Quantitative Assessment of Image Segmentation Quality by Random Walk Relaxation Times

Björn Andres[1], Ullrich Köthe[1], Andreea Bonea[1], Boaz Nadler[2,⋆], and Fred A. Hamprecht[1]

[1] University of Heidelberg, Germany
[2] Weizmann Institute of Science, Israel

Abstract. The purpose of image segmentation is to partition the pixel grid of an image into connected components termed segments such that (i) each segment is homogenous and (ii) for any pair of adjacent segments, their union is not homogenous. (If it were homogenous the segments should be merged). We propose a rigorous definition of segment homogeneity which is scale-free and adaptive to the geometry of segments. We motivate this definition using random walk theory and show how segment homogeneity facilitates the quantification of violations of the conditions (i) and (ii) which are referred to as under-segmentation and over-segmentation, respectively. We describe the theoretical foundations of our approach and present a proof of concept on a few natural images.

1 Introduction and Related Work

Image segmentation is an important step in many applications, sometimes even the ultimate goal of the analysis. It remains a challenging problem which requires the search for an optimal partition of the pixel grid of an image. Even under simple model assumptions, this problem is NP-hard [13]. The task of image segmentation has thus been addressed by various constructive algorithms, e.g., watershed segmentation [11] as well as by spectral methods such as normalized cuts [13]. All segmentation algorithms have design parameters which need to be tuned for each specific application. Hence, a quantitative validation of segmentations is all the more important. Given a segmentation, there are two possible types of errors: (i) under-segmentation – a segment contains parts which belong to different regions and should be split; (ii) over-segmentation, two adjacent segments in fact belong to the same region and should be merged. Most image segmentations suffer from at least one if not both types of errors.

It is thus desirable to develop a quantitative tool to assess the quality of each individual segment, and to provide a diagnostic measure of whether a given segment should be split, or whether two adjacent segments should be joined. In an interactive segmentation framework, this measure has the potential to focus user attention on segmentation errors whereas in an automated procedure,

⋆ The research of Boaz Nadler was supported by the Israel Science Foundation (grant 432/06).

J. Denzler, G. Notni, and H. Süße (Eds.): DAGM 2009, LNCS 5748, pp. 502–511, 2009.

incorrect segments can be subjected to further processing. In this paper, we develop such a measure. A key feature of this measure is scale-invariance. It can therefore be applied to both small and large segments. As a proof of concept, we present results on segmentations of Berkeley database images [8]. Both watershed segmentation [11] and a graph-based segmentation algorithm [2] are used to illustrate that the method works with any segmentation. The only prerequisite is a *boundary indicator function* b giving the probability $b(r_1, r_2)$ that adjacent pixels r_1, r_2 are separeted by a boundary. A simple way to obtain such a function is to interpolate the output of an edge detector to inter-pixel positions and then normalize the result. Given an image annotation, i.e. a function a with $a(r, c)$ the probability that the pixel r is associated with class c (e.g. sky, ground, car, building), $b(r_1, r_2)$ can be a distance of the distributions $a(r_1, \cdot)$ and $a(r_2, \cdot)$.

The measure of segment homogeneity we propose is based on the spectral analysis of random walks on weighted graphs. Random walks on graphs are the foundation of diffusion distances and diffusion maps [14,1] which have been applied for clustering, dimensionality reduction and signal denoising. The close relation between random walks and normalized cuts [13] is noted in [9,10]. Random walks have recently been used for interactive image segmentation [4,5,6]. Similar to [4,5,6], we consider weighted graphs in which vertices represent image pixels and edges connect neighboring pixels. The novelty we propose is to relate the relaxation time of a random walk on a weighted graph to the relaxation time on the topologically identical graph with all weights set to an equal value. This approach facilitates scale-invariance and adaptivity to the geometry of segments.

2 Segment Homogeneity

In this section, the spectral analysis of random walks on graphs is outlined first, followed by the description of the particular graphs we consider for segmentation analysis. A rigorous definition of segment homogeneity is given at the end.

