Abstract: Timely and accurate information about spatial distribution of tree species in urban areas provides crucial data for sustainable urban development, management and planning. Very high spatial resolution data collected by sensors onboard Unmanned Aerial Vehicles (UAV) systems provide rich data sources for mapping tree species. This paper proposes a method of tree species mapping from UAV images over urban areas using similarity in tree-crown object histograms and a simple thresholding method. Tree-crown objects are first extracted and used as processing units in subsequent steps. Tree-crown object histograms of multiple features, i.e., spectral and height related features, are generated to quantify within-object variability. A specific tree species is extracted by comparing similarity in histogram between a target tree-crown object and reference objects. The proposed method is evaluated in mapping four different tree species using UAV multispectral ortho-images and derived Digital Surface Model (DSM) data collected in Shanghai urban area, by comparing with an existing method. The results demonstrate that the proposed method outperforms the comparative method for all four tree species, with improvements of 0.61–5.81% in overall accuracy. The proposed method provides a simple and effective way of mapping tree species over urban area.

Keywords: UAV; DSM; object histogram; VBSD; tree species; urban area

1. Introduction

Urban tree cover plays an important role in sustainable urban development and planning by providing a range of environmental and ecological services, and social and economic benefits [1]. For example, urban trees absorb carbon dioxide, improve air quality, mitigate urban heat island effect, reduce urban flood risk, embellish urban environments and provide recreational spaces [2–5]. The diversity, structure and spatial distribution of tree species are closely related to the quality of these services and benefits. For example, some tree species are suitable for wildlife, or tolerate water logging, or have ornamental characters [6,7]. Therefore, accurate and timely information about urban tree species and their spatial distributions is critical for supporting strategies of urban development and for planting and maintaining city greening.

Remote sensing data and derived data provide useful data sources for mapping tree species [8]. In particular, high spatial resolution data with fine spatial details are more suitable for identification at species level [9]. Many existing studies used different image data, such as satellite and aerial multispectral images [10,11], airborne hyperspectral images [12,13] of very high spatial resolution (VHR), and three-dimensional data from airborne LiDAR (Light Detection and Ranging) data [14–17]. VHR multispectral images, such as IKONOS, QuickBird and WorldView-2, and hyperspectral images have been widely used in distinguishing different tree species [10–13,18,19]. However, due to spectral
similarity of tree species, additional features are often used, such as texture and shape features derived from multispectral images [10] and red edge bands and vegetation indices from hyperspectral images [13,19]. In addition, hyperspectral data and LiDAR data were jointly used to extract different tree species [14–17], providing promising results.

Recently, Unmanned Aerial Vehicles (UAV), as a low-altitude airborne platform, have been widely used to collect VHR images [20]. Three-dimensional (3D) data are generated from the overlapped UAV images collected, i.e., photogrammetric point clouds and Digital Surface Model (DSM). These 3D data generated, in turn, are also used to produce UAV ortho-images [21]. The acquisition of VHR UAV images also provides an opportunity for mapping individual tree species.

For identification of tree species from UAV images, different image features were used to quantify characteristics of different tree species. Franklin et al. identified four tree species from UAV images acquired by a Mini-MCA6 multispectral camera using visible and near-infrared bands, texture features from Gray-Level Co-occurrence Matrix (GLCM), and object shape features [22]. Lisein et al. used phenological differences between species from multi-temporal UAV images collected by visible and color infrared cameras to improve accuracy of tree species mapping [23]. Similarly, in other studies, spectral features, vegetation indices and textures from UAV imagery were used to identify invasive plant species [24–26].

In some studies of tree species mapping from UAV images, photogrammetric point clouds and DSM generated from overlapped UAV images were used to delineate the structure of tree crowns and the height related information of tree species. For example, Saarinen et al. used height related features, such as maximum value, standard deviation, coefficient of variation and percentiles of heights, from 3D point clouds acquired by a UAV-based visible camera, and spectral features from a hyperspectral frame camera to monitor biodiversity of various tree species [27]. Cao et al. separated different mangrove species using a combination of UAV hyperspectral image collected by a UHD 185 hyperspectral sensor and height data from UAV-derived DSM [28]. Liu et al. used UAV multispectral ortho-image and topographical features from DSM, such as slope and aspect, to detect an endangered tree species, *Firmiana danxiaensis* [29]. It is found from these studies that spectral and height data provide supplementary information in tree species mapping.

From the literature survey, an object-based method has been widely used in mapping tree species in most existing studies [22,23,28,29]. The commonly used features of image objects are the mean and standard deviation of pixel values within an image object. However, due to significant heterogeneity of image objects in VHR UAV images, these features may not well quantify characteristics of image objects [30–33]. Therefore, different ways of expressing the characteristics of image objects are needed to exploit for tree species identification.

In addition, in most existing studies of tree species mapping, the multi-class classification methods have been widely used, such as Nearest Neighbor (NN) [26], Support Vector Machine (SVM) [29], and Random Forest (RF) [22–25] classifiers, to identify tree species. Although these multi-class classifiers provide promising results, there are two problems with multi-class classifiers. One is that several training samples are required by these multi-class classifiers. The other is that samples of non-target classes (other tree species and vegetation) should be selected in classification training. Therefore, other methods to overcome these problems are worth exploiting to extract tree species.

It should be mentioned that most existing studies of tree species mapping using VHR images focus on forest or rural areas [18,19,22,23]. There are few studies using VHR images in mapping tree species in urban areas [10,11]. It is difficult to identify tree species in urban areas using VHR images, since the trees are generally scattered and different tree species are mixed in urban areas.

Considering the aforementioned problems, in this paper, we propose a novel method of tree species mapping from UAV multispectral images and derived height data over urban area using tree-crown object histogram and a simple thresholding method.
2. Study Area and Data

2.1. Study Area and Tree Species

In this study, a small area of Tongji University campus in Shanghai urban area (31°19′N, 121°30′E), China (Figure 1a) was selected as study area, which is very flat. The main land cover types in the area include building, road, water, tree and grass. Different types of trees species disperse in this dense urban area and several different tree species are generally mixed.

Three major types of tree species were identified in the area, i.e., Metasequoia, Platanus and Camphora. Metasequoia is a unique plant in China and has a certain resistance to sulfur dioxide. It is also an important tree species of timber forest, shelter forest, urban greening and landscape forest [34]. This tree species has tall, straight trunks and crown forms of circular cone. The crown diameters of Metasequoia are much smaller than those of other species in the study area. Platanus is one of the world-famous roadside trees with dense branches and leaves. The trunk diameters of Platanus vary greatly and large amounts of fruit balls are produced when the trunk diameter reaches about 30 cm. The seminal hair and pollen from fruit balls floating in the air from April to June may cause air pollution problems and allergic reactions to some people [35]. Given that Platanus trees have significant difference in trunk diameter, crown width and crown form, Platanus is divided into two sub-types in this study, i.e., Platanus I and Platanus II. The tree crowns of Platanus II are wide and thick, whereas Platanus I has sparse tree crowns and is of relatively low height. Camphora is an evergreen ornamental tree species sending out fragrance, which can drive mosquitoes away. It also has a strong resistance to carbon dioxide, chlorine gas and some other toxic gases [36]. The crown forms of Platanus and Camphora are usually of similar shapes as egg or ball.
Tree crowns of these two species are mostly wide, providing shade for people. Therefore, crowns of trees planted in rows are generally connected to each other.

These four tree species mentioned above are planted regularly, and are all distributed in rows along the northwest or northeast directions. The tree species of Metasequoia and Camphora are mainly planted around the buildings. The tree species of Platanus I and Platanus II are mainly planted on roadsides. Except for these four tree species mentioned above, there are a few other tree species, which are scattered in the area and are not target tree species of this study.

2.2. Data

UAV multispectral images and derived DSM data were used in this study. The UAV images were collected using a customized multispectral imaging system [29], in September, 2016, when the weather was fine and the trees flourished in the study area. The spatial resolution of the UAV multispectral images is about 4 cm. Red, Green, Blue and near-infrared bands were used in this study.

The UAV multispectral images collected were processed to generate ortho-image and DSM image. First, optical and radiometric calibrations were carried out on the UAV images using the methods presented in Reference [29]. The optical calibration corrected the geometric distortion and removed the vignetting effect, while the radiometric calibration transformed the digital numbers to radiance. Point clouds were computed from these calibrated stereo UAV images. Then ortho-images were produced and were mosaicked to a stacked image over study area. DSM was generated from the derived point clouds using the inverse distance weighted interpolation [37]. These processing steps were implemented using Pix4D mapper software. The multispectral ortho-images and the DSM images generated were from the same source with geometric consistency.

The image size of UAV multispectral ortho-image and the derived DSM image used in this study is 6000 × 11,000 pixels (Figure 1b,c). It should be noted that the DSM is used to represent the relative height in this study, since the terrain of the study area is very flat.

3. Methods

In this study, a novel method of tree species mapping from UAV images over urban areas using object-level histogram of multiple features was proposed. Image objects were first generated from image segmentation. Tree-crown objects were then generated from the initial image objects obtained and used as processing units in subsequent steps. Instead of using mean or standard deviation of pixel features in image object as in the conventional object-based method, object histograms were used to quantify distributions of different features in tree-crown object. Multiple features, i.e., the spectral features and height related features derived from UAV images, were used to quantify histograms of tree-crown objects. A specific tree species was extracted by quantitatively comparing similarity in histograms between a target tree-crown object and reference objects, measured using the Variable Bin Size Distance (VBSD) [38], a recently proposed histogram similarity measure. Specifically, four main steps are included, namely, extraction of tree-crown objects, generation of object histogram, comparison of histogram similarity and urban tree species mapping (Figure 2). These steps are described in detail in the following sub-sections.
3.1. Tree-Crown Object Extraction

For mapping tree species from UAV images over urban area, tree-crown objects were first extracted in this study. Five steps were implemented to generate tree-crown objects (Figure 3). In each step, different features were used, which were selected for different purposes.

Image segmentation was first done using UAV multispectral ortho-image and DSM image to produce initial image objects. Image segmentation is a common method of generating homogeneous and disjoint image objects [39,40]. A widely used multiresolution segmentation method, implemented in eCognition Developer software (Trimble) [41,42], was used for image segmentation. The method is a region-based segmentation method. It starts with each pixel forming one image object. At each step, neighboring image objects are merged into one larger object. The merging decision is based on local homogeneity criteria, describing the similarity of adjacent image objects. As scale parameter increases, different levels of segmentation are generated. According to both spectral and height homogeneity criteria, the image was segmented into lots of image objects, under the premise of avoiding under-segmentation.

After initial segmentation, potential tree-crown objects were generated from these image objects obtained. Vegetation index (e.g., Normalized Difference Vegetation Index, NDVI) and height features of image objects were used in this step. Specifically, if the mean NDVI value of an image object was higher than the specified NDVI threshold and its mean height value was also higher than the height threshold, the image object was identified as a potential tree-crown object. Otherwise, the image object was identified as a non-tree-crown object and was masked out.
threshold value determined, the image object was identified as potential tree-crown object. Otherwise, the image object was identified as non-tree-crown object and was masked out.

To make the shape of potential tree-crown objects generated in previous step more complete, in the third step, adjacent and homogeneous potential tree-crown objects were merged into one tree-crown object. Specifically, if both spectral and height homogeneity values for two neighboring tree-crown objects were greater than the specified threshold values, these two neighboring tree-crown objects were merged into a larger and more complete tree-crown object, by using a larger scale of segmentation.

It is found that these relatively complete tree-crown objects extracted may still contain some non-tree-crown objects. For example, bright patches on roofs or shadow objects around the trees were wrongly included. Therefore, these misidentified objects should be removed. In the fourth step, Brightness, which is the mean value of all the spectral bands [43], was used as an object feature to remove these misidentified objects. Given that Brightness of bright patches on roofs was very high and Brightness of shadow objects around the trees was very low, while Brightness of tree crowns were between them, two threshold values were determined, i.e., a high threshold value and a low threshold value. Those objects with Brightness values greater than the high threshold or less than the low threshold were labeled as misidentified objects and were eliminated.

