Oversegmentation Methods: A New Evaluation

Bérengère Mathieu[1(✉)], Alain Crouzil[1], and Jean Baptiste Puel[2]

[1] UPS, IRIT, Université de Toulouse, Toulouse, France
{berengere.mathieu,alain.crouzil}@irit.fr
[2] ENFA, IRIT, Université de Toulouse, Toulouse, France
jean-baptiste.puel@irit.fr

Abstract. Using superpixels instead of pixels has become a popular pre-processing step in computer vision. Currently, about fifteen oversegmentation methods have been proposed. The last evaluation, realized by Stutz *et al.* in 2015, concludes that the five more competitive algorithms achieve similar results. By introducing HSID, a new dataset, we point out unexpected difficulties encountered by state-of-the-art oversegmentation methods.

Keywords: Superpixel · Segmentation · Comparative study

1 Introduction

The idea and definition of superpixel is given for the first time by Ren *et al.*, in [12]. The two authors describe a segmentation method including an *oversegmentation* pre-processing grouping pixels into small homogeneous and regular regions called *superpixels*. By using them instead of pixels, they significantly reduce the complexity of their algorithm. Subsequently, superpixels have been integrated with success to a lot of methods, [4,20]. Currently, oversegmentation is an active research field, with steady publication of new methods [1,2,8].

Previous Work. Made in 2012 by Achanta *et al.*, the first review of oversegmentation methods [1] compare six algorithms: Normalized Cut (NC) [12], Felzenszwalb algorithm (FZ) [3], Quick Shift (QS) [17], TurboPixels (TP) [6], Veksler method (VK) [18] and the Simple Linear Iterative Clustering method (SLIC) [1]. Tests were carried out using 100 images of Berkeley Segmentation Dataset [9] (BSD[1]). For each method Achanta *et al.* [1] analyze the complexity, the execution time and the quality of the produced oversegmentation results. They use two metrics: the undersegmentation error rate (UE) and the boundary recall measure (BR). Both of them compare the oversegmentation result S to a ground-truth G. The UE score takes into account for each object G_i in G the set of superpixels required to cover it and counts the number of pixels

[1] https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/.

leaking: $UE(S, G) = \frac{1}{N} \sum_{G_i \in G} \sum_{S_j \cap G_i \neq \emptyset} min(|S_j \cap G_i|, |S_j - G_i|)$ where N is the pixel number and $|E|$ denotes the cardinality of the set E. The result is a rate between 0 and 1, 0 denoting an error-free oversegmentation result. The BR score checks if whether boundary pixels around objects in G match with boundary pixels in S. We indicate by B_G the set of boundary pixels in G and B_S the set of boundary pixels in S. If we assume that there is no doubt about the fact that a pixel belongs or not to a boundary, the quality of an oversegmentation can be evaluated by calculating the rate of boundary pixels in G corresponding to boundary pixels in S: $BR(S, G) = \frac{|B_S \cap B_G|}{|B_G|}$.

In fact, even for a human, it is sometimes difficult to know whether a pixel belongs or not to a boundary. Achanta *et al.* allow a distance τ_{br} of 2 pixels between pixels in B_G and in B_S. The BR score is in the range $[0, 1]$, 1 meaning that all boundaries in G match with boundaries in S. In 2015, Stutz mades a second evaluation [14]. Their first contribution is to include seven supplementary oversegmentation methods: Entropy Rate Superpixels (ERS) [7], Superpixels via Pseudo-Boolean Optimization (SPBO) [21], Contour Relaxed Superpixels (CRS) [2], Superpixels Extracted via Energy-Driven Sampling (SEEDS) [16], Topology Preserved Superpixels (TPS) [15], Depth-Adaptive Superpixels (DAS) [19] and Voxel Cloud Connectivity Segmentation (VCCS) [11]. They also use as additional dataset, 400 images of New York University [13], NYU[2]. As the dimensions of the two datasets photographs are not identical (481×321 pixels for BSD, 640×480 pixels for NYU), Stutz *et al.* modify the threshold τ_{br}, allowing matching between boundary pixels $0.0075 \times diag$ away, where $diag$ is the image diagonal length. In Achanta *et al.* review [1], FZ and SLIC methods achieve the best scores. Stutz *et al.* evaluation [14] corroborates this result and shows that QS, CRS and ERS algorithms achieve performances similar to those of FZ and SLIC. On the BSD images, best scores are achieved with oversegmentation containing approximately 1000 superpixels. For FZ, SLIC, QS, ERS and CRS methods, UE is lower than or equal to 0.04 and BR is greater than or equal to 0.99. On NYU dataset, with about 1500 superpixels, UE is lower than or equal to 0.09 and BR is greater than or equal to 0.99. For these two datasets, execution times are about one second on a computer with a 3.4 GHz Intel Core i7 processor and 16 Go RAM.

