

# Segmentation and Clustering of 3D Forest Point Cloud Using Mean Shift Algorithms

Xingbo Hu<sup>1, a</sup> and Ying Xie<sup>1, b</sup>

<sup>1</sup>School of Information Science and Technology, East China Normal University, Shanghai, China

<sup>a</sup>xbhu@ee.ecnu.edu.cn, <sup>b</sup>xiying1219@sina.com

**Keywords:** Point cloud, mean shift, forest segmentation, LiDAR.

**Abstract:** Segmenting individual trees from the forest point cloud has significant implications in forestry inventory. This paper presents a novel computational scheme to segment and cluster the 3D point cloud data acquired by an airborne LiDAR. The scheme employs a mean shift-based iterative procedure on the data sets in a defined complex multimodal feature space to cluster points with similar modes together. Experimental results reveal that the proposed scheme can work effectively and the average accuracy of tree detection (88.6%) can meet the requirements of forest inventory.

## 1. Introduction

Airborne LiDAR is an active remote sensing technology that provides distance measurements between an aircraft (or any other platform) and the surface illuminated by the laser beam. Compared with passive imaging, LiDAR has the advantage of directly measuring the three-dimensional (3D) coordinates of forest canopies [1]. Due to its ability to generate 3D point cloud data with high spatial resolution and accuracy, the airborne small-footprint LiDAR is becoming a promising technique for modeling the forest canopies and thus for completing several inventory tasks [2].

Segmenting individual trees and extracting relevant structure information from point cloud data have significant implications in forestry. Once accurately segmented, tree structural attributes such as tree height, crown diameter, basal area, diameter at breast height (DBH), wood volume, biomass, and species type can be derived [2]. Yet the existing segmentation schemes are nearly all based on the canopy height model (CHM) derived from the raw LiDAR data, which is the difference between canopy surface height and a digital elevation model (DEM) of the earth surface [1].

Mean shift is a non-parametric feature space analysis technique for locating the maxima of a density function, a so-called mode-seeking algorithm [3]. As a filtering and clustering approach, it has recently found much interest in the image processing community. However, applying the mean shift algorithm in segmentation of forest point cloud has been rarely reported. In this paper, we present a novel mean shift-based scheme dedicated to forest segmentation and clustering.

## 2. Study Area and Data

The study zone is in the center of Dinghushan National Nature Reserve, which is situated in Zhaoqing, Guangdong Province, China. It is located in 112°30'39" - 112°33'41"E, 23°09'21" - 23°11'30"N and covers a total area of 1155 ha with the altitude of 14.1 - 1000.3 m. The site is covered by a low subtropical evergreen broadleaved forest more than 400 year old [4].

The full-waveform ALS (Airborne Laser Scanning) data along with high-resolution CCD image data were acquired through a flight experiment conducted on the Dinghushan study site. The point cloud data directly generated by the system are delivered in the LiDAR standard file exchange format – Log ASCII Standard (LAS), which can record three-dimensional coordinates, intensity values and the pulse return numbers for all discrete target points.

### 3. Methodology

#### 3.1 Mean Shift Theory

The mean shift is a recursive computing procedure using a non-parametric probability density estimator based on the Parzen window kernel function [5]. It can estimate the stationary points (modes) of the underlying probability density function without explicitly estimating the density. Given  $n$  data points  $\mathbf{x}_j$  ( $j = 1, \dots, n$ ) in the  $d$ -dimensional space  $R^d$ , the kernel estimator for the probability density function  $f(\mathbf{x})$  at point  $\mathbf{x}$  can be written as [5]

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{j=1}^n k\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|\right), \quad (1)$$

where,  $c_{k,d}$  is the normalization constant making  $K(\mathbf{x})$  (the  $d$ -variate kernel function) integrate to one,  $h$  is the kernel bandwidth and  $k(\cdot)$  is called the profile of the kernel that describes how strongly the data points are taken into account in the estimation.