Although non-tree-crown objects were masked out, the shapes of tree-crown objects were not complete enough. A merging step was further implemented with a larger scale of segmentation than that used in third step. Spectral features were not used as homogeneity criteria, because of the existence of illumination variation in UAV images. Instead, height was used as homogeneity criterion. Neighboring tree-crown objects were merged into more complete ones when their height homogeneity value was greater than the threshold value.

![Figure 3. The process of tree-crown object extraction and features used in each step.](image)

### 3.2. Generation of Tree-Crown Object Histogram

The tree-crown objects extracted from UAV images in the previous step show significant internal heterogeneity. Object histogram, which represents frequency distribution of pixel features within a tree-crown object, was used to quantify the within-object heterogeneity [32,33]. The object histogram is generated by grouping all the feature values within an object into different intervals and counting occurrence of different intervals.

Different features were exploited to quantify object histograms. To select appropriate features used for object histogram, various features, including spectral features, texture, vegetation index, height, slope and aspect, were compared and analyzed for different tree species. After comparison, it was found that spectral feature, height and slope of tree-crown objects are more discriminative features. Therefore, these three features were selected to constitute multiple features for quantifying characteristics of tree-crown objects in this study.
To make different features have similar value ranges of occurrence, the percentage frequency of occurrence of each bin size was generated. The bin sizes of object histograms were selected according to the overall ranges of spectral features or height related features. Fine bin size was selected to better reflect object features. After that, object histograms of multiple different features were combined to represent the characteristics of different tree species.

Considering that object histograms generally show noise or undulation, which impedes feature combination, histogram smoothing was adopted. Specifically, Savitzky–Golay filter (S–G filter) \[44\] was adopted to smooth object histograms. The filter is based on local polynomial least square fitting, by which the noise is removed while the shape and width of the signal are preserved \[45\]. It should be noted that other filters could also be used for smoothing.

### 3.3. Histogram Similarity Comparison

After multiple feature histograms of tree-crown objects were generated, the histogram of a tree-crown object was compared with those of reference tree-crown objects to extract a specific tree species. Therefore, an appropriate histogram similarity measure should be selected. In this study, VBSD \[38\], was used to measure the similarity between tree-crown object histograms.

VBSD is based on bin-by-bin distances and achieves the effect of cross-bin distances by varying the bin size from fine scale to coarse scale \[38\]. Therefore, it could be considered as a cross-bin extension of corresponding bin-by-bin distances, e.g., the VBSD for \(L_1\) distance.

The basic principle of VBSD can be summarized as follows \[38\]. First, bin-by-bin distance for the most refined bin size is computed as the first sub-distance \(d_1\). After that, the intersection part of two histograms is subtracted from each histogram. Second, the remained two histograms are changed into coarser histograms with an increased bin size. The bin-by-bin distance is computed again for current bin size to generate the second sub-distance \(d_2\), and so does the intersection part subtracted. This process continues until the bin size is large enough. Finally, the fine-to-coarse bin-by-bin sub-distances \(D = \{d_1, d_2, \ldots, d_{t_{\text{max}}}\}\), where \(t_{\text{max}}\) denotes the largest bin size, are obtained and a weighting function is used for all these sub-distances to obtain a summation. This summation, i.e., VBSD, is used to represent the similarity between these two histograms. The VBSD is expressed as

\[
VBSD = f(D) = \sum_{t=1}^{t_{\text{max}}} (w_t \cdot d_t),
\]

where \(w_t\) is the weight of \(d_t\). In this study, \(w_t = 1/t_{\text{max}}\), the mean value of all these distances, is used to calculate VBSD. The smaller VBSD represents that the target histogram is more similar with the reference histogram. The major advantage of VBSD is that it is insensitive to both the histogram translation and the variation of histogram bin size \[38\].

In this study, three bin-by-bin distances were used, \(L_1\) distance, \(L_2\) distance and \(\chi^2\) statistic distance. Suppose that \(R\) denotes the reference object histogram and \(T\) is the histogram of a target tree-crown object with \(n\) bins. These three distances are expressed as

\[
D_{L_1}(R, T) = \sum_{i=1}^{n} |r_i - t_i|,
\]

\[
D_{L_2}(R, T) = \sqrt{\sum_{i=1}^{n} |r_i - t_i|^2},
\]

\[
D_{\chi^2}(R, T) = \sum_{i=1}^{n} \frac{|r_i - t_i|^2}{2(r_i + t_i)}.
\]

### 3.4. Tree Species Mapping

As mentioned previously, VBSD was used as a histogram similarity measure in mapping tree species. Specifically, to determine if a target tree-crown object is a specific tree species, VBSD between the
target tree-crown object and reference object is computed and compared with a specified threshold value. If the VBSD obtained is less than the selected threshold, the tree-crown object is identified as target class. Otherwise, the target object is identified as non-target class. Tree-crown objects with smaller VBSDs with the reference object (i.e., higher similarity) have greater probability of being the same tree.

For histogram similarity comparison, it is important to obtain representative reference object histograms. Reference samples (tree-crown objects) were selected by visual interpretation and field investigation in this study. Since most tree species usually show significant variability in multiple features, such as the spectral and height differences in tree crowns and different planting patterns of trees (i.e., tree crowns alone or in clusters), multiple reference samples reflecting the intra-class variability of tree species were selected. A reference object histogram used in this study was expressed by the average histogram of two similar samples. After that, all the reference object histograms were used to compute VBSDs with a target object histogram. Therefore, multiple VBSDs each for a reference object histogram were obtained for the target tree-crown object. The minimum of these VBSDs was considered as the final VBSD for the target object.

To determine an appropriate VBSD threshold for extracting a specific tree species, a threshold range was first determined by analyzing the distribution (histogram) of VBSDs image obtained. A bimodal model of VBSDs histogram [46] was assumed in this study. A VBSD threshold was then selected from the threshold range obtained. Specifically, according to VBSD threshold range, the optimal threshold was determined near to the intersection of two peaks of VBSDs histogram by trial and error. After that, tree-crown objects with VBSDs smaller than the threshold for a specific tree species were labeled as this species.