Consider an undirected weighted graph $G = (V, E, s)$ with the *edge weight function* $s : E \to (0, 1]$. Interpret this function as a measure of similarity of the two vertices which are incident to the edge. Let V be finite and identify each vertex with a positive integer, $V = \{1, \ldots, n\}$. Let $S \in \mathbb{R}^{n \times n}$ be the *similarity matrix* of the graph G with

$$S_{jk} = \begin{cases} s(\{j,k\}) & \text{if } \{j,k\} \in E \\ \delta_{jk} & \text{otherwise} \end{cases} \tag{1}$$

In addition, define the *degree* of any vertex $j \in V$ as the sum of weights of all edges which are incident to j, i.e.

$$\deg : V \to \mathbb{R} \quad \text{such that} \quad \forall j \in V : \deg(j) = \sum_{k \in V} S_{jk} \ . \tag{2}$$

Let $D \in \mathbb{R}^{n \times n}$ be the diagonal matrix with $\forall j \in V : D_{jj} = \deg(j)$ and let

$$M = D^{-1}S \ . \tag{3}$$

M is obtained from S by an L1-normalization of each row and is thus a row-stochastic matrix. M is adjoint to the symmetric matrix $M_s = D^{1/2}SD^{-1/2}$. Thus, M and M_s share the same eigenvalues. Moreover, since M_s is symmetric, the matrix has n real eigenvalues $\lambda_1, \ldots, \lambda_n$ whose corresponding eigenvectors $v_1, \ldots, v_n$ form an orthonormal basis of $\mathbb{R}^n$. The left and right eigenvectors of M denoted ϕ_j and ψ_j are related to those of M_s according to

$$\phi_j = v_j D^{1/2}, \quad \psi_j = v_j D^{-1/2} \ . \tag{4}$$

As the eigenvectors v_j are orthonormal under the standard innner product in $\mathbb{R}^n$, it follows that ϕ_j and ψ_j are bi-orthonormal, i.e. $\langle \phi_j, \psi_k \rangle = \delta_{jk}$. One normalized eigenvector of M is $(1, \ldots, 1)^T/\sqrt{n}$ and the corresponding eigenvalue is 1. In addition, no eigenvalue of a stochastic matrix can be larger than one in magnitude (this is a consequence of the Gelfand spectral radius theorem, cf. [12]). Thus, the eigenvalues $\lambda_1, \ldots, \lambda_n$ can be ordered such that

$$1 = |\lambda_1| \geq \ldots \geq |\lambda_n| \geq 0 \ . \tag{5}$$

A *random walk* on the weighted graph G is a discrete stochastic process whose states are the vertices of G. Let $P(X_t = j)$ denote the probability that the process attains the state $j \in V$ at the discrete point $t \in \mathbb{N}_0$ in time. Assume that the process is Markovian, i.e. the probability of moving from a vertex $j \in V$ to a vertex $k \in V$ does not depend on the history of the process,

$$\forall t \in \mathbb{N}_0 : \quad P(X_{t+1}|X_t, \ldots, X_0) = P(X_{t+1}|X_t) \ . \tag{6}$$

In addition, let the transition probabilities be given by the matrix M,

$$\forall t \in \mathbb{N}_0 \ \forall j, k \in V : \quad P(X_{t+1} = k|X_t = j) = M_{jk} \ . \tag{7}$$

The evolution of the random walk is then given by the repeated multiplication of the transition matrix M with the probability row vector of the initial state,

$$P(X_t) = P(X_{t-1})M = P(X_{t-2})M^2 = \ldots = P(X_0)M^t \ . \tag{8}$$

This evolution is elucidated by the spectral decomposition of M:

$$M^t = \left(\sum_{j=1}^{n} \lambda_j \phi_j \psi_j^T \right)^t = \sum_{j=1}^{n} \lambda_j^t \phi_j \psi_j^T = \phi_1 \psi_1^T + \sum_{j=2}^{n} \lambda_j^t \phi_j \psi_j^T \ . \tag{9}$$