Contributions. In Stutz *et al.* [14], evaluation the BR results achieved by the five best oversegmentation methods are so close to the maximal score, that it seems difficult to suggest a new method allowing a significant improvement. However, one can ask whether the two used datasets are sufficient for an exhaustive evaluation. The BSD and NYU images cover a wide panel of situations, containing both outdoor (BSD) and indoor (NYU) images with weak local contrast, important noise and lighting problems. However, images of the two datasets have similar sizes (some thousand of pixels) and small size in comparison to images taken with common cameras. Hence, our first contribution is a Heterogeneous Size Image Dataset (HSID), mainly containing big size images

---

[2] http://cs.nyu.edu.

(millions of pixels). HSID allows to check that algorithms do not suffer from a bias related to image dimensions and to better evaluate their scaling up. As HSID contains images with heterogeneous sizes, we need to transform BR into a fuzzy boundary adherence measure, FBR, which is the second contribution of this article. Moreover, we show that UE is not suitable for HSID. The demonstration leading us to this conclusion allows to better understand the behavior of UE and should be valuable for other works. The major contribution of this paper is a careful analysis of the result of the five best oversegmentation methods (FZ, QS, SLIC, ER and CRS) and of a recently proposed algorithm, Waterpixel (WP) [8]. Unlike previous evaluations, we show that, applied on HSID, each method encounters specific difficulties. The remainder of the article is organized as follows: in Sect. 2 we indicate properties of a suitable oversegmentation algorith and we describe the algorithms that we will compare. In Sect. 3 we describe HSID and we discuss about UE and BR measures. In Sect. 4 we analyze the evaluation results. We conclude with a discussion about perspectives of this work.

2   State-of-the-Art Methods

The review of works using superpixels [4,5,10,20,22,23] shows that a good oversegmentation method must satisfy five properties: **validity** (an oversegmentation must be an image partition into connected components) **boundary adherence** (superpixels must not overlap different objects of the image) **conciseness** (an oversegmentation must give as few superpixels as possible) **simplicity** (the number of neighbors of each superpixel must be as small as possible, to avoid a complex adjacency graph) **efficiency**   (an oversegmentation algorithm must have an execution time as low as possible). Simplicity and efficiency properties ensure that the time spent to oversegment the image and the time necessary to take superpixel neighborhood into account, will not be longer than the time saved by the usage of superpixels instead of pixels. Because boundary adherence is much more difficult to satisfy with large superpixels, this property is generally in contradiction with conciseness. We call *adaptivity*, the ability of an algorithm find the best compromise between these two properties, by reducing the number of superpixels in wide homogeneous regions.