3.5. Accuracy Assessment

The confusion matrix was used to assess accuracy of the proposed method. The producer’s accuracy (PA), user’s accuracy (UA), overall accuracy (OA), Kappa Coefficient (Kappa) and F1 score, computed from confusion matrix, were used as the accuracy measures. The PA, UA, OA and Kappa are commonly used measures in remote sensing [47]. The F1 score is the harmonic mean of precision and sensitivity and is usually used as an accuracy measure of a dichotomous model [48], which is suitable for one-class classification method.

For a complete comparative analysis, VBSDs for three different bin-by-bin distances were adopted in the proposed method and accuracies using these three VBSDs were compared. Moreover, a conventional object-based method for tree species mapping was used for comparison. In this comparative method, tree-crown objects were first generated using the method proposed in this study. Spectral and height related features for each tree-crown objects were used, i.e., four spectral features, height and slope features were used. The mean values of all six features within each tree-crown object were used in extraction process. A recently developed one-class classifier, One-class Support Vector Machine (OCSVM) [49], was used in object-based classification. It should be noted that more training samples of the target class are needed in training of OCSVM than the reference samples of the proposed method.

To fully evaluate the proposed method, five repeated tests were implemented with different validation samples on the premise of the same quantity and proportion of the training samples and the validation samples. In each test, the same validation samples were used for different methods. The validation samples include both samples of the target tree species and samples of other non-target tree species. Because the target tree species were a small portion of all trees, the non-target class samples were randomly selected as twice of the target class samples. The average and standard deviation for five tests were computed for each mapping method.
4. Results

4.1. Generation of Tree-Crown Objects

As described in Section 3.1, tree-crown objects were first generated. The threshold values used in each step are shown in Table 1. A total of 888 tree-crown objects were generated and the other objects were masked out.

Table 1. The threshold values used in each step for generation of tree-crown object.

| Steps | Parameters and Threshold Values |
|-------|---------------------------------|
| 1     | Initial segmentation scale: 50  |
| 2     | NDVI: −0.4; Height: −24        |
| 3     | Merging scale: 120             |
| 4     | Brightness: 57 (low) and 180 (high) |
| 5     | Merging scale: 140             |

Figure 4 shows the tree-crown objects extraction results at each step. From the figure, the initial image segmentation shows significant over-segmentation (Figure 4a), because small scale of segmentation is selected to avoid under-segmentation. After second step, most of potential tree-crown objects are extracted, while non-tree areas, such as buildings, roads and grasses, are masked out (Figure 4b). After merging was done (third step), the potential tree-crown objects obtained became relatively complete. However, it is worth noting that some non-tree objects are wrongly identified as tree-crown objects (Figure 4c). After fourth step, most of these misidentified objects are eliminated and tree-crown objects are more accurately extracted, but these tree-crown objects extracted are not complete enough (Figure 4d). After merging in final step, complete tree-crown objects are generated and the boundaries between tree-crown objects obtained are more accurate (Figure 4e).

Figure 4. Portions of tree-crown object extraction results at each step (a–e) and final result of study area (f): The initial image objects with white boundaries (a); The potential tree-crown objects shown in white are extracted and the others are masked out (b); The merging objects with white boundaries (c); The misidentified objects shown in blue are eliminated (d); The merging objects with white boundaries (e). The extracted tree-crown objects in the study area colored in green (f).
The final result of tree-crown object extraction is shown in Figure 4f. In general, all tree-crown objects are accurately and completely extracted. It is also found from the figure that the isolated trees extracted are more complete than the trees in clusters. A part of a tree or a combination of several trees are extracted as one tree-crown object when the boundaries between these trees in clusters are not clearly discernable.

4.2. Object Histogram Analysis

Object histograms of multiple features were generated to quantify characteristics of different tree species. In this study, six features were used, namely four spectral features (blue, green, red and near-infrared) and two height related features (height and slope). The spectral and height features were directly derived from four multispectral bands and DSM from UAV images, respectively. The slope feature was obtained by a moving window of 25 by 25 (i.e., 1 m by 1 m) acting on DSM to compute the slope value at each pixel position. The kernel sizes of S–G filter for smoothing were 5 for the spectral and slope features and 10 for the height feature.

Figure 5 shows multiple feature histograms of selected reference objects of four tree species. There are different variations in peaks and shapes of these reference object histograms. Generally, four target tree species show different object histograms. From Figure 5, the object histograms of Metasequoia are the most different from other tree species in most features, i.e., the height, slope, green, red and near-infrared features (Figure 5a). The object histograms of Platanus II also show different characteristics, in particular in height, slope, green and red features (Figure 5c). The object histograms of Platanus I (Figure 5b) and Camphora (Figure 5d) are generally similar, but with subtle differences in green, red and height features.

It is also found from Figure 5 that histograms of four tree species in green and red bands are more different. For example, histograms of Metasequoia have one or two peaks, while the histograms of Camphora only have one obvious peak and the right sides of the histograms become flat gradually. The histograms of Platanus I in green and red bands rapidly increase to peak and gradually decline to the end, but the changing rates of histograms of Platanus II are different. As for histograms in near-infrared band, differences between tree species are evident in shapes and trends of the histograms. The value range of height for Metasequoia is the widest and the value range of height for Platanus I is the narrowest. A narrow value range reflects a relatively uniform height distribution. The value range of slope for Metasequoia, which mainly concentrates on large values, is obviously different from that of other tree species.

It is also worth noting that there are similarities in different features among four tree species. For example, histograms in blue band of Metasequoia and Platanus II are slightly similar and the histograms in blue band of Platanus I and Camphora are also slightly similar. The histograms of slope for Platanus I, Platanus II and Camphora only have some subtle differences, since they have similar crown shapes of egg or ball. In summary, four types of tree species show generally different tree-crown object histogram characteristics.
Figure 5. The reference histograms of tree-crown objects for four tree species colored in different colors. Multiple features are labeled from left to right: spectral features, i.e., blue, green, red and near-infrared features, height feature and slope feature. The numbers of reference object histograms are 5, 6, 7, 5, respectively, for *Metasequoia* (a), *Platanus I* (b), *Platanus II* (c) and *Camphora* (d). Their corresponding UAV images are circled in red on the right.