If, for any sufficiently large number of steps, any vertex in the graph is reachable from any other vertex via a sequence of edges, the process is called *irreducible*. If, for any large enough number of steps, the probability of returning to the start vertex is non-zero, the process is called *aperiodic*. If a process is both aperiodic

and irreducible, there exists a $t_0 \in \mathbb{N}$ such that for all $t \in \mathbb{N}^{>t_0}$, M^t is a positive matrix. It then follows from Perron's theorem [7] that $|\lambda_2| < 1$ and thus $M^t \to \phi_1 \psi_1^T$ as $t \to \infty$. The rate of convergence depends on $|\lambda_2|$. We refer to

$$\tau := \frac{1}{1 - |\lambda_2|} \tag{10}$$

as the *characteristic relaxation time* of the random walk. Since M is a square, sparse matrix, λ_2 can be efficiently computed using the Lanczos algorithm [3].

Consider the pixel grid Γ and the 8-neighborhood system of pixels which we denote by $\sim$. For any subset of pixels $U \subseteq \Gamma$, let pairs(U) be the set of all (unordered) pairs of pixels of U which are 8-neighbors, pairs$(U) := \{\{u, u'\} \subseteq U | u \sim u'\}$. We consider segmentation algorithms which output a number $m \in \mathbb{N}$ of segments and a function seg : $\Gamma \to \{1, \ldots, m\}$ which maps any pixel to the index of a segment. Our goal is to provide a quantitative measure of segment homogeneity. To this end, we consider two different random walks on each candidate segment. First, we study the geometric properties of the segment, and second the possible presence of significant inner boundaries in it. For any segment $j \in \{1, \ldots, m\}$, we construct the weighted graph $G_j = (V_j, E_j, s_j)$ where $V_j = \text{seg}^{-1}(j)$ consists of all pixels of the segment, $E_j = \text{pairs}(V_j)$ is the set of all pairs of pixels in the segment which are 8-neighbors, and s_j is a similarity measure. An example is depicted in Fig. 1. First, we chose s_j such that all edge weights are equal, which yields a graph that depends exclusively on the size and shape of the segment and is independent of the boundary indicator function. We term this graph the *geometric graph* of the segment and denote the relaxation time of the associated random walk by τ_g (*geometric relaxation time*). Second, we define s_j depending on the boundary indicator. Since we consider similarity graphs (as opposed to the dissimilarity measured by the boundary indicator), we invert the indicator by the function $f(x) = 1 - (1 - s_0)x$. The design parameter $s_0 \in [0, 1]$ sets the minimal permitted similarity of neighboring pixels. A positive value $s_0 > 0$ ensures irreducibility of the random walk associated with the weighted graph. Such a simple linear transform is sufficient if the boundary indicator function is appropriately scaled. Otherwise, the exponential function $f(x) = \exp(-\alpha x)$ can be used with α adapted to the scale of b. We denote the relaxation time of the random walk associated with the weighted graph by τ_w and define segment homogeneity as the ratio

$$H := \frac{\tau_w}{\tau_g} \ . \tag{11}$$

Under-segmentation is quantified by computing H for each segment. If a segment is split by a boundary, the random walk on the weighted graph stays trapped inside one sub-region, or "well", of the segment for a long time and only occasionally escapes across a boundary into another sub-region. Thus, it takes longer for the random walk to converge to the stationary distribution on the weighted graph than it does on the graph with equal weights. Hence, $H \gg 1$ is indicative of under-segmentation. Conversely, $H = 1$ holds for homogenous segments, regardless of their size and shape. H thus constitutes a scale-free measure of segment homogeneity.

