full resolution images. In this article, we propose to use a simplified version of the image so that to achieve faster segmentation solution. To this aim, a quadtree-based partitioning process is applied, leading to a non uniform subsampled image representation.

Moreover, our segmentation performs both multiresolution and hierarchy scalability, as shown on figure 1. Multiresolution representation consists of representing a given image as a set of images at different levels of resolution. In this image representation, the segmentation process is typically based on *Top-down* approaches. In the literature, multiresolution segmentations rely on regular pyramids.²⁻⁴ The relationships and properties between pixels from different levels of representation are thus preserved. As for region-based hierarchical representations they rely on a bottom-up approach: seeds are first created then the number of regions is progressively reduced. Hierarchical segmentations are classically represented by irregular pyramids.^{5,6} Semantic or pseudo-semantic region extraction takes advantage of multiresolution representation in terms of region definition at the expense of the computational complexity. Then an effective tradeoff should be found so that to obtain both content compliant representation and a fast segmentation solution.

This paper is organized as follow: section 2 describes quadtree partitioning which is the key data structure and the regular image pyramid construction driven by the quadtree. Section 3 presents the global JHMS framework including the description of the multiresolution RAG on which inter and intra level region links rely. In section 4, experiments on Berkeley benchmark and obtained results are discussed.

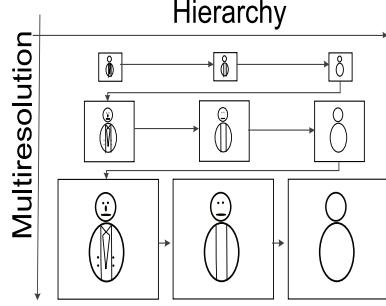


Figure 1. Multiresolution and hierarchical segmentation representations

2. RESOLUTION SCALABLE REPRESENTATION

In this section, we describe the preprocess steps involved in JHMS technique, *i.e.* the structure of the quadtree and the multiresolution pyramid.

The quadtree represents the key acceleration mechanism of our algorithm. The goal of the quadtree is to partition the image in order to reduce the processing space while preserving the semantic content of the image.

The quadtree corresponds to a block-based partition that depends on local cues of the image. With a top-down approach, each node that does not respect a given homogeneity criterion is decomposed into four sub-blocks. Figure 2 shows different decomposition steps of the quadtree. A block based image is then built following the quadtree partitioning where each block of the image corresponds to the local block mean value.

The resolution scalable representation is obtained by building a multiresolution pyramid that uses the block based image at different levels of resolution. At each level, the image is then composed of blocks of different sizes according to the quadtree partition. Let Res^l be the block image at the l^{th} resolution level of the pyramid.

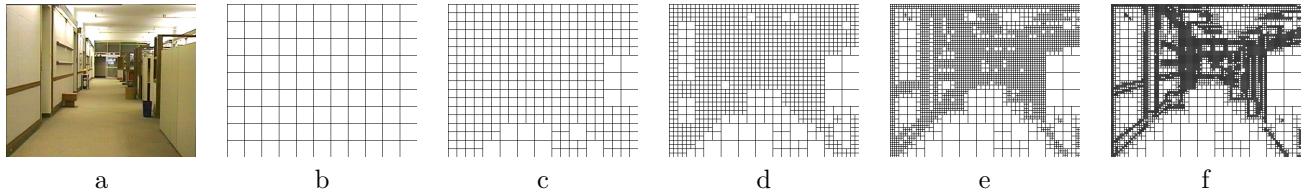


Figure 2. Quadtree construction on the first frame of *hall* in cif (352x288) resolution using 5 levels of multiresolution. a) The original image, b) c) d) e) f) quadtree representations

3. JHMS FRAMEWORK

In this section, we introduce the general JHMS framework. First we present the general principles of the segmentation technique. Secondly, we focus on the multiresolution Region Adjacency Graph (RAG) process. Finally hierarchy is considered.

3.1 General principles

The JHMS technique relies on an iterative process of the pyramid until reaching the desired level of resolution l_{min} . As described in figure 3, an initialization process from the low resolution image is first required.

Let RAG^l be the Region Adjacency Graph at resolution level l , with $l \in \{l_{max}, \dots, l_{min}\}$. Regions of the initial $RAG^{l_{max}}$ correspond to the blocks of the lowest resolution image. Then, a hierarchical segmentation modifies the $RAG^{l_{max}}$ by including hierarchical description.

Iteratively, for each level l of the pyramid, the current RAG^l is obtained by projecting the RAG^{l+1} onto the RAG^l resolution. RAG^l is then updated by taking into account both the projected RAG^{l+1} and changes within neighborhood relationships. Finally, a hierarchical segmentation is performed onto the current RAG^l .