4.3. Tree Species Mapping Results

Following the method mentioned in the previous section, the threshold ranges were all determined by reference to VBSDs histograms (Figure 6). To better understand their distributions, the VBSDs histograms are smoothed and shown by solid lines in Figure 6. It is found from the figure that most VBSDs histograms have obvious separations between two peaks of histogram, i.e., showing bi-mode patterns, except those shown in Figure 6f, h, k, i. For example, the threshold range of VBSDs for $\chi^2$ distance for *Metasequoia* is roughly near 1.0, which is located between two peaks (Figure 6c). However, when the VBSDs histogram is hard to separate as two obvious peaks, more uncertain and wider range of threshold is selected (Figure 6f). In addition, the histograms and ranges of VBSDs for three distances are different. It is found that the distributions of VBSDs for $L_1$ distance (Figure 6a, d, g, j) are similar to those of VBSDs for $\chi^2$ distance (Figure 6c, f, i, l). The threshold range obtained provides a guidance for VBSD threshold selection. The final thresholds for four tree species used are shown in Table 2.
Figure 6. The distributions of VBSDs images for three distances and four tree species: Metasequoia (a–c); Platanus I (d–f); Platanus II (h–i); Camphora (j–l). The smoothed histograms are shown by solid lines. The ranges of the threshold are labeled in red.

Table 2. The VBSD thresholds for four tree species in the proposed method.

| Tree Species | VBSD for $L_1$ | VBSD for $L_2$ | VBSD for $\chi^2$ |
|--------------|----------------|----------------|-------------------|
| Metasequoia  | 2.23           | 0.40           | 1.02              |
| Platanus I   | 1.73           | 0.35           | 0.77              |
| Platanus II  | 1.81           | 0.37           | 0.82              |
| Camphora     | 1.78           | 0.31           | 0.78              |

The validation samples included 42, 82, 89 and 77 tree-crown objects for four tree species all over the study area, respectively. The samples of non-target class were twice as these target class samples to guarantee a consistent sample proportion.

The accuracies of mapping results of the proposed method using VBSDs for three distances are listed in Table 3. From the table, the OA, Kappa and F1 score mostly have a consistent trend. Specifically, the proposed method performs differently in mapping different tree species (OAs are from 95.75–82.65%), but is mostly similar when using VBSDs for three different distances in each species mapping. The VBSD for $\chi^2$ distance performs the best for tree species of Metasequoia, Platanus I and Platanus II. The OAs from VBSD for $\chi^2$ distance for these three tree species are 0.34%, 1.86% and 3.77% higher than those from VBSD for $L_2$ distance, respectively. The VBSD for $L_2$ distance performs the best for tree species of Camphora, with OA higher than that from VBSD for $\chi^2$ distance (1.55%). The accuracies using VBSD for $L_1$ distance are between the other two distances for all the four tree species.
with reference tree-crown objects, it is found that many under-estimated areas are distributed on

Remote Sens. 2019, 11, 1982  

over-estimated areas are around the target tree species (in blue circle of Figure 7). These di

areas of target tree species shown in the reference maps are correctly recognized. From Figure 7,

standard deviations than those from the comparative method.

higher than that from the comparative method (4–6%). This demonstrates that the proposed method

comparatively few using both methods (Figure 7a). By comparing

the OA(%) Kappa F1

Table 3. Mapping accuracies of four tree species using the proposed method with VBSDs for

Three Distances

Metasequoia

| Tree Species  | VBSDs for Three Distances | Accuracy Measures |
|---------------|---------------------------|-------------------|
|               | L1                        | PA(%) | UA(%)   | OA(%)     | Kappa | F1   |
| Metasequoia   | 90.11 ± 2.196             | 96.35 ± 0.883    | 95.57 ± 0.731 | 0.899 ± 0.017 | 0.931 ± 0.012 |
|               | L2                        | 89.87 ± 1.340    | 96.10 ± 0.557 | 95.41 ± 0.446 | 0.895 ± 0.011 | 0.929 ± 0.007 |
|               | χ²                        | 90.91 ± 1.743    | 96.14 ± 0.068 | 95.75 ± 0.581 | 0.903 ± 0.014 | 0.934 ± 0.009 |

For comparative method using object-based method and OCSVM, 30 tree-crown objects were

Table 4. Mapping accuracies of four tree species using the comparative method.

| Tree Species  | Accuracy Measures |
|---------------|-------------------|
|               | PA(%) | UA(%)   | OA(%) | Kappa | F1   |
| Metasequoia   | 86.80 ± 5.480    | 98.45 ± 0.355 | 95.14 ± 1.775 | 0.887 ± 0.043 | 0.922 ± 0.031 |
| Platanus I    | 83.39 ± 6.857    | 83.04 ± 1.432 | 88.78 ± 1.860 | 0.747 ± 0.047 | 0.831 ± 0.035 |
| Platanus II   | 71.34 ± 5.912    | 84.16 ± 3.218 | 85.93 ± 2.033 | 0.671 ± 0.050 | 0.771 ± 0.038 |
| Camphora      | 67.20 ± 13.014   | 67.71 ± 5.089  | 78.39 ± 4.235 | 0.510 ± 0.109 | 0.669 ± 0.088 |

For comparative method using object-based method and OCSVM, 30 tree-crown objects were

selected as training samples for each tree species. The mapping results from the comparative method

are shown in Table 4. By comparing Tables 3 and 4, the proposed method using VBSDs for three
different distances produces higher accuracies with smaller standard deviations than those from the
comparative method for all four tree species. For example, OAs of mapping results using the proposed
method are 0.61%, 2.57%, 4.01% and 5.81% higher than those from the comparative method for
tree species of Metasequoia, Platanus I, Platanus II and Camphora, respectively. In addition, it is worth
noting that the OAs of mapping results for Camphora from the proposed method are significantly
higher than that from the comparative method (4–6%). This demonstrates that the proposed method
significantly outperforms the comparative method for the target class with lower mapping accuracy.
In addition, the PAs and UAs using the proposed method are also higher or comparable with smaller
standard deviations than those from the comparative method.

