Computer Vision and Image Understanding xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Robust superpixels using color and contour features along linear path

Rémi Giraud[*,a,b,c,d], Vinh-Thong Ta[a,b,e], Nicolas Papadakis[c,d]

[a] Univ. Bordeaux, LaBRI, UMR 5800, PICTURA, Talence F-33400, France
[b] CNRS, LaBRI, UMR 5800, PICTURA, Talence F-33400, France
[c] Univ. Bordeaux, IMB, UMR 5251, Talence F-33400, France
[d] CNRS, IMB, UMR 5251, Talence F-33400, France
[e] Bordeaux INP, LaBRI, UMR 5800, PICTURA, Talence F-33400, France

ARTICLE INFO

Keywords:
Superpixels
Linear path
Segmentation
Contour detection

ABSTRACT

Superpixel decomposition methods are widely used in computer vision and image processing applications. By grouping homogeneous pixels, the accuracy can be increased and the decrease of the number of elements to process can drastically reduce the computational burden. For most superpixel methods, a trade-off is computed between 1) color homogeneity, 2) adherence to the image contours and 3) shape regularity of the decomposition. In this paper, we propose a framework that jointly enforces all these aspects and provides accurate and regular Superpixels with Contour Adherence using Linear Path (SCALP). During the decomposition, we propose to consider color features along the linear path between the pixel and the corresponding superpixel barycenter. A contour prior is also used to prevent the crossing of image boundaries when associating a pixel to a superpixel. Finally, in order to improve the decomposition accuracy and the robustness to noise, we propose to integrate the pixel neighborhood information, while preserving the same computational complexity. SCALP is extensively evaluated on standard segmentation dataset, and the obtained results outperform the ones of the state-of-the-art methods. SCALP is also extended for supervoxel decomposition on MRI images.

1. Introduction

The use of superpixels has become a very popular technique for many computer vision and image processing applications such as: object localization (Fulkerson et al., 2009), contour detection (Arbelaez et al., 2011), face labeling (Kae et al., 2013), data associations across views (Sawhney et al., 2014), or multi-class object segmentation (Giraud et al., 2017b; Gould et al., 2008; 2014; Tighe and Lazebnik, 2010; Yang et al., 2010). Superpixel decomposition methods group pixels into homogeneous regions, providing a low-level representation that tries to respect the image contours. For image segmentation, where the goal is to split the image into similar regions according to object, color or texture priors, the decomposition into superpixels may improve the segmentation accuracy and decrease the computational burden (Gould et al., 2014). Contrary to multi-resolution approaches, that decrease the image size, superpixels preserve the image geometry, since their boundaries follow the image contours. Hence, the results obtained at the superpixel level may be closer to the ground truth result at the pixel level.

Many superpixel methods have been proposed using various techniques. Although the definition of an optimal decomposition depends on the tackled application, most methods tend to achieve the following properties. First, the boundaries of the decomposition should adhere to the image contours, and superpixels should not overlap with multiple objects. Second, the superpixel clustering must group pixels with homogeneous colors. Third, the superpixels should have compact shapes and consistent sizes. The shape regularity helps to visually analyze the image decomposition and has been proven to impact application performances (Reso et al., 2013; Strassburg et al., 2015; Veksler et al., 2010). Finally, since superpixels are usually used as a preprocessing step, the decomposition should be obtained in limited computational time and allow the control of the number of produced elements.

To achieve the aforementioned properties, most state-of-the-art methods compute a trade-off between color homogeneity and shape regularity of the superpixels. Nevertheless, some approaches less consider the regularity property and can produce superpixels of highly irregular shapes and sizes. In the following, we present an overview of the most popular superpixel methods, defined as either irregular or regular ones. Note that although some methods can include terms into their models to generate for instance more regular results, e.g., Van den Bergh et al. (2012), we here consider methods in their default

* Corresponding author at: LaBRI, 351 cours de la libération, Talence, France.
E-mail address: remi.giraud@labri.fr (R. Giraud).

https://doi.org/10.1016/j.cviu.2018.01.006
Received 1 July 2017; Received in revised form 13 January 2018; Accepted 15 January 2018
1077-3142/

settings, as described by the authors.

The regularity criteria can be seen as the behavior to frequently produce irregular regions, in terms of both shapes and sizes (Giraud et al., 2017c). Methods such as Felzenszwalb and Huttenlocher (2004) and Vedaldi and Soatto (2008) generate very irregular regions in terms of both size and shape while SLIC can generate a few irregular shapes but their sizes are constrained into a fixed size window.

1.1. Irregular superpixel methods

With irregular methods, superpixels can have very different sizes and stretched shapes. For instance, small superpixels can be produced, without enough pixels to compute a significant descriptor. Too large superpixels may also overlap with several objects contained in the image. First segmentation methods, such as the watershed approach, e.g., Vincent and Soille (1991), compute decompositions of highly irregular size and shape. Methods such as Mean shift (Comaniciu and Meer, 2002) or Quick shift (Vedaldi and Soatto, 2008) consider an initial decomposition and perform a histogram-based segmentation. However, they are very sensitive to parameters and are obtained with high computational cost (Vedaldi and Soatto, 2008). Another approach considers pixels as nodes of a graph to perform a faster agglomerative clustering (Felzenszwalb and Huttenlocher, 2004). These methods present an important drawback: they do not allow to directly control the number of superpixels. This is particularly an issue when superpixels are used as a low-level representation to reduce the computational time.