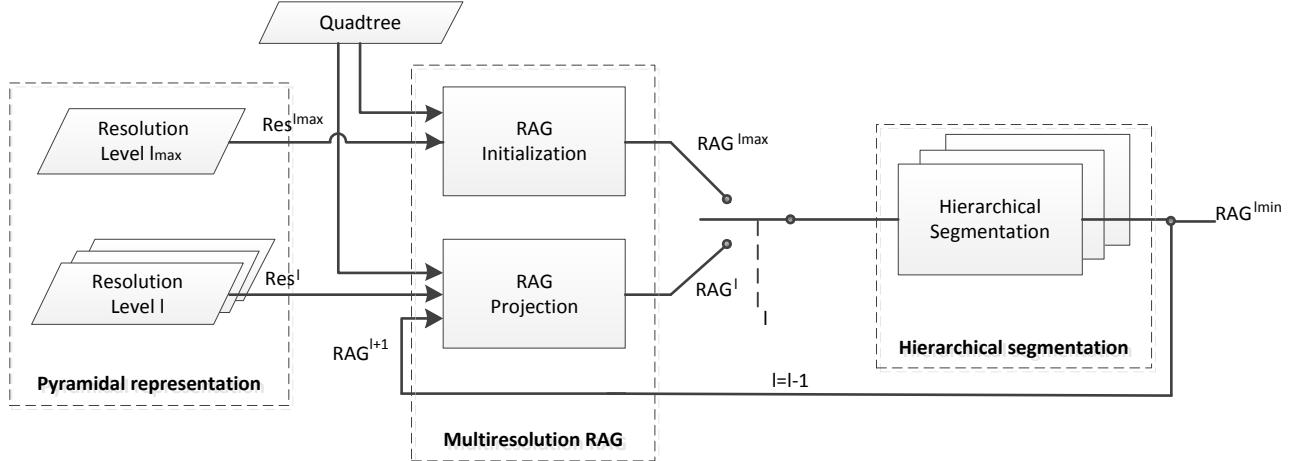


Figure 3. JHMS general scheme

3.2 Multiresolution RAG

Figure 4 depicts the different steps of the multiresolution RAG algorithm. To project RAG^{l+1} onto RAG^l across resolution, the algorithm projects the regions labels by using the quadtree partition. Indeed it is used as a reference to decide which blocks are kept unchanged and which ones are split. Thus regions composed of at least one unchanged block are considered as fixed regions, and are directly projected in the current level. Labels of fixed regions are maintained for the next segmentation step. Inheriting labels ensures the label consistency of the segmentation. Unchanged blocks correspond to leaves of the quadtree. Thus if an unchanged block is detected at l^{th} level, it means that 2^{l+1} pixels in the full resolution will be discarded from the computation. Consequently, the computation is reduced thanks to quadtree partitioning.

Region relationships between two successive levels are shown in figure 4. Two kinds of region relationships are obtained: regions created in RAG^l with blocks from the same region in the RAG^{l+1} (Figure 4.a) or regions with blocks that belong to different regions in RAG^{l+1} (Figure 4.b).

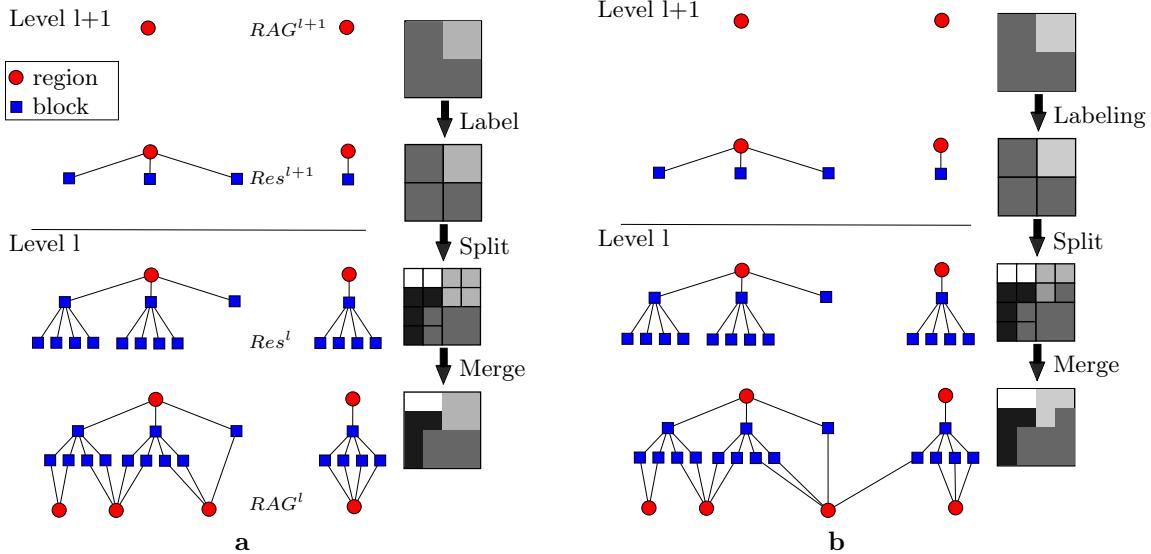


Figure 4. Inter-level regions relationships. Four regions in Res^l , in a) regions are composed with blocks of the same region parent. However in b) there is one region composed from blocks belonging to the two regions parents

3.3 Hierarchical segmentation

An Extension to hierarchical representation at each level of multiresolution can be designed to overcome the natural resulting over-segmentation by selecting the segmentation granularity. This global solution called Joint Hierarchical and Multiresolution Segmentation (JHMS) provides a highly scalable region representation.

4. EXPERIMENTS AND RESULTS

4.1 Region merging criteria

The proposed algorithm combines two criteria. The first criterion is the difference between the mean values of two adjacent regions. The second criterion computes the gradient between blocks along the shared contour between two regions. For our experiments, we used the Locally Adaptive Resolution structure to build the quadtree and the hierarchical segmentation function.⁷ Although a learning step is recommended to find the optimal set of parameters, they have been empirically tuned relying on experiments when trying to get effective results.

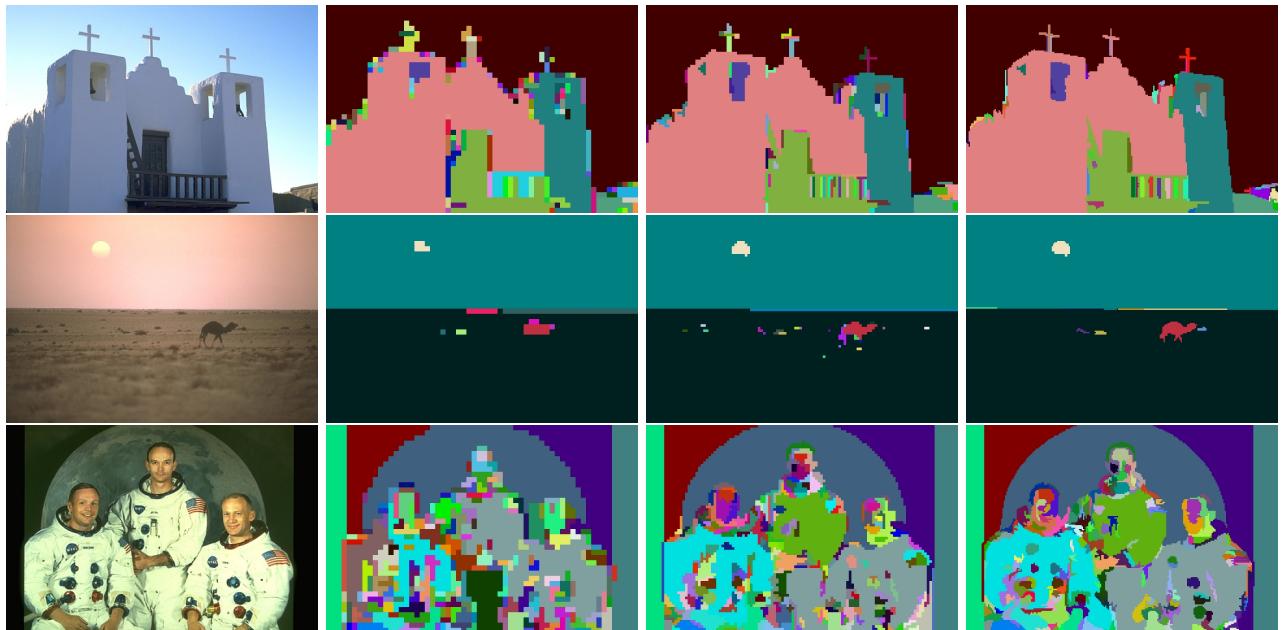


Figure 5. Scalable segmentation results

4.2 Visual results

In Figure 5, resulting segmented images are shown. Regions are here presented in false color. From the left to the right, original image is first presented, then the corresponding oversampled image of labels at different resolution levels (*i.e.* 3, 2, 1). As can be observed, results at this resolution are satisfactory in terms of representation. In addition, the consistency of region labels is well preserved across resolutions, so that object tracking can be envisaged throughout the multiresolution. Furthermore, color gradated regions such as the sky of the first image are well detected as a single region thanks to the local gradient feature.