The mean shift tries to determine the local modes of the density function  $f(\mathbf{x})$ , which correspond to the zeros of the gradient  $\nabla f(\mathbf{x}) = 0$ . According to [5], the density gradient estimator has the following mathematical expression:

$$\hat{\nabla} f_{h,K}(\mathbf{x}) = 2\hat{f}_{h,G}(\mathbf{x}) \cdot \mathbf{m}_{h,G}(\mathbf{x}) / h^2 c, \quad (2)$$

where  $c = c_{g,d} / c_{k,d}$  and  $G$  is the kernel profile defined as  $g(x) = -k'(x)$ . In (2), the first term is the density estimate at point  $\mathbf{x}$  computed with kernel  $G$ :

$$\hat{f}_{h,G}(\mathbf{x}) = \frac{c_{g,d}}{nh^d} \sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|\right), \quad (3)$$

and the second term is the so-called mean shift vector:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{j=1}^n \mathbf{x}_j g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|\right)}{\sum_{j=1}^n g\left(\left\|\frac{\mathbf{x} - \mathbf{x}_j}{h}\right\|\right)} - \mathbf{x}, \quad (4)$$

which is the difference between the weighted mean, using the kernel  $G$  for weights, and  $\mathbf{x}$ , the center of the kernel.

Rewriting (2) can yield

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{1}{2} h^2 c \frac{\hat{\nabla} f_{h,K}(\mathbf{x})}{\hat{f}_{h,K}(\mathbf{x})}. \quad (5)$$

This equation shows that the mean shift vector at point  $\mathbf{x}$  computed with kernel  $G$  is proportional to the normalized density gradient estimate obtained with kernel  $K$ . The mean shift vector thus always points towards the direction of the maximum increase in the probability density. By setting

$$\mathbf{x}(t+1) \leftarrow \mathbf{x}(t) + \mathbf{m}_{h,G}[\mathbf{x}(t)], \quad (6)$$

where  $t$  denotes the iteration number, the iterative process converge towards the local maxima. Thus it can be seen that the mean shift procedure does not need to evaluate the density function  $\hat{f}_{h,K}(\mathbf{x})$  itself but only the induced kernel  $g(x)$ .

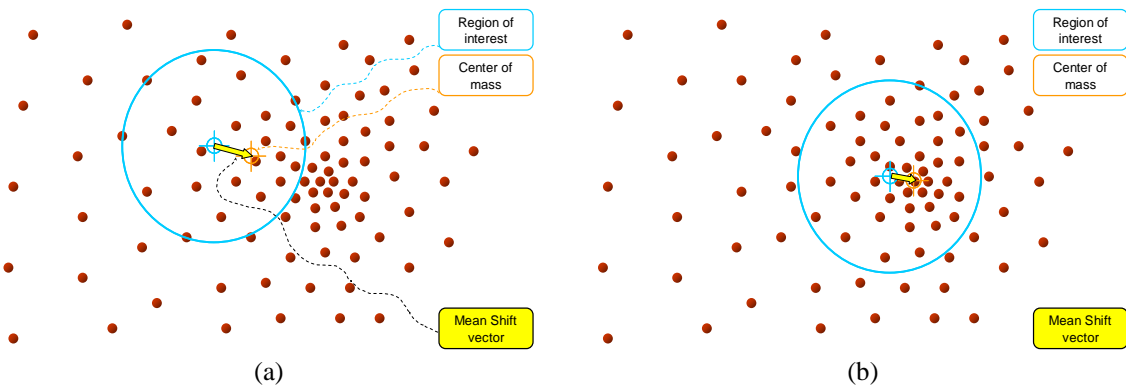


Fig. 1. Intuitive description of the mean shift procedure. (a) Distribution of identical billiard balls. (b) Find the densest region.

Mean shift is essentially a hill climbing algorithm which involves shifting a kernel iteratively to a higher density region until convergence. The relation captured in (5) is intuitive and the local mean is shifted toward the region in which the majority of the points reside, as depicted in Fig. 1.

### 3.2 Mean Shift-Based Segmentation

The mean shift analysis is performed on the well-prepared data sets in the defined complex multimodal feature space to cluster points with similar modes (in the forestry scenario, the term “mode” has close relations with each single tree’s structure and shape) together and delineate individual trees. As the Gaussian function has similar characteristics with the forestry point cloud in statistical distributions and peak shapes, we choose the truncated, multivariate Gaussian function as the kernel (or window) function. Employing the joint domain representation in this study, the multivariate Gaussian kernel can be defined as the product of three radially symmetric kernels and the Euclidean metric allows a single kernel bandwidth parameter for each domain