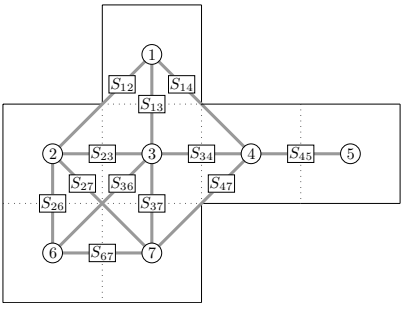

Fig. 1. Similarity graph of a simple segment consisting of seven pixels. Each pixel corresponds to one vertex (enumerated circles). Edges (gray lines) connect any two pixels which are 8-neighbors. A similarity S_{jk} is associated with every edge.

In over-segmentation analysis, all pairs (V_1, V_2) of adjacent segments are considered and for each of these pairs, H is computed on the merged segment $V_1 \cup V_2$. A homogenous pair is indicative of over-segmentation. V_1 and V_2 should thus be merged if $H \approx 1$.

3 Experiments

To illustrate the principle, the two segments depicted in the rows of Fig. 2 exemplify the under-segmentation analysis. In the first column, the segment itself is depicted in black, indicating that the boundary indicator is zero throughout the segment. Consequently, $\tau_g = \tau_w$ and hence, $H = 1$. In the second and third column, the segment is split by a boundary depicted in gray. In these cases, $H > 1$ quantifies the degree of under-segmentation. In the fourth column, there is a pronounced boundary which does however not split the segment. H is very close to one in this case because τ_w differs only slightly from τ_g. This property of H is essential for segmentation analysis where split segments must not be confused with segments that include non-splitting boundaries.

In order to quantify the under-segmentation of natural images of the Berkeley database [8], the boundary indicator function was taken to be the normalized squared gradient magnitude of the three channels of the L*a*b* color space. Figures 4 and 5 show segmentations obtained from this boundary indicator by means of a seeded region growing variant of the watershed algorithm [11]. Seeds were set at those points of a regular grid (with a grid point distance of 15 pixels) where the normalized boundary indicator did not exceed 0.4. Under-segmentation thus occurs especially if no seed is placed inside an object in the image. In Fig. 4, a coarse, graph-based segmentation (cf. [2]) is shown in addition, stressing that the measure of segment homogeneity is scale-free and can be used with any segmentation algorithm. Along with the segmentations, H is depicted for each segment in shades of gray. Under-segmentation occurs in those segments for which H is much larger than 1. The distribution of relaxation times for the watershed segmentation shown in Fig. 4 is depicted in Fig. 3a. Homogenous segments are concentrated along the diagonal which means that the

relaxation times are similar ($H \approx 1$). Segments which are to be split lie significantly above the diagonal ($H \gg 1$). It can happen that $\tau_w < \tau_g$ if the weighted graph distributes the random walk faster than equal weights do. Clearly, this is not the case if a segment is split, so $H < 1$ can be taken as a strong indication of not having under-segmentation. Over-segmentation as depicted in Fig. 6 is a well-known property of the watershed algorithm where regions are grown from all local minima of the boundary indicator function. Along with the segmentation, a boundary map is depicted. The shade of a boundary separating two segments V_1 and V_2 is proportional to the segment homogeneity H of the merged pair $V_1 \cup V_2$. White (removed) boundaries indicate over-segmentation. The corresponding distribution of relaxation times is depicted in Fig. 3b.

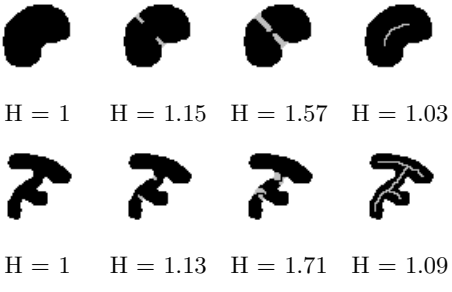

Fig. 2. Illustration of two segments (rows) and four different boundary indicator functions on these segments (columns). The gray edges inside of the segments indicate a boundary probability of 0.6. Segments which are split by boundaries exhibit a segment homogeneity $H > 1$ whereas $H \approx 1$ for segments which are not split.