4.3 Objective quality of segmentation

In order to compare the proposed algorithm with the literature, obtained segmentation maps are tested against the Berkeley benchmark⁸. This benchmark is usually used for comparing contour detector algorithms. Contour maps are thus compared with contour maps designed by human beings which are then considered as ground truth. In our experiment the BSDS300 dataset is used.

The proposed algorithm is strictly based on local gradient and mean merging criteria. To provide a fair comparison, only color-based contour detectors, from the Berkeley benchmark, that share the same features have been compared. The algorithm results have been then compared with ground truth images. As shown in table 1, JHMS obtained $F=0.60$ score. As for it, GPB⁹ (*Global Probability of Boundary*) combines the use of local information derived from brightness, color, and texture signals to produce a contour detector. It provides the best performance on the benchmark with $F=0.70$. CG¹⁰ (*Color Gradient*), another algorithm based on same features as the JHMS provides a score of $F=0.57$.

JHMS provides a pseudo semantic segmentation and is not able to reach object level granularity by itself. Typically, the difference of score with GPB is due to the fact that additional texture information are used in the GPB segmentation. Texture improves the segmentation results, however, extracting texture features makes the algorithm more computationally complex.

Algorithms	Ground Truth	GPB	CG	JHMS
Scores Average	0.79	0.70	0.57	0.59

Table 1. Quantitative scores on Berkeley database BSDS30

In figure 6, different contour maps are presented from GPB, CG and JHMS techniques. Most of contours in JHMS are similar to those found in the ground truth images. Neither global information nor texture features are used in JHMS. In consequence, any strong brightness or color variations within a single object lead to over-segmentation, thus penalizing the score of our algorithm. For example, to segment the tiger in the last row of figure 6, hand segmented images consider the whole tiger as a single region and one consistent contour. With the proposed algorithm, the tiger is detected as multiple adjacent regions corresponding to the stripes of the skin.

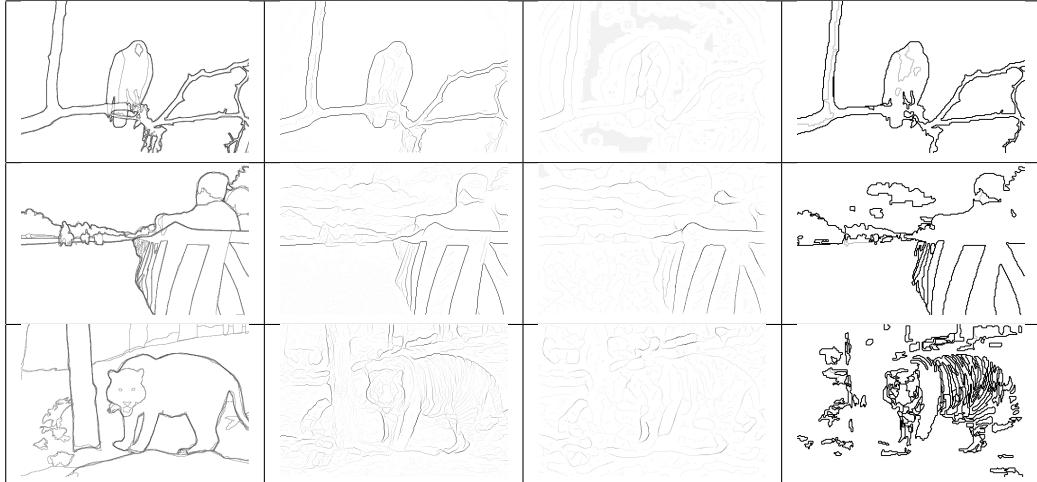


Figure 6. Image Boundaries, from the left: ground truth, Global Probability of Boundary GPB, Color Gradient (CG) and our algorithm JHMS

4.4 Multiresolution and quadtree partitioning influence on complexity and objective scores

In order to exhibit the influence of the quadtree partitioning and the influence of the multiresolution scalability on the computational complexity and objective scores, segmentations in different configurations have been performed on the BSDS300 dataset images. Results are here obtained using an Intel Core i7 @2.67GHz with a single thread. The average time of segmenting the 100 images of the dataset and mean scores of Berkeley benchmark have been measured.

First the influence of the quadtree partitioning is evaluated without multiresolution. Segmentations are directly performed on the full resolution image, either on pixel based images or on block based images following the quadtree partitioning. Table 2 shows that the quadtree partitioning enables a 5.5 times speed up compared

to the pixel based images. In addition the objective score with quadtree partitioning is better. This can be explained by the fact that the quadtree helps by guiding the segmentation algorithm in the image by providing a first pseudo semantic description.