According to Stutz *et al.* review [14] five methods outperform other algorithms: FZ [3], QS [17], ERS [7], SLIC [1] and CRS [2]. The FZ [3] and ERS [7] algorithms use a graph-based representation of the image $G < V, E >$, with V the set of elements to be grouped (*i.e.* the pixels) and E, the set of edges linking pairs of neighboring elements. Each edge is weighted using a dissimilarity measure. FZ uses a predicate checking that the dissimilarity between elements along the boundary of two components is greater than the dissimilarity between neighboring elements within each of the two components, to produce a partition of G in K connected components corresponding to superpixels. ERS is a greedy algorithm selecting a subset $A \subset E$ and removing these edges. The result is a partition of G, which maximizes an entropy rate. The QS method [17] is a modification of the medoid-shift algorithm to efficiently find modes of a Parzen

density estimate P. Color and location of each pixel are used as feature vectors that are clustered by linking each vector to its nearest neighbor which increases P. The SLIC method [1] is an adaptation of the k-means algorithm. Starting from an image oversegmentation into a regular grid, the average color and location features of each superpixel are computed. Then, each pixel is re-assigned to the most similar superpixel and the average features of superpixels are recomputed. Finally a post-processing step reassigns disjoints pixel sets to nearby superpixels, to ensure an image partition into connected components. The CRS algorithm [2] finds a partition S into superpixels, which has a high likelihood of having generated the observed image. Starting from an initial segmentation into rectangular superpixels, CRS maximizes the probability function by reallocating some boundary pixels to another superpixel. In our evaluation, we add to these five state-of-the-art methods, WP, a watershed transformation based algorithm using a spatially regularized gradient, which has been recently suggested by Machairas *et al.* [8].

3   New Oversegmentation Evaluation Benchmark

We provide a new oversegmentation evaluation dataset, including 100 images from Wikimedia Common[3]. Photographs have been selected to cover a wide variety of difficulties, including blur, noise, shadow, weak contrast and objects with similar colors. For each image, a hand drawn ground truth is provided. First, objects to extract are identified. Then a segmentation is designed by locating and fitting each region corresponding to a same object with a same color. Finally, boundaries of regions are automatically extracted allowing to visually check and remove some mistakes in the previous step. The images, ground truth and a file giving image licenses are available online[4].

Need for a Cautious Use of Undersegmentation Error. In reviews [1,14] UE was one of the two measures used to check boundary adherence of superpixels. However, our investigation shows that this measure must be used with cautions, in particular for dataset like HSID containing images of highly varying complexities. Figure 1 shows an oversegmentation into regular squared superpixels of two kind of images: a portrait-like in Fig. 1a with a unique big object on foreground and a panoramic-like in Fig. 1b, with multiple small objects on foreground. The foreground areas in the two images are the equal and a visual analysis show that superpixels similarly fail to match object boundaries. However, the UE score equals to 0.09 in the first case (Fig. 1a) and to 0.23 in the second case (Fig. 1b). This difference is explained by the fact that in Fig. 1a, superpixels wholly included within the objects boundaries are more numerous than in Fig. 1b (105 against 68). These superpixels do not carry information about boundary adherence and yet are taken into account by UE score, decreasing it. Hence the average UE score for these two images is not relevant to measure

---

[3] https://commons.wikimedia.org/wiki/Accueil.
[4] http://image.ensfea.fr/hsid/.

(a) Portrait-like synthetic image.
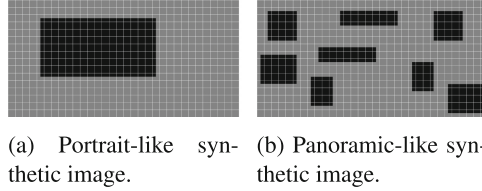
(b) Panoramic-like synthetic image.

Fig. 1. Background is in gray, foreground in black and superpixel boundaries in white.

boundary adherence. We encountered the same problem with HSID images and chose not to use UE.