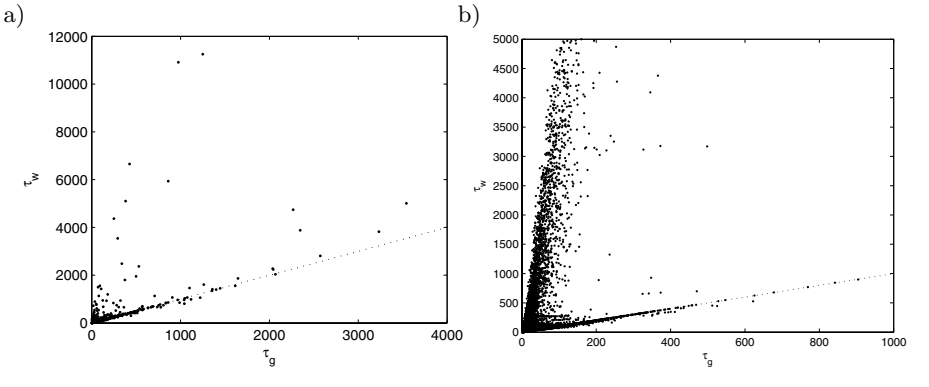

Fig. 3. a) Relaxation times τ_g and τ_w of all segments of Fig. 4. Points close to the diagonal correspond to homogenous segments. Points significantly above the diagonal indicate under-segmentation. b) Relaxation times of all merged pairs of adjacent segments of Fig. 6. Points on the diagonal indicate over-segmentation.