Secondly the multiresolution segmentations are compared with an increasing number of resolution levels. Table 2 depicts the results of our experiments with segmentations performed without multiresolution, and with up to 4 levels of embedded resolution levels. Objective scores remain almost identical when increasing the number of resolution levels and remain close to the performances of the segmentations without multiresolution. However, using multiresolution, scalability strongly impacts the complexity. Complexity comes from both the RAG projection from one resolution level to the next and the segmentation itself at one level of resolution. With more levels of resolution the segmentation at each resolution level is simplified, therefore, more levels tend to reduce the overall complexity. As for the RAG projection mechanism, its complexity is proportional to the number of blocks. From the half resolution up to the full resolution, the RAG projection handles much more small blocks than during other previous projection steps. Therefore, the last RAG projection onto the full resolution level explains for the most part the complexity of the method. Future plans are to optimize the RAG projection and reduce the overall complexity.

Multiresolution	none	none	2 levels	3 levels	4 levels
Quadtree partitioning	no	yes	yes	yes	yes
BSDS300 scores	0.58	0.60	0.575	0.577	0.578
Execution time (s)	1.100	0.200	1.429	1.355	1.300

Table 2. Multiresolution and quadtree partitioning influence on complexity and objective scores

5. CONCLUSION

We present in this paper a fast segmentation method characterized by both multiresolution and hierarchical scalability. At each level, we apply a region merging approach. In the meantime, we construct a hierarchical RAG. Through the successive levels of resolution, the label consistency is insured by the dedicated RAG projection mechanism. As our approach relies on a quadtree structure, thus simplifying the image, our approach exhibits a low computational complexity. Results obtained by the proposed algorithm are effective when considering the limited amount of information used as input. Regions remain coherent with the original image content. Contours are well detected, leading to an efficient pseudo-semantic representation.

Future works will be performed to improve the objective results, in order to find the optimal set of parameters. In particular, a statistical study has to be performed on the learning set of the Berkeley dataset.

REFERENCES

- [1] Lucchese, Mitra, S. K., and Barbara, S., “Color image segmentation: A State-of-the-Art survey,” *Proceedings of the Indian National Science Academy* **67**(2), 207–221 (2001).
- [2] Kropatsch, W., Haxhimusa, Y., and Ion, A., “Multiresolution image segmentations in graph pyramids,” in [*Applied Graph Theory in Computer Vision and Pattern Recognition*], Kandel, A., Bunke, H., and Last, M., eds., *Studies in Computational Intelligence* **52**, 3–41, Springer Berlin / Heidelberg (2007).
- [3] Vantaram, S. R., Saber, E., Dianat, S., Shaw, M., and Bhaskar, R., “An adaptive and progressive approach for efficient gradient-based multiresolution color image segmentation,” in [*Proceedings of the 16th IEEE international conference on Image processing*], *ICIP’09*, 2345–2348, IEEE Press, Piscataway, NJ, USA (2009).
- [4] Stojmenovic, M., Solis-Montero, A., and Nayak, A., “Colour and texture based pyramidal image segmentation,” in [*2010 International Conference on Audio Language and Image Processing (ICALIP)*], 778–786, IEEE (Nov. 2010).
- [5] Kropatsch, W. G., Haxhimusa, Y., Pizlo, Z., and Langs, G., “Vision pyramids that do not grow too high,” *Pattern Recogn. Lett.* **26**, 319–337 (February 2005).
- [6] Mignotte, M., “Mds-based multiresolution nonlinear dimensionality reduction model for color image segmentation,” *Neural Networks, IEEE Transactions on* **22**, 447 –460 (march 2011).

- [7] Déforges, O., Babel, M., Bédat, L., and Ronsin, J., “Color LAR codec: a color image representation and compression scheme based on local resolution adjustment and self-extracting region representation,” *IEEE Transactions on Circuits and Systems for Video Technology* **17**, 974–987 (Aug. 2007).
- [8] Martin, D., Fowlkes, C., Tal, D., and Malik, J., “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in [*Proc. 8th Int'l Conf. Computer Vision*], **2**, 416–423 (July 2001).
- [9] Maire, M., Arbelaez, P., Fowlkes, C. C., and Malik, J., “Using contours to detect and localize junctions in natural images,” in [*CVPR*], (2008).
- [10] Martin, D. R., Fowlkes, C. C., and Malik, J., “Learning to detect natural image boundaries using local brightness, color, and texture cues,” *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 530–549 (May 2004).