Fuzzy Boundary Recall Measure. The metric BR measures the capacity for a method to give superpixels whose boundaries match with ground truth boundaries. To take into account the uncertainty about boundary location in hand-drawn ground truth, previous approaches use a static threshold. They accept matching between boundary pixels at a distance of $0.0075 \times diag$ pixels, where $diag$ is the image diagonal length. On big images of HSID (several millions of pixels) this distance (more than 30 pixels) is clearly too large. Rather than choosing another static threshold, we suggest to amend standard BR formula using fuzzy-set theory to introduce some tolerance error near the border pixels. Let G be a partition of the image into L connected components corresponding to objects $(G_1, \cdots, G_L)$ and S an oversegmentation into K superpixels $(S_1, \cdots, S_K)$, with $L << K$. A pixel p_i in S is on a boundary if $\exists p_j$ such as $p_j \in \text{nei}(p_x) \wedge (p_i \in R_n^S \Rightarrow p_j \notin R_n^S)$ where nei is a function giving for each pixel the set of its neighbors. Likewise, p_i is a boundary pixel in G, if $\exists p_j$ such as $p_j \in \text{nei}(p_x) \wedge (p_i \in R_n^G \Rightarrow p_j \notin R_n^G)$. Let B_G be the set of boundary pixels in G and B_S the set of boundary pixels in S. The rate of boundary pixels in G matching with boundary pixels in S is given by the classic BR measure. From B_G we define the fuzzy set $B^*_{G \cap S}$ with the membership function $f_{G \cap S}(p_i) = \exp(-\frac{d(p_i - p_{i'})^2}{2\sigma^2})$ where $d(p_i - p_j)$ is the distance between p_i and p_j locations and $p_{i'} = \arg\min_{p_j \in B_S}(d(p_i - p_j))$. The function $f_{G \cap S}$ returns a value in the range $[0, 1]$, a value of 1 meaning a perfect coincidence between an element in B_G and an element in B_S. Finally, we propose the fuzzy boundary recall measure $FBR(S, G) = \frac{1}{|B_G|} \sum_{p \in B_G} f_{G \cap S}(p)$.

4   Experimental Results

We evaluate the ability of the algorithms FZ [3], QS [17], SLIC [1], ERS [7], CRS [2] and WP [8] to satisfy the properties defined in Sect. 2, using the implementations provided by their authors. By design all the tested algorithms satisfy the validity property. The boundary adherence of the superpixels is evaluated using FBR score, the simplicity by computing the average number of neighbors

by superpixel (Nei), the conciseness by giving the average number of superpixels by image (K) and the efficiency by measuring the execution time (T), on a desktop computer with a 2.6 GHz Intel Core i7 processor and 16 Go of RAM. We focus this evaluation on two distinct aspects: the adaptivity and the scale up of the algorithm. To measure adaptivity, we analyze the evolution of mean and standard variation for both FBR and K scores. The fact that HSID contains a majority of big images allows us to check the ability to scale up of each method, by comparing the scores achieved on HSID with the results obtained in previous evaluations. Table 1 gives the mean and the standard deviation achieved by the six algorithms for all these measures. We made 8 tests where methods are configured to respectively produce about 500 (test 1), 700 (test 2), 900 (test 3), 1100 (test 4), 1300 (test 5), 1500 (test 6), 1700 (test 7) and 1900 (test 8) superpixels. Because the execution time strongly depends on the image size, the standard deviation for this measure is high, often similar to the mean. Even if these two measures must be analyzed with caution, they are sufficient to compare the ability of algorithms to scale up. Visual results for the complete dataset and values for each method parameters are available online[5].

Table 1. Quantitative results of the evaluated algorithms.

Test	K	Nei	T	FBR
1	539±94	6±0.11	12±12	0.52±0.17
2	756±135	6±0.12	11±11	0.54±0.17
3	974±184	6±0.12	11±11	0.55±0.16
4	1185±236	6±0.13	10±11	0.57±0.16
5	1390±285	6±0.14	10±10	0.57±0.16
6	1596±338	6±0.15	10±11	0.59±0.16
7	1801±391	6±0.16	10±10	0.59±0.15
8	1998±446	6±0.17	10±10	0.60±0.15

(a) FZ, parameters of [10].