Figure 7 shows portions of mapping results using different methods for four tree species. Most
areas of target tree species shown in the reference maps are correctly recognized. From Figure 7,
there is a significant reduction of the under-estimated areas and the over-estimated areas when
using the proposed method than the comparative method. However, the under-estimated areas and
over-estimated areas for Metasequoia are comparably few using both methods (Figure 7a). By comparing
with reference tree-crown objects, it is found that many under-estimated areas are distributed on
the edges of trees (in red circle of Figure 7), which are the boundaries of continuous trees. Many
over-estimated areas are around the target tree species (in blue circle of Figure 7). These different tree
species are planted too close and their tree crowns tend to obscure each other.
Figure 7. Portions of mapping results for four tree species using the proposed method (using VBSD for \( \chi^2 \) distance) and the comparative method according to the reference maps labeling tree-crown objects for a specific tree species: *Metasequoia* (a); *Platanus I* (b); *Platanus II* (c); *Camphora* (d). Some under-estimated areas on the edges of trees are circled in red and some over-estimated areas around the target tree species are circled in blue.

The best mapping results over study area using the proposed method are shown in Figure 8. Most target tree species in reference map are correctly extracted in these four mapping results (shown in different colors for each tree species). The spatial distributions of each tree species are also confirmed. There are fewer over-estimated and under-estimated areas for *Metasequoia*, because this species has more distinct characteristics than the others (Figure 8a). From Figure 8b, most of the tree crowns of *Platanus I* are accurately mapped. Similarly, major tree crowns of *Platanus II* are accurately extracted, although some of these continuous tree crowns are under-estimated (Figure 8c). The mapping result for *Camphora* is not as accurate as the other species. However, the target tree crowns distributed in the northeast of this study area are mainly extracted (Figure 8d). From Figure 8b,d, it is found that the tree species of *Platanus I* shows a confusion with the tree species of *Camphora*, which appears in their over-estimated areas. This is consistent with the previous object histograms shown in Figure 5 that *Platanus I* and *Camphora* are more similar.
5. Discussion

In this study, we proposed a new method of tree species mapping from UAV images. Tree-crown object histograms of spectral and height related features were used to quantify characteristics of different tree species. VBSD [38], a recently proposed histogram similarity measure, was used to quantitatively compare target object histograms with reference object histograms. A state-of-the-art method using conventional object-based method and OCSVM, was used for comparison. The proposed method was evaluated in mapping four different tree species over an urban area. The results showed that the proposed method outperformed the comparative method for all the four tree species.

The proposed method shows the following advantages. First, instead of using image objects, tree-crown objects are extracted and used as processing units. The use of tree-crown objects in the proposed method reduces confusion with other irrelevant objects in study area, such as grass and other non-vegetation, and also avoids merging different tree species into one image object. Moreover, the histograms of tree-crown objects are more representative of a specific tree species.

Second, joint use of spectral and height related features of tree-crown objects is effective in tree species mapping. Although spectral features are helpful in mapping tree species (e.g., Figure 5), the
sole use of spectral features in mapping tree species may produce limited accuracy due to spectral similarity between different tree species and illumination variation of UAV images. The height related features provide useful structural information of tree-crown objects. However, using height related features alone could not distinguish between tree species with similar heights and structures. Therefore, the combination of spectral features and height related features will provide complementary information for mapping tree species in urban area. In particular, object-level histogram of spectral and height related features is found to well quantify characteristics of tree-crown objects.

Third, VBSD used in this study is a feasible histogram similarity measure when mapping tree species. The VBSDs for three different distances all performed well. Because the bin size of the histogram is variable, the cross-bin distance considers the correlations between neighboring bins and reduces the sensitivity to bin size [38]. Generally, relatively refined bin size of object histogram is selected as initial bin size in practical applications.

Fourth, the proposed method based on a simple threshold on VBSD only needs a small amount of reference samples of a specific tree species and does not need to select non-target samples. The representativeness of reference samples is important in the proposed method, but not the quantity of samples. However, in the conventional object-based mapping method, OCSVM classifier is affected by quantity of training samples.

It should be noted that there are several thresholds to be determined in the proposed method. In this study, we selected a threshold range by analyzing the distribution of VBSD values. The threshold range provides a guidance to determine the optimal threshold. The threshold values for other variables were determined by trial and error [22,25,26,28]. The reference for determining threshold and more automatic threshold selection method would be further explored.

Although the proposed method shows very promising results in tree species mapping using UAV images over urban area, there are still some problems to be further addressed in the future. For example, illumination variation of the UAV images is an important concern. Because ground targets are densely distributed over urban area, some trees are in shadow of tall buildings. There are spectral differences between the side of tree crowns towards the sun and the side away from the sun. This may affect the shape of tree-crown objects, the similarity of object histograms and the selection of reference samples.

Given that the evaluation was only done in a study area with limited coverage, the proposed method will be further evaluated by applying more UAV images in other areas with different tree species. VHR spectral and height data obtained in other ways, such as airborne LiDAR data, and other features will be explored in the future.

6. Conclusions

In this paper, a tree species mapping method using UAV images was proposed. In the proposed method, multiple feature histograms of tree-crown objects were used to quantify different tree species, where spectral features and height related features were selected as object features. The VBSD was used to measure the histogram similarity in each tree-crown object. A specific tree species was extracted by thresholding the VBSDs obtained. The experimental results demonstrated that the proposed method produced higher accuracies for all the four tree species than the existing method in the study area. The proposed method provides a simple and effective way for tree species mapping in urban area. More studies are needed to further validate the performance of the proposed method in other areas.

**Author Contributions:** P.L. and X.F. conceived and designed the paper. X.F. performed the experiments, analyzed the data, and wrote the paper. P.L. analyzed the data and revised the paper.