The SEEDS method (Van den Bergh et al., 2012) proposes a coarse-to-fine approach starting from a regular grid. However, this method may provide superpixels with irregular shapes. Although a compactness constraint can be set to compute regular superpixels, the authors report degraded results of decomposition accuracy with such approach.

1.2. Regular superpixel methods

For superpixel-based object recognition methods, e.g., Gould et al. (2008) and Gould et al. (2014), or video tracking, e.g., Reso et al. (2013) and Wang et al. (2011), the use of regular decompositions is mandatory, i.e., decompositions with superpixels having approximately the same size and compact shapes. For instance, for superpixel-based video tracking applications, the tracking of object trajectories within a scene is improved with consistent decompositions over time (Chang et al., 2013; Reso et al., 2013).

Most of the regular methods consider an initial regular grid, allowing to set the number of superpixels, and update superpixels boundaries while applying spatial constraints. Classical methods are based on region growing, such as Turbopixels (Levinshtein et al., 2009) using geometric flows, or eikonal-based methods, e.g., ERGC (Buyssens et al., 2014), while other approaches use graph-based energy models (Liu et al., 2011; Veksler et al., 2010). In Machairas et al. (2015), a watershed algorithm is adapted to produce regular decompositions using a spatially regularized image gradient. Similarly to SEEDS (Van den Bergh et al., 2012), a coarse-to-fine approach has recently been proposed in Yao et al. (2015), producing highly regular superpixels.

The SLIC method (Achanta et al., 2012) performs an iterative accurate clustering, while providing regular superpixels, in order of magnitude faster than graph-based approaches (Liu et al., 2011; Veksler et al., 2010). The SLIC method has been extended in several recent works, e.g., Chen et al. (2017); Huang et al. (2016); Rubio et al. (2016); Zhang et al. (2016) and Zhang and Zhang (2017). However, it can fail to adhere to image contours, as for other regular methods, e.g., Levinshtein et al. (2009) and Yao et al. (2015), since it is based on simple local color features and globally enforces the decomposition regularity using a fixed trade-off between color and spatial distances.

1.3. Contour constraint

In the literature, several works have attempted to improve the decomposition performances in terms of contour adherence by using gradient or contour prior information. In Mori et al. (2004), a contour detection algorithm is used to compute a pre-segmentation using the normalized cuts algorithm (Shi and Malik, 2000). The segmentation may accurately guide the superpixel decomposition, but such approaches based on normalized cuts are computationally expensive (Mori et al., 2004). Moreover, the contour adherence of the produced decompositions are far from state-of-the-art results (Achanta et al., 2012). In Moore et al. (2008), the superpixel decomposition is constrained to fit to a grid, also called superpixel lattice. The decomposition is then refined using graph cuts. However, this method is very dependent on the used contour prior. Moreover, although the superpixels have approximately the same sizes, they have quite irregular shapes and may appear visually unsatisfactory.

In Machairas et al. (2015), the image gradient information is used to constrain the superpixel boundaries, but the results on superpixel evaluation metrics are lower than the ones of SLIC (Achanta et al., 2012). In Zhang et al. (2016), the local gradient information is considered to improve the superpixel boundaries evolution. However, the computational cost of the method is increased by a $10\times$ order of magnitude compared to SLIC.

1.4. Segmentation from contour detection

Contour detection methods generally do not enforce the contour closure. To produce an image segmentation, a contour completion step is hence necessary. Many contour completion methods have been proposed (see for instance Arbelaez et al., 2011 and references therein). This step may improve the accuracy of the contour detection, since objects are generally segmented by closed curves.

Methods such as Arbelaez and Cohen (2008) and Arbelaez et al. (2009), propose a hierarchical image segmentation based on contour detection. This can be considered as a probability contour map, that produces a set of closed curves for any threshold. Although such methods enable to segment an image from a contour map, they do not allow to control the size, the shape and the number of the produced regions, while most superpixel methods enable to set the number of superpixels which approximately have the same size. Moreover, the performances of the contour detection is extremely dependent on the fixed threshold parameter, which depends on the image content (Arbelaez et al., 2009). Hence, they are mainly considered as segmentation methods and cannot be considered as relevant frameworks to compute superpixel decompositions.

1.5. Robustness to noise

Superpixel decompositions are usually used as a pre-processing step in many computer vision applications. Therefore, they tend to be applied to heterogeneous images that can suffer from noise. Moreover, image textures and high local gradients may also mislead the superpixel decomposition. Most of the state-of-the-art superpixel methods are not robust to noise, and provide degraded decompositions when applied to slightly noised images or images with low resolution. With such approaches, a denoising step is necessary to compute a relevant decomposition. For instance, the watershed approach of Machairas et al. (2015) uses a pre-filtering step to smooth local gradients according to the given size of superpixels. Nevertheless, this step is only designed to smooth local gradients of initial images and the impact of this filtering is not reported (Machairas et al., 2015).

1.6. Contributions

In this paper, we propose a method that produces accurate, regular and robust Superpixels with Contour Adherence using Linear Path