Test	K	Nei	T	FBR
1	478±456	6±0.43	215±213	0.42±0.13
2	662±625	6±0.36	171±168	0.45±0.13
3	859±808	6±0.36	141±140	0.48±0.13
4	1174±1095	6±0.37	116±113	0.51±0.13
5	1443±1341	6±0.37	102±100	0.52±0.13
6	1676±1556	6±0.39	97±95	0.53±0.13
7	2583±2383	6±0.45	79±77	0.57±0.12
8	10030±9129	6±0.72	53±52	0.66±0.1

(b) QS, parameters of [14].

Test	K	Nei	T	FBR
1	474±28	6±0.08	3±3	0.51±0.14
2	663±37	6±0.1	3±3	0.53±0.13
3	860±42	6±0.12	3±3	0.55±0.13
4	1053±54	6±0.12	3±3	0.56±0.13
5	1253±63	6±0.14	3±3	0.56±0.13
6	1446±68	6±0.15	3±3	0.57±0.13
7	1648±79	6±0.14	3±3	0.58±0.13
8	1854±101	6±0.18	3±3	0.59±0.13

(c) SLIC, parameters of [1].

Test	K	Nei	T	FBR
1	500±0	6±0.19	32±34	0.54±0.16
2	700±0	6±0.18	32±35	0.56±0.16
3	900±0	6±0.18	31±33	0.57±0.15
4	1100±0	6±0.17	31±33	0.58±0.15
5	1300±0	6±0.18	31±33	0.59±0.15
6	1500±0	6±0.18	36±39	0.60±0.15
7	1700±0	6±0.18	36±39	0.61±0.15
8	1900±0	6±0.19	36±38	0.61±0.15

(d) ERS, parameters of [14].

Test	K	Nei	T	FBR
1	530±12	6±0.13	16±11	0.21±0.17
2	743±21	6±0.15	17±11	0.23±0.19
3	959±34	6±0.15	19±12	0.28±0.19
4	1167±47	6±0.16	19±12	0.3±0.2
5	1381±41	6±0.14	20±13	0.32±0.21
6	1601±63	6±0.16	21±12	0.34±0.21
7	1797±87	6±0.17	22±13	0.35±0.21
8	2025±92	6±0.16	23±13	0.37±0.22

(e) CRS, parameters of [2].

Test	K	Nei	T	FBR
1	613±11	6±0.09	24±42	0.43±0.17
2	858±20	6±0.07	25±43	0.45±0.17
3	1104±29	6±0.08	26±45	0.46±0.17
4	1346±44	6±0.08	26±46	0.47±0.18
5	1587±38	6±0.07	27±47	0.49±0.17
6	1840±60	6±0.07	27±49	0.50±0.17
7	2069±84	6±0.08	28±50	0.51±0.17
8	2320±77	6±0.08	29±51	0.51±0.17

(f) WP, parameters of [8].

Initially designed as a segmentation method, **FZ** [3] cannot be used with the parameters suggested by its authors. We first use the parameters learned by Stutz *et al.* [14]. But a visual analyze of the produced results shows that they have been over-fitted for BSD and NYU datasets and are not suitable for HSID, where they produce segmentation-like results. A second attempt with the parameters learned by Mathieu *et al.* [10] gives much better results and allows a fair comparison with the rest of the state-of-the-art methods. The analysis of superpixel numbers and FBR scores shows that FZ has a good adaptability. When its parameters are set to produce a lot of superpixels ($K > 1500$), both mean of FBR

---

[5] http://image.ensfea.fr/hsid/.