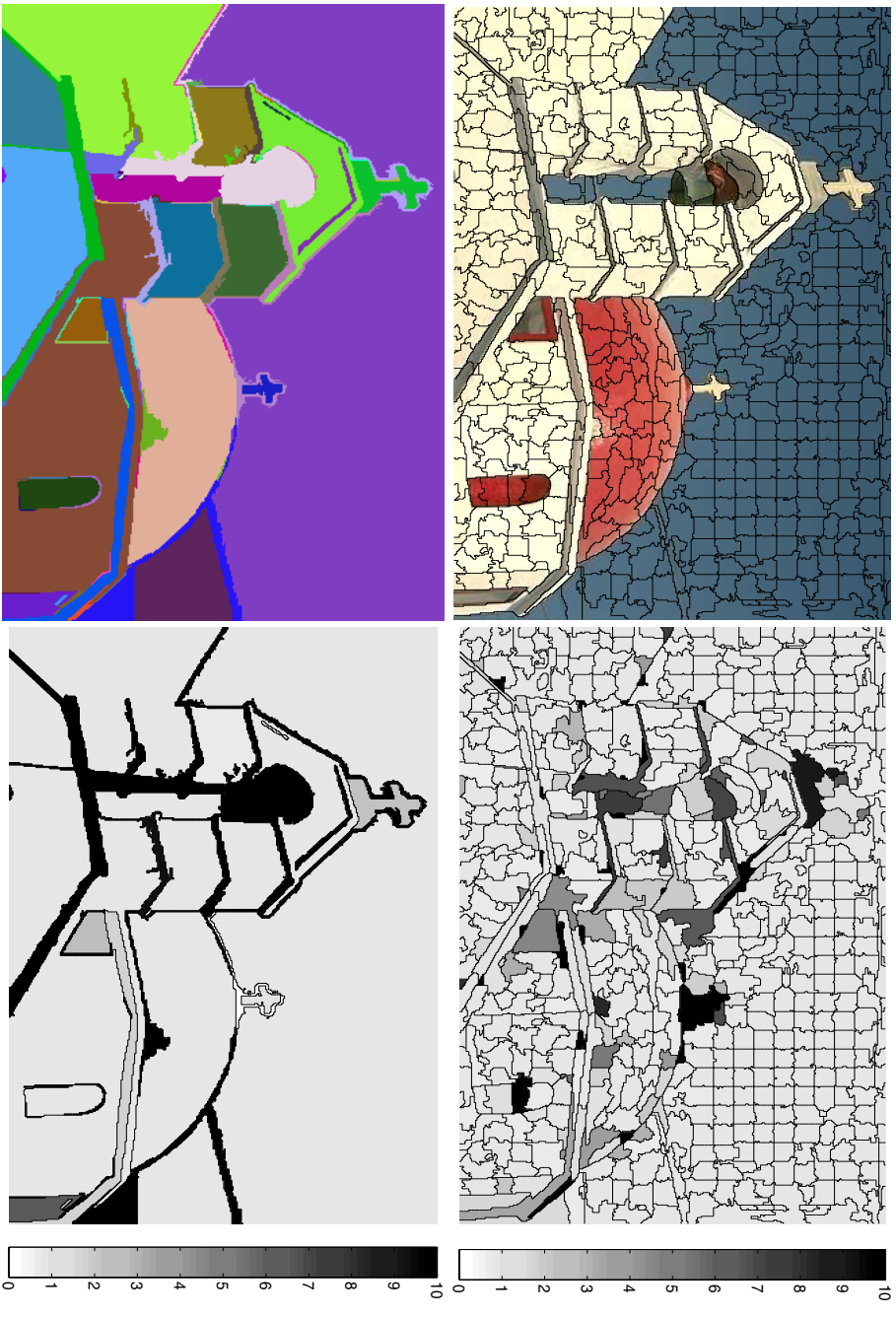

Fig. 4. Quantification of under-segmentation. One fine watershed-segmentation and one coarser graph-based segmentation [2] are shown. For each segment, its homogeneity H is depicted in shades of gray. Under-segmentation occurs in those darker segments for which H is significantly larger than 1.

