ARTICLE IN PRESS

ISPRS Journal of Photogrammetry and Remote Sensing xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Toward better boundary preserved supervoxel segmentation for 3D point clouds☆

Yangbin Lin[a], Cheng Wang[b,*], Dawei Zhai[b], Wei Li[b], Jonathan Li[c]

[a] *Computer Engineering College, Jimei University, Xiamen, China*
[b] *Fujian Key Laboratory of Sensing and Computing for Smart Cities, Department of Computer Science, Xiamen University, Xiamen FJ 361005, China*
[c] *Department of Geography & Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada*

A R T I C L E I N F O

Keywords:
Supervoxel segmentation
Point clouds
Subset selection
Over-segmentation

A B S T R A C T

Supervoxels provide a more natural and compact representation of three dimensional point clouds, and enable the operations to be performed on regions rather than on the scattered points. Many state-of-the-art supervoxel segmentation methods adopt fixed resolution for each supervoxel, and rely on the initialization of seed points. As a result, they may not preserve well the boundaries of the point cloud with a non-uniform density. In this paper, we present a simple but effective supervoxel segmentation method for point clouds, which formalizes supervoxel segmentation as a subset selection problem. We develop an heuristic algorithm that utilizes local information to efficiently solve the subset selection problem. The proposed method can produce supervoxels with adaptive resolutions, and dose not rely the selection of seed points. The method is fully tested on three publicly available point cloud segmentation benchmarks, which cover the major point cloud types. The experimental results show that compared with the state-of-the-art supervoxel segmentation methods, the supervoxels extracted using our method preserve the object boundaries and small structures more effectively, which is reflected in a higher boundary recall and lower under-segmentation error.

1. Introduction

As with superpixels in 2D image processing, the use of supervoxels greatly reduces the number of points. This is beneficial to applications that are time consuming with the original 3D points. Moreover, supervoxels provide a more natural and compact representation for 3D points, which enables operations (such as feature computing) to be performed on regions rather than on scattered points. For these reasons, supervoxels have become increasingly popular in many 3D remote sensing applications, such as object detection (Guan et al., 2016; Wang et al., 2015), semantic labeling (Luo et al., 2016), and saliency detection (Yun and Sim, 2016).

Here, we define the supervoxel as a compact point cluster, which is slightly different from the one in Papon et al. (2013). The general desirable properties of superpixel segmentation are also suitable for supervoxel segmentation. First, supervoxel segmentation should preserve object boundaries, and each supervoxel should overlap with only one object. Second, supervoxel segmentation must be efficient, and at least should not reduce the achievable performance of an application that is dependent on it. Third, each supervoxel should have a regular shape, which is convenient for subsequent applications.

Many state-of-the-art supervoxel segmentation methods adopt fixed resolution for each supervoxel, and rely on initialization of seed points. As a result, they may not preserve well the boundaries of the point cloud with a non-uniform density. In this paper, we formalize supervoxel segmentation as a subset selection problem, and present a simple but effective method to solve the problem. The major advantage of our method is that it adopts an adaptive resolution for each supervoxel and can preserve object boundaries more effectively than existing methods. An example is presented in Fig. 1, the supervoxels extracted by the proposed method better adhere to the ground-truth boundaries, even for road curbs with slight height differences (see the red[1] boxes in Fig. 1(c)).

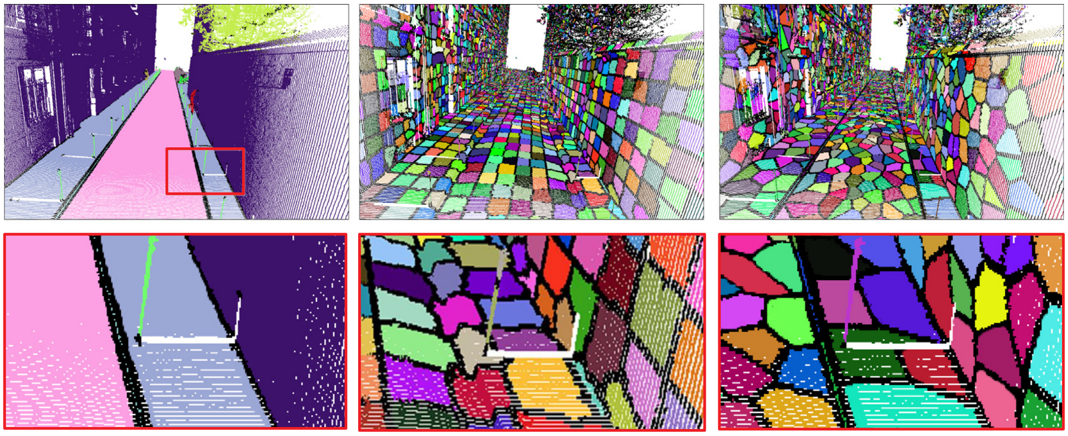
The main contributions of this paper are as follows:

First, we formalize supervoxel segmentation as a subset selection problem. Our formalization involves an explicit energy function, which can be optimized directly. Second, in order to minimize the energy function for subset selection, we propose a simple but effective method that does not require seed points initialization and does not contain internal parameters. Finally, the proposed method significantly outperforms the state-of-the-art supervoxel methods with respect to

☆ This work was supported by the National Science Foundation of China (Project No. 61701191, 41471379, and U1605254).
* Corresponding author.
E-mail address: cwang@xmu.edu.cn (C. Wang).
[1] For interpretation of color in 'Fig. 1', the reader is referred to the web version of this article.

https://doi.org/10.1016/j.isprsjprs.2018.05.004
Received 14 July 2017; Received in revised form 24 April 2018; Accepted 8 May 2018
0924-2716/

Please cite this article as: Lin, Y., ISPRS Journal of Photogrammetry and Remote Sensing (2018), https://doi.org/10.1016/j.isprsjprs.2018.05.004



(a) Point cloud with the groud-truth labels (b) VCCS (c) Proposed

Fig. 1. Comparison results for VCCS (Papon et al., 2013) and the proposed method. As emphasized in red boxes, the proposed method can preserve the boundaries more effectively.

boundary recall and under-segmentation error metrics on three publicly available point cloud segmentation benchmarks.