and standard deviation of K increase, showing that images, that are correctly oversegmented when setting FZ to produce less superpixels (for example 500), are always partitioned in a small number of superpixels. Boundary adherence is satisfactory with FBR scores better than those of SLIC and slightly worse than those of ERS. The only drawback of FZ is its execution time. With a FBR score significantly lower than FZ, SLIC and ERS algorithms, and the longer execution times of all the evaluated methods, **QS** is the first case of algorithm failing to oversegment HSID. A visual analysis of the produced superpixels shows that QS has strong difficulties with images where some boundaries are slightly blurred. Moreover, QS parameters are related to the image size. Thus, the large standard deviation of K is not the consequence of a good adaptivity, but the result of these size-dependent parameters, reducing the superpixel number when dealing with a small image. Evaluation on HSID including images with several millions of pixels, allow to highlight the huge advantage of the linear complexity of **SLIC**, which has an execution times from 3 to 71 times faster than the other algorithms. This main strength is offset by the FBR results, SLIC requiring more superpixels to achieve performance similar to ERS or FZ algorithms. Moreover, the standard deviation for K is low, revealing that, even if configuring SLIC to produce more superpixels reduces boundary adherence errors, this improvement suffers from an oversegmentation of simple (with a few objects) images in much more superpixels than necessary. In other terms, SLIC is not adaptive. **ERS** is the method achieving the best compromise between conciseness and boundary adherence. Unfortunately this result come with the second highest execution time after QS. The second drawback of ERS is its complete lack of adaptability, with a standard deviation of superpixel numbers equals to 0, meaning that the superpixel numbers is fixed by the user, without any possibility for the method to adapt it to the image complexity. Thus, to reduce errors in images with a lot of tiny details, photographs with large homogeneous areas are oversegmented into numerous small superpixels. Conversely, even if with 1900 superpixels, thin elements of some images are not correctly segmented. The algorithm **CRS** is the second case of algorithm failing to oversegment HSID. Even with more than 2000 superpixels, it achieves poor FBR scores, lower than 0.5. This result is easily explained by the study of CRS algorithm. Even with numerous superpixels, the initial partition in regular rectangles, corresponds to a significant error on HSID biggest images. Consequently, the convergence to a more relevant solution by only moving boundary pixels is very slow. In addition, statistical distributions of color inside superpixels are often so wrong that the algorithm remains stuck in local optimum far away from a correct oversegmentation. For example, when multiplying the number of iteration by 100, no visible improvement is shown, but the execution time raises over 2000 s. **WP** method is the last case of method achieving good results on BSD and failing to have similar performance on HSID. While evaluation of Machairas *et al.* [8] shows that WP and SLIC have similar boundary adherence, with execution time lower for WP, FBR results and computational times are, in our evaluation of WP, far away from those of SLIC.

In addition, WP is the only algorithm unable to oversegment the totality of HSID, failing for *img-012*, *img-066* and *img-072*.

5   Conclusion and Perspectives

Evaluation of QS [17], CRS [2] and WP [8] shows that even if these algorithms achieve good results on BSD and NYU datasets, they fail to correctly oversegment HSID images. Moreover, none of the remaining algorithms (FZ, SLIC and ERS) reaches a satisfactory compromise between boundary adherence, conciseness and efficiency. Thus, the proposed dataset HSID shows that is room for improvement and new propositions in oversegmention research area. We are currently working on a method based on a region merging approach. Our goal is to obtain a new algorithm able of adapting to the image content, reaching a compromise between conciseness and boundary adherence, while keeping short execution time. Regarding HSID, even if this dataset is sufficient to make interesting conclusion the state-of-the-art oversegmentation methods, we think that enlarge it with some supplementary images and the associated ground truth should be valuable. We hope that a collaborative effort, will be made. Finally, we invite all researchers working on a new oversegmentation method to not only evaluate their algorithms using previous benchmarks but also to show that they are competitive when dealing with HSID.

References

1. Achanta, A., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: Slic superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
2. Conrad, C., Mertz, M., Mester, R.: Contour-relaxed superpixels. In: Heyden, A., Kahl, F., Olsson, C., Oskarsson, M., Tai, X.-C. (eds.) EMMCVPR 2013. LNCS, vol. 8081, pp. 280–293. Springer, Heidelberg (2013). doi:10.1007/978-3-642-40395-8_21
3. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. Int. J. Comput. Vis. **59**(2), 167–181 (2004)
4. Gould, S., Rodgers, J., Cohen, D., Elidan, G., Koller, D.: Multi-class segmentation with relative location prior. Int. J. Comput. Vis. **80**(3), 300–316 (2008)
5. Irie, K., Tomono, M.: Road recognition from a single image using prior information. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 1938–1945 (2013)
6. Levinshtein, A., Stere, A., Kutulakos, K.N., Fleet, D.J., Dickinson, S.J., Siddiqi, K.: Turbopixels: fast superpixels using geometric flows. IEEE Trans. Pattern Anal. Mach. Intell. **31**(12), 2290–2297 (2009)
7. Liu, M.Y., Tuzel, O., Ramalingam, S., Chellappa, R.: Entropy rate superpixel segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (2011)
8. Machairas, V., Faessel, M., Cárdenas-Peña, D., Chabardes, T., Walter, T., Decenciére, E.: Waterpixels. IEEE Trans. Image Process. **24**(11), 3707–3716 (2015)

9. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: 8th International Conference on Computer Vision, vol. 2, pp. 416–423 (2001)

10. Mathieu, B., Crouzil, A., Puel, J.B.: Segmentation interactive pour l'annotation de photographies de paysages (2016)

11. Papon, J., Abramov, A., Schoeler, M., Worgotter, F.: Voxel cloud connectivity segmentation-supervoxels for point clouds, pp. 2027–2034 (2013)

12. Ren, X., Malik, J.: Learning a classification model for segmentation. In: Ninth IEEE International Conference on Computer Vision, vol. 1, pp. 10–17 (2003)

13. Silberman, N., Hoiem, D., Pushmeet, K., Fergus, R.: Indoor segmentation and support inference from RGBD images, pp. 746–760 (2012)

14. Stutz, D.: Superpixel segmentation: an evaluation. In: German Conference on Pattern Recognition, pp. 555–562 (2015)

15. Tang, D., Fu, H., Cao, X.: Topology preserved regular superpixel. In: IEEE International Conference on Multimedia and Expo, pp. 765–768 (2012)

16. Bergh, M., Boix, X., Roig, G., Capitani, B., Gool, L.: SEEDS: superpixels extracted via energy-driven sampling. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012. LNCS, vol. 7578, pp. 13–26. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33786-4_2

17. Vedaldi, A., Soatto, S.: Quick shift and kernel methods for mode seeking. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5305, pp. 705–718. Springer, Heidelberg (2008). doi:10.1007/978-3-540-88693-8_52

18. Veksler, O., Boykov, Y., Mehrani, P.: Superpixels and supervoxels in an energy optimization framework. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6315, pp. 211–224. Springer, Heidelberg (2010). doi:10.1007/978-3-642-15555-0_16

19. Weikersdorfer, D., Gossow, D., Beetz, M.: Depth-adaptive superpixels. In: 21st International Conference on Pattern Recognition, pp. 2087–2090 (2012)

20. Zhang, L., Yang, Y., Wang, M., Hong, R., Nie, L., Li, X.: Detecting densely distributed graph patterns for fine-grained image categorization. IEEE Trans. Image Process. **25**(2), 553–565 (2016)

21. Zhang, Y., Hartley, R., Mashford, J., Burn, S.: Superpixels via pseudo-boolean optimization. In: CIEEE International Conference on Computer Vision, pp. 1387–1394 (2011)

22. Zhou, C., Liu, C.: Interactive image segmentation based on region merging using hierarchical match mechanism. In: International Conference on Computer Science & Service System, pp. 1781–1784 (2012)

23. Zitnick, C.L., Kang, S.B.: Stereo for image-based rendering using image oversegmentation. Int. J. Comput. Vis. **75**(1), 49–65 (2007)