$$G_N(\mathbf{x}) = \frac{C}{h_s^3 h_n^3 h_e} g\left(\left\|\frac{\mathbf{x}^s}{h_s}\right\|^2\right) g\left(\left\|\frac{\mathbf{x}^n}{h_n}\right\|^2\right) g\left(\left|\frac{x^e}{h_e}\right|^2\right), \quad (7)$$

where,  $C$  is the corresponding normalization constant,  $\mathbf{x}_s$  is the spatial feature vector,  $\mathbf{x}_n$  is the shape feature vector,  $x_e$  is the energy component,  $h_s$ ,  $h_n$  and  $h_e$  are the employed kernel bandwidths for the three domains, respectively, and  $k(\cdot)$  is the common profile of the induced kernel  $G(\mathbf{x})$  used in all three domains. So the mean shift vector defined in (4) can be rewritten as:

$$\mathbf{m}_{h,G}(\mathbf{x}) = \frac{\sum_{j=1}^n \mathbf{x}_j g\left(\left\|\frac{\mathbf{x}^s - \mathbf{x}_j^s}{h_s}\right\|^2\right) g\left(\left\|\frac{\mathbf{x}^n - \mathbf{x}_j^n}{h_n}\right\|^2\right) g\left(\left|\frac{x^e - x_j^e}{h_e}\right|^2\right)}{\sum_{j=1}^n g\left(\left\|\frac{\mathbf{x}^s - \mathbf{x}_j^s}{h_s}\right\|^2\right) g\left(\left\|\frac{\mathbf{x}^n - \mathbf{x}_j^n}{h_n}\right\|^2\right) g\left(\left|\frac{x^e - x_j^e}{h_e}\right|^2\right)} - \mathbf{x}. \quad (8)$$

Let  $\mathbf{X} = \{\mathbf{x}_i\}$  ( $i = 1, \dots, N$ ) be the input data set of the multi-dimensional (7D in this study) feature space, which is transferred from the original LiDAR point cloud. Run the mean shift computation on the data set  $\mathbf{X}$  for each data point  $\mathbf{x}_i$  and each iterative process will converge towards their corresponding stationary points or modes. Let  $\mathbf{Y} = \{\mathbf{y}_i\}$  ( $i = 1, \dots, N$ ) denote the data set of mean shift points, i.e. values at point  $\mathbf{x}_i$  after its mean shift iteration. The detailed mean shift procedure proposed for the individual tree segmentation is described as follows:

1). Specify the spatial-domain kernel bandwidth  $h_s$ , the geometric-domain kernel bandwidth  $h_n$  and the energy-domain kernel bandwidth  $h_e$ . Specify a value for the allowable error -  $\varepsilon$ . Kernels are also associated with data points in the multimodal feature space.

2). For each point  $\mathbf{x}_i$ , execute the following computations:

– Initialization: associate a mean shift point  $\mathbf{y}_i$  with the point  $\mathbf{x}_i$ , and initialize it to coincide with that point.

Let  $t = 0$  and  $\mathbf{y}_i(0) = \mathbf{x}_i$ , where  $t$  denotes the current number of mean shift iterations.

– For  $0 \leq t \leq c$  ( $c$  represents the word “convergence”), repeat the following two steps until convergence.

– Calculate the mean shift vector:

$$\mathbf{m}_{h,G}[\mathbf{y}_i(t)] = \frac{\sum_{j=1}^n \mathbf{y}_j(t) g\left(\left\|\frac{\mathbf{y}_i^s(t) - \mathbf{x}_j^s}{h_s}\right\|^2\right) g\left(\left\|\frac{\mathbf{y}_i^n(t) - \mathbf{x}_j^n}{h_n}\right\|^2\right) g\left(\left|\frac{y_i^e(t) - x_j^e}{h_e}\right|^2\right)}{\sum_{j=1}^n g\left(\left\|\frac{\mathbf{y}_i^s(t) - \mathbf{x}_j^s}{h_s}\right\|^2\right) g\left(\left\|\frac{\mathbf{y}_i^n(t) - \mathbf{x}_j^n}{h_n}\right\|^2\right) g\left(\left|\frac{y_i^e(t) - x_j^e}{h_e}\right|^2\right)} - \mathbf{y}_i(t). \quad (9)$$

– Update the mean shift point:

$$\mathbf{y}_i(t+1) = \mathbf{y}_i(t) + \mathbf{m}_{h,G}[\mathbf{y}_i(t)]. \quad (10)$$

stopping when  $\mathbf{m}_{h,G}[\mathbf{y}_i(t)]$  is less than a specified value of  $\varepsilon$ .