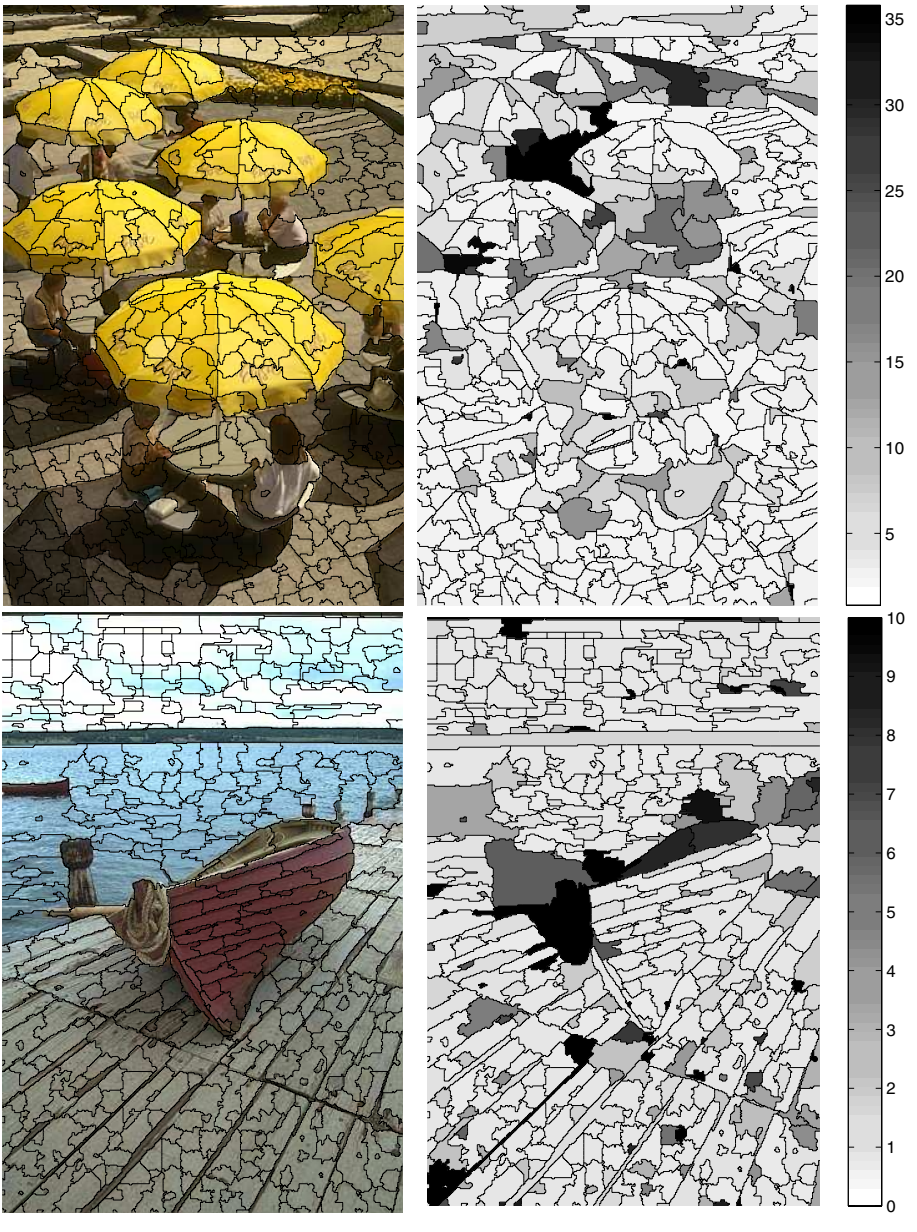

Fig. 5. Quantification of under-segmentation (continued)