**Funding:** This research was funded by the National Science Foundation of China (Grant Number 41371329).

**Acknowledgments:** The authors thank Chun Liu from Tongji University for providing the data used in this study. We also thank Tongfan Surveying Engineering and Technology Co. Ltd for their technical support in data generation.

**Conflicts of Interest:** The authors declare no conflicts of interest.
References

1. Dwyer, J.F.; McPherson, E.G.; Schroeder, H.W.; Rowntree, R.A. Assessing the benefits and costs of the urban forest. *J. Arboric.* 1992, 18, 227.

2. Davies, Z.G.; Edmundson, J.L.; Heinemeyer, A.; Leake, J.R.; Gaston, K.J. Mapping an urban ecosystem service: Quantifying above-ground carbon storage at a city-wide scale. *J. Appl. Ecol.* 2011, 48, 1125–1134. [CrossRef]

3. Armson, D.; Stringer, P.; Ennos, A. The effect of tree shade and grass on surface and globe temperatures in an urban area. *Urban For. Urban Green.* 2012, 11, 245–255. [CrossRef]

4. Berland, A.; Shiflett, S.A.; Shuster, W.D.; Gar mestani, A.S.; Goddard, H.C.; Herrmann, D.L.; Hopton, M.E. The role of trees in urban stormwater management. *Landscape Urban Plan.* 2017, 162, 167–177. [CrossRef] [PubMed]

5. Dallimer, M.; Irvine, K.N.; Skinner, A.M.J.; Davies, Z.G.; Rouquette, J.R.; Maltby, L.L.; Warren, P.H.; Armsworth, P.R.; Gaston, K.J. Biodiversity and the Feel-Good Factor: Understanding Associations between Self-Reported Human Well-being and Species Richness. *Bioscience* 2012, 62, 47–55. [CrossRef]

6. Escobedo, F.J.; Kroeger, T.; Wagner, J.E. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. *Environ. Pollut.* 2011, 159, 2078–2087. [CrossRef] [PubMed]

7. Yan, P.; Yang, J. Performances of Urban Tree Species under Disturbances in 120 Cities in China. *Forest* 2018, 9, 50. [CrossRef]

8. Fassnacht, F.E.; Latifi, H.; Stere´nczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* 2016, 186, 64–87. [CrossRef]

9. Ardila, J.P.; Tolpekin, V.A.; Bijker, W.; Stein, A. Markov-random-field-based super-resolution mapping for identification of urban trees in VHR images. *ISPRS J. Photogramm. Remote Sens.* 2011, 66, 762–775. [CrossRef]

10. Li, D.; Ke, Y.; Gong, H.; Li, X. Object-Based Urban Tree Species Classification Using Bi-Temporal WorldView-2 and WorldView-3 Images. *Remote Sens.* 2015, 7, 16917–16937. [CrossRef]

11. Pu, R.; Landry, S. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sens. Environ.* 2012, 124, 516–533. [CrossRef]

12. Xiao, Q.; Ustin, S.L.; McPherson, E.G. Using AVIRIS data and multiple-masking techniques to map urban forest tree species. *Int. J. Remote Sens.* 2004, 25, 5637–5654. [CrossRef]

13. Alonzo, M.; Roth, K.; Roberts, D. Identifying Santa Barbara’s urban tree species from AVIRIS imagery using canonical discriminant analysis. *Remote Sens. Lett.* 2013, 4, 513–521. [CrossRef]

14. Zhang, Z.; Qiu, F. Mapping Individual Tree Species in an Urban Forest Using Airborne Lidar Data and Hyperspectral Imagery. *Photogramm. Eng. Remote Sens.* 2012, 78, 1079–1087. [CrossRef]

15. Alonzo, M.; Bookhagen, B.; Roberts, D.A. Urban tree species mapping using hyperspectral and lidar data fusion. *Remote Sens. Environ.* 2014, 148, 70–83. [CrossRef]

16. Zhang, Z.; Kazakova, A.; Moskal, L.M.; Styers, D.M. Object-Based Tree Species Classification in Urban Ecosystems Using LiDAR and Hyperspectral Data. *Forest* 2016, 7, 122. [CrossRef]

17. Liu, L.; Coops, N.C.; Aven, N.W.; Pang, Y. Mapping urban tree species using integrated airborne hyperspectral and LiDAR remote sensing data. *Remote Sens. Environ.* 2017, 200, 170–182. [CrossRef]

18. Inmmitzer, M.; Atzberger, C.; Koukal, T. Tree Species Classification with Random Forest Using Very High Spatial Resolution 8-Band WorldView-2 Satellite Data. *Remote Sens.* 2012, 4, 2661–2693. [CrossRef]

19. Boschetti, M.; Boschetti, L.; Oliveri, S.; Casati, L.; Canova, I. Tree species mapping with Airborne hyper-spectral MIVIS data: The Ticino Park study case. *Isprs J. Photogramm. Remote Sens.* 2011, 66, 762–775. [CrossRef]

20. Lu, B.; He, Y.; Liu, H. Investigating Species Composition in a Temperate Grassland Using Unmanned Aerial Vehicle-Acquired Imagery. In Proceedings of the 4th International Workshop on Earth Observation and Remote Sensing Applications (EORSA), Guangzhou, China, 4–6 July 2016; pp. 107–111.

21. Gini, R. Use of Unmanned Aerial Systems for multispectral survey and tree classification: A test in a park area of northern Italy. *Eur. J. Remote Sens.* 2014, 47, 251–269. [CrossRef]

22. Franklin, S.E.; Ahmed, O.S. Deciduous tree species classification using object-based analysis and machine learning with unmanned aerial vehicle multispectral data. *Int. J. Remote Sens.* 2018, 39, 5236–5245. [CrossRef]

23. Lisein, J.; Michez, A.; Claessens, H.; Lejeune, P. Discrimination of Deciduous Tree Species from Time Series of Unmanned Aerial System Imagery. *PLoS ONE* 2015, 10, e0141006. [CrossRef] [PubMed]
24. Hill, D.J.; Tarasoff, C.; Whitworth, G.E.; Baron, J.; Bradshaw, J.L.; Church, J.S. Utility of unmanned aerial vehicles for mapping invasive plant species: A case study on yellow flag iris (Iris pseudacorus L.). Int. J. Remote Sens. 2017, 38, 2083–2105. [CrossRef]

25. Michez, A.; Piégay, H.; Jonathan, L.; Claessens, H.; Lejeune, P. Mapping of riparian invasive species with supervised classification of Unmanned Aerial System (UAS) imagery. Int. J. Appl. Earth Obs. Geoinf. 2016, 44, 88–94. [CrossRef]

26. Alvarez-Taboada, F.; Paredes, C.; Julián-Pelaz, J. Mapping of the Invasive Species Hakea sericea Using Unmanned Aerial Vehicle (UAV) and WorldView-2 Imagery and an Object-Oriented Approach. Remote Sens. 2017, 9, 913. [CrossRef]

27. Saarinen, N.; Vastaranta, M.; Näsi, R.; Rosnell, T.; Hakala, T.; Honkavaara, E.; Wulder, M.A.; Luoma, V.; Tommaselli, A.M.G.; Imai, N.N.; et al. Assessing Biodiversity in Boreal Forests with UAV-Based Photogrammetric Point Clouds and Hyperspectral Imaging. Remote Sens. 2018, 10, 338. [CrossRef]

28. Cao, J.; Leng, W.; Liu, K.; Liu, L.; He, Z.; Zhu, Y. Object-Based Mangrove Species Classification Using Unmanned Aerial Vehicle Hyperspectral Images and Digital Surface Models. Remote Sens. 2018, 10, 89. [CrossRef]

29. Liu, C.; Ai, M.; Chen, Z.; Zhou, Y.; Wu, H. Detection of Firmiana danxiaensis Canopies by a Customized Imaging System Mounted on an UAV Platform. J. Sens. 2018, 2018, 1–12. [CrossRef]

30. Stow, D.A.; Toure, S.I.; Lippitt, C.D.; Lippitt, C.C.; Lee, C.R. Frequency distribution signatures and classification of within-object pixels. International J. Appl. Earth Obs. Geoinf. 2012, 15, 49–56. [CrossRef]

31. Toure, S.I.; Stow, D.A.; Weeks, J.R.; Kumar, S. Histogram curve matching approaches for object-based image classification of land cover and land use. Photogramm. Eng. Remote Sens. 2013, 79, 433–440. [CrossRef]

32. Zhou, Y.; Qiu, F. Fusion of high spatial resolution WorldView-2 imagery and LiDAR pseudo-waveform for object-based image analysis. ISPRS J. Photogramm. Remote Sens. 2015, 101, 221–232. [CrossRef]

33. Liu, J.; Li, P. Extraction of Earthquake-InducedCollapsed Buildings from Bi-Temporal VHR Images Using Object-Level Homogeneity Index and Histogram. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2019, 1–16. [CrossRef]

34. Liu, Z.; Jiang, J.; Yang, H.; Wen, S.; Li, A. Study on bio-ecological characteristics of Platanus acerifolia in Blooming and Bearing II Changing of flower buds’ development with increase of DBH. J. Henan Agric. Univ. 2002, 36, 54–58. (In Chinese with English abstract)

35. Wang, X.; Ma, L.; Guo, B. History and Research Process on the Silviculture of Metasequoia glyptostroboides in China. J. Northwest For. Univ. 2004, 19, 82–88. (In Chinese with English abstract).

36. Mao, C. Research on Introduction and Domestication of Camphora Camphora (Linn.). Presl. J. Shandong Agric. Univ. (Nat. Sci.) 2004, 35, 534–539. (In Chinese with English abstract).

37. Mesas-Carrascosa, F.J.; Torres-Sánchez, J.; Clavero-Rumbao, I.; García-Ferrer, A.; Peña, J.M.; Borra-Serrano, I.; López-Granados, F. Assessing optimal flight parameters for generating accurate multispectral orthomosaics by UAV to support site-specific crop management. Remote Sens. 2015, 7, 12793–12814. [CrossRef]

38. Ma, Y.; Gu, X.; Wang, Y. Histogram similarity measure using variable bin size distance. Comput. Vis. Image Underst. 2010, 114, 981–989. [CrossRef]

39. Liu, J.; Li, P.; Wang, X. A new segmentation method for very high resolution imagery using spectral and morphological information. ISPRS J. Photogramm. Remote Sens. 2015, 101, 145–162. [CrossRef]

40. Belgiu, M.; Drăguţ, L. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. ISPRS J. Photogramm. Remote Sens. 2014, 96, 67–75. [CrossRef] [PubMed]

41. Baatz, M.; Schäpe, A. Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In Angewandte Geographische Informationsverarbeitung XII, 1st ed.; Strobl, J., Blaschke, T., Griesebner, G., Eds.; Herbert Wichmann Verlag: Heidelberg, Germany, 2000; pp. 12–23.

42. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 2004, 58, 239–258. [CrossRef]

43. Trimble eCognition Developer 9.0 Reference Book. 2014. Available online: www.eCognition.com (accessed on 6 July 2014).

44. Savitzky, A.; Golay, M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Anal. Chem. 1964, 36, 1627–1639. [CrossRef]
45. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. Remote Sens. Environ. 2004, 91, 332–344. [CrossRef]
46. Zhou, W.; Huang, G.; Troy, A.; Cadenasso, M. Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: A comparison study. Remote Sens. Environ. 2009, 113, 1769–1777. [CrossRef]
47. Congalton, R.G.; Green, K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices; Lewis Publishers: Boca Raton, FL, USA, 1999; 137p.
48. Koga, Y.; Miyazaki, H.; Shibasaki, R. A CNN-Based Method of Vehicle Detection from Aerial Images Using Hard Example Mining. Remote Sens. 2018, 10, 124. [CrossRef]
49. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.; Smola, A.J.; Williamson, R.C. Estimating the Support of a High-Dimensional Distribution. Neural Comput. 2001, 13, 1443–1471. [CrossRef]