– Set  $\mathbf{y}_i = \mathbf{y}_i(c)$ .

3). Merge all modes whose mutual distance is less than the kernel bandwidth, i.e., closer than  $h_s$  in the spatial domain,  $h_n$  in the geometric domain and  $h_e$  in the energy domain, by concatenating the basins of attraction of the corresponding stationary points, to produce the clusters  $\{\mathbf{C}_p\}$  ( $p = 1, \dots, M$ ).

#### 4. Results and Analysis

The mathematical processing flow is performed on the MATLAB platform by calling a C++ wrapper which implements a mean shift-based iterative procedure. Experimental results are presented to validate the effectiveness of the proposed mean shift-based segmentation scheme. Fig. 2 shows the two different views (from top and side, respectively) of segmented point cloud after the first round mean shift-based iterative computations, with spatially adjacent clusters rendered in different colors.

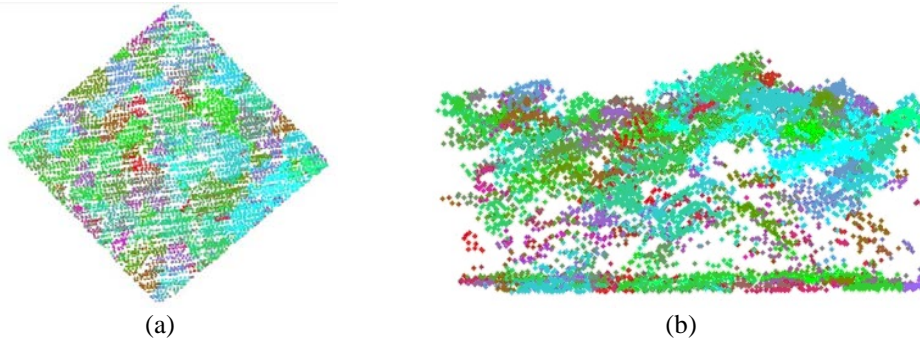


Fig. 2. Segmented point cloud after the mean shift procedure. (a) Top view. (b) Side view.

To assess the accuracy of the proposed segmentation scheme, we calculate the tree detection rates for three typical sample plots. From the data listed in Table 1, we can see that the average accuracy of this scheme would reach 88.6% and could meet the requirements of forest survey.

Table 1. Tree Detection Rates in Three Test Plots.

Plot No.	Actual Tree Numbers	Detected Tree Numbers	Detection Rate
502	124	111	89.5%
605	104	115	90.4%
1802	113	97	85.8%

#### 5. Conclusion

This paper presents a novel computational scheme to segment and cluster the 3D forest point cloud data. The scheme employs a mean shift-based recursive computing procedure on the data sets in a defined complex multimodal feature space to cluster points with similar modes together. Experimental results show that the proposed scheme can work effectively and the accuracy could meet the requirements of forest inventory.

#### Acknowledgements

This work was financially supported by the Special Scientific Research Fund of Public Welfare Profession of China (Grant No. 201313009).

#### References

- [1] Q. Chen, D. Baldocchi, P. Gong, and M. Kelly, "Isolating individual trees in a savanna woodland using small footprint lidar data," *Photogrammetric Engineering & Remote Sensing*, vol. 72, no 8, pp. 923-932, Aug. 2006.
- [2] W. Li, Q. Guo, M. K. Jakubowski, and M. Kelly, "A new method for segmenting individual trees from the lidar point cloud," *Photogrammetric Engineering & Remote Sensing*, vol. 78, no. 1, pp. 75-84, Jan. 2012.
- [3] Y. Cheng, "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Anal. & Mach. Intel.*, vol. 17, no. 8, pp. 790 - 799, Aug. 1995.

- [4] Y. Shen, L. S. Santiago, L. Ma, G. Lin, J. Lian, H. Cao, and W. Ye, "Forest dynamics of a subtropical monsoon forest in Dinghushan, China: recruitment, mortality and the pace of community change," *J. Trop. Ecol.*, vol. 29, no. 2, pp. 131-145, Mar. 2013.
- [5] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603-619, May 2002.