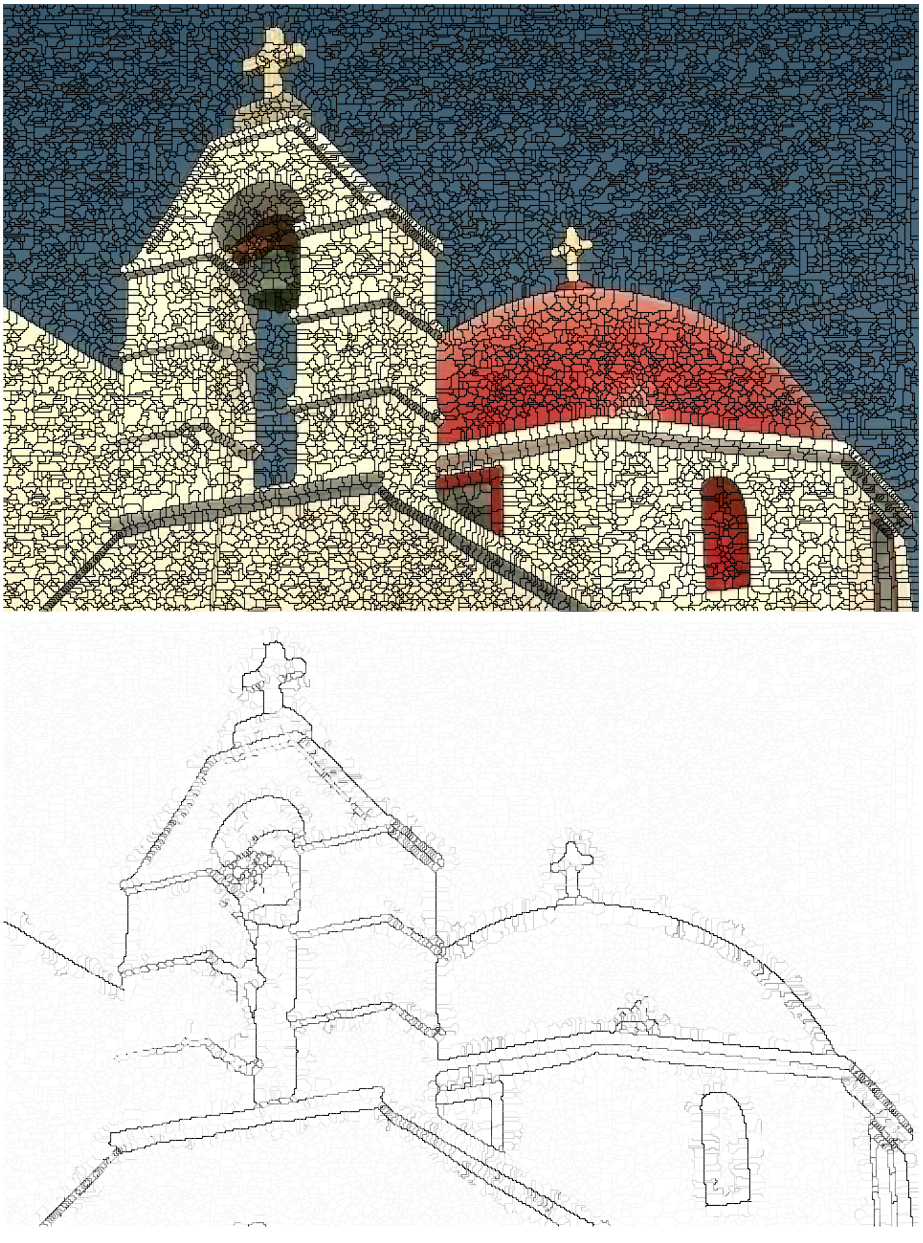

Fig. 6. Quantification of over-segmentation. Each boundary in the lower plot corresponds to a pair (V_1, V_2) of segments which are adjacent via that boundary. The shade of a boundary is proportional to the segment homogeneity H of the merged pair $V_1 \cup V_2$. Black corresponds to high values, white corresponds to $H \leq 1$. White (removed) boundaries indicate over-segmentation.

4 Conclusion

In this paper, we developed a quantitative measure of segment homogeneity, based on random walks. Random walks on two different segment graphs were considered: A weighted graph with weights obtained from a boundary indicator function and a graph where all weights are set to an equal value. Segment homogeneity was defined as the ratio of the relaxation times of these two random walks. This definition is scale-free and adaptive to the geometry of segments. It facilitates the quantitative assessment of segmentation quality, in particular the quantification of under- and over-segmentation. We are currently applying the concept to 3D segmentations and explore its potential as a criterion in an unsupervised split-and-merge segmentation algorithm.

References

1. Coifman, R.R., Lafon, S., Lee, A.B., Maggioni, M., Nadler, B., Warner, F., Zucker, S.W.: Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. PNAS 102(21), 7426–7431 (2005)
2. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. Int. J. Comput. Vision 59(2), 167–181 (2004)
3. Golub, G.H., van Loan, C.F.: Matrix Computations, 3rd edn. The Johns Hopkins University Press, Baltimore (1996)
4. Grady, L.: Random walks for image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence 28, 1768–1783 (2006)
5. Grady, L., Sinop, A.K.: A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In: ICCV. IEEE, Los Alamitos (2007)
6. Grady, L., Sinop, A.K.: Fast approximate random walker segmentation using eigenvector precomputation. In: CVPR, IEEE, Los Alamitos (2008)
7. MacCluer, C.R.: The many proofs and applications of Perron's theorem. SIAM Rev. 42(3), 487–498 (2000)
8. 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: ICCV, vol. 2, pp. 416–423 (2001)
9. Meila, M., Shi, J.: Learning segmentation by random walks. In: NIPS, pp. 873–879. MIT Press, Cambridge (2000)
10. Nadler, B., Lafon, S., Kevrekidis, I., Coifman, R.: Diffusion maps, spectral clustering and eigenfunctions of fokker-planck operators. In: NIPS, pp. 955–962. MIT Press, Cambridge (2005)
11. Roerdink, J., Meijster, A.: The watershed transform: definitions, algorithms and parallelization strategies. Fundamenta Informaticae 41(1-2), 187–228 (2000)
12. Rudin, W.: Functional Analysis. McGraw-Hill, New York (1991)
13. Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Trans. Pattern Analysis and Machine Intelligence 22(8), 888–905 (2000)
14. Singer, A., Shkolnisky, Y., Nadler, B.: Diffusion interpretation of non-local neighborhood filters for signal denoising. SIAM J. on Imaging Sciences 2, 118–139 (2009)