2. Related work

Unlike superpixel segmentation, which is already a well studied topic in image processing (Moore et al., 2008; Veksler et al., 2010; Liu et al., 2011; Bergh et al., 2013), supervoxel segmentation remains in the development stage.

In video and 3D image segmentation, a supervoxel is usually defined as a stack of 2D image regions (Moore et al., 2008; Xu and Corso, 2012; Zhou et al., 2015). In this case, many superpixel segmentation methods can be directly extended to compute supervoxels. Moore et al. (2008) presented a graph-based superpixel segmentation method, where superpixels are iteratively partitioned from a 2D grid graph by horizontal and vertical cutting. The authors also extended this method to the computation of supervoxels on a 3D grid graph, which can be employed for video over-segmentation. Veksler et al. (2010) presented a framework for both superpixel and supervoxel segmentation. They formulated the superpixel or supervoxel segmentation problem as an energy minimization problem, and solved it using graph cut. Weikersdorfer et al. (2012) transferred superpixels to 3D space, by taking into account depth information for RGB-D image over-segmentation. This method has been further extended to RGB-D video over-segmentation. Zhou et al. (2015) proposed a multiscale superpixel and supervoxel algorithm using hierarchical edge weighted centroidal Voronoi tessellation. Here, superpixels or supervoxels in higher levels are clustered from superpixels or supervoxels in lower levels. In addition to video segmentation, Picciau et al. (2015) develop an adaptation of the SLIC superpixel algorithm (Achanta et al., 2012) for tetrahedral mesh over-segmentation.

However, the methods above are designed for regular data, where the primitives are uniformly distributed. More related to the proposed method is the VCCS algorithm developed in Papon et al. (2013). It first voxelizes the point cloud using octree, and then extracts the initial supervoxels by evenly partitioning the 3D space. These initial supervoxels are then grown using the local k-means clustering method (Achanta et al., 2012). VCCS is reported to be highly efficient, and achieves reasonably good results on RGB-D test data. However, the results of VCCS depends on the setting of voxel resolution. For the point cloud with non-uniform density (typically acquired by the current laser devices), more than one object can overlap with the same voxel. In this case, VCCS may not preserve well the object boundaries.

To make supervoxels conform better to object boundaries, Song et al. (2014) present a boundary-enhanced supervoxel segmentation (BESS) method. The method has two stages. In the first stage, it detects the boundary points by estimating the discontinuity of consecutive points along the scan line. In the second stage, it constructs a neighborhood graph that excludes the edges connected by boundary points, and then performs a clustering process on the graph to segment the point clouds into supervoxels. Although BESS can be used for outdoor scene data with the non-uniform density, it depends on the assumption that the points are sequentially ordered in one direction. This assumption greatly reduces the practicality of BESS method to general point cloud data.

3. Problem formulation

Given a point set $\mathscr{P} = \{p_1,\ldots,p_N\}$ with N points, the partitioning of $\mathscr{P}$ into K supervoxels $\mathscr{S} = \{S_1,\ldots,S_K\}$ can be regarded as a mapping from each point to a label of a supervoxel, i.e.,

$$s: \{p_1,\ldots,p_N\} \rightarrow \{1,\ldots,K\}, \tag{1}$$

where $s(p)$ represents the label of the supervoxel to which the point p belongs. In addition, the supervoxel S_k is defined as a set of points whose label is equal to k:

$$S_k = \{p|s(p) = k\}. \tag{2}$$

Note that any mapping in the form of Eq. (1) can result in a partition with no more than K supervoxels. Furthermore, the total number of different possible partitioning solutions is $\frac{K^N}{K!}$ (Bergh et al., 2013), which is an extremely large number. In order to reduce the solution space, we consider a representative point $r_i \in S$ for each supervoxel S. Assuming that we have already obtained K representative points $\{r_1,\ldots,r_K\}$, the mapping function s can easily be computed according to the following equation:

$$s(p) = \arg\min_i D(p,r_i). \tag{3}$$

where D is a distance metric to measure the dissimilarity between two points. Therefore, the problem of seeking a partitioning is transformed into the problem of selecting K representative points from N original points, which is known as the subset selection problem (Elhamifar et al., 2016; Tropp, 2008). Then, the solution space is reduced from $\frac{K^N}{K!}$ to $\binom{N}{K}$. Note that because $K \ll N, \binom{N}{K} \ll \frac{K^N}{K!}$.

The subset selection problem can be encoded as an optimization problem on unknown binary variables $z_{ij} \in \{0,1\}$. Here, $z_{ij} = 1$ has two meanings: first that p_i is a representative point, and second that p_j is a non-representative point that is represented by p_i. Therefore, the definition of a supervoxel in Eq. (2) can be rewritten as: