Review on the Use of Remote Sensing for Urban Forest Monitoring

Razieh Shojanoori and Helmi Z. M. Shafri

Abstract. Urban forests are vital in urban areas because they clean the air, absorb water, and protect the environment from intense heat. Destruction of the urban forest by increased urbanization is a considerable threat to the ecosystem. Hence, urban planners must obtain and manage information about urban forests, but the complexity of urban areas has made these tasks difficult. With developments in remote-sensing technologies, the monitoring and detection of urban forests can be achieved without performing any field measurements. In this study, different remote-sensing imageries and various methods are evaluated to obtain urban forest information. This review demonstrates that very high resolution (VHR) satellite imagery, such as from WorldView-2, is the most efficient data that can be used to obtain urban forest information. The use of the combination of LiDAR data with VHR imagery increases the accuracy of information, particularly about tree crown delineation. Traditional pixel-based classification methods are not effectively applicable to obtain urban tree information because of significant spectral variability in urban areas. An object-based classification technique, which uses spatial, textural, and color information, can be a potential method to detect urban forest and tree species discrimination. The new VHR imaging method, which uses the object-based technique, is recommended to overcome limitations of collecting urban forest information.

Key Words. LiDAR; Object-Based Classification; Pixel-Based Classification; Satellite Imagery; Tree Crown Delineation; Tree Species Detection.

Urban area is commonly defined as having spaces with artificial surfaces and a dense population, with vegetation that covers an urban area is considered an urban forest. Urban forests may include any type of vegetation in a metropolitan area; types of vegetation include trees, shrubs, and woody plants on the roadside, or plants on a larger scale, such as a forest park.

Urban forests not only result in social and economic advantages, such as recreational spaces and tourism, but they also provide several benefits to the ecosystem. The most important effect of urban forests in the ecosystem is with regard to the urban atmosphere, and can be divided into four main issues: 1) influence on microclimate and decrease in temperature because shadows are created on surfaces, wind speeds vary, and water is transpired; 2) air pollution removal; 3) release of volatile carbon-based compounds; and 4) energy waste minimization by providing shade to buildings to reduce consumption of heating and cooling energy.

Urban forests also affect urban hydrology. Urban trees can help solve issues related to water quality by reducing the overflow volume of storm water and minimizing the destruction of flooding because water is absorbed by the roots and soil erosion is avoided. Aside from problems related to the urban atmosphere and hydrology, sound pollution is another issue in urban areas. The leaves and stems of urban trees and shrubs can reduce noise by scattering the sound; thus, planting trees and shrubs in an urban area leads to a decrease in sound pollution (Nowak and Dwyer 2007). In conclusion, the most important advantages of the urban forest include the following: protection of biodiversity, avoidance of soil erosion, carbon storage, nutrient cycling, improvement of air and water quality, slowing wind and reducing water volume caused by storms, moderating local climate, providing shade to homes, and decreasing heat islands (Conine et al. 2004; Gobster and Westphal 2004; Xiao and
McPherson 2005; Huang et al. 2007; Shahidan et al. 2010; Ma and Ju 2011; Akamphon 2014).

As urban forests are vital in ecology, their management must include strategic and appropriate urban design and planning (Kong and Nakagoshi 2005; Huang et al. 2007; Iovan et al. 2008). Nevertheless, rapid urbanization, which poses a threat to the safety of an ecosystem (Hepinstall-Cymerman et al. 2013), has compelled scholars to focus on urban greenspaces. Human society seems to have realized that living without nature is difficult and unsafe (Kong and Nakagoshi 2005; Li et al. 2010).

With progressing urbanization, managing urban forests has become a significant concern. The growth of residential and commercial areas can negatively affect vegetation and ecology. Hence, one of the issues in urban forestry is determining the state and quantity of urban vegetation and buildings and controlling their growth and deterioration (Kong and Nakagoshi 2005; Iovan et al. 2008; Gillespie et al. 2012). Sufficient knowledge on urban forests, such as knowledge of tree location, size, and species, is essential for effective urban forestry (Ardila et al. 2012).

Manual field measurement was the earliest method used to study urban forests (Francis 1987). In this method, the whole city or some parts of an area are randomly selected for sampling (Nowak et al. 2008). However, an urban space is a complex area, and obtaining information on all the trees and vegetation spaces via field surveying is difficult, time consuming, and may provide inaccurate results. These limitations can be overcome by using remote sensing to obtain accurate information by monitoring and controlling urban areas and vegetation (Ardila et al. 2012). The present article reviewed several studies that utilized remote-sensing techniques to investigate urban forests; evaluate the potential of remote sensors and different methods, such as pixel-based and object-based; and obtain accurate urban forest information.

EVALUATION OF DIFFERENT SATELLITE IMAGING TECHNIQUES

Field surveys and visual interpretations of aerial imagery are conventionally used to generate a vegetation cover map, but these methods are costly and time consuming. Therefore, scientists integrated traditional tools into new remote-sensing systems (Huang et al. 2007; Iovan et al. 2008); such tools include different satellite images, passive optical systems, and active sensors. Moderate resolution imaging spectroradiometer (MODIS) is medium-resolution imagery used to monitor urban forests with a spatial resolution of 250–500 m. MODIS and Landsat imageries are multi-temporal remote-sensing tools, and their most important characteristic is the ability to obtain seasonal and annual information on different types of vegetation and land covers (Peijun et al. 2010; Zheng and Qui 2012; Qu et al. 2014). Vegetation covers have been shown to have different patterns in time-series experiments under various conditions, such as humidity, because of their potential to combine various information or species compositions (Zheng and Qui 2012). Nevertheless, MODIS and Landsat exhibit temporal limitations (16 days of repeat cycles) (Shouse et al. 2013). Moderate-resolution imageries often have mixed pixels because of their low spatial resolution (approximately 30 m), and thus they cannot be defined as a specific pure class (Peijun et al. 2010) and can only detect land-use types at the city level (Huang et al. 2007; Zheng and Qui 2012).

The other medium-resolution satellite systems used to study urban forests are Landsat systems, which can provide a means to rapidly monitor urban forests (Huang et al. 2007; Zhang et al. 2007). The visual interpretation of Landsat TM shows that bands 2 and 4 provide sufficient information on land-cover types, and an image with a false color composite of band 4-5-3 (R-G-B) clearly differentiates vegetation types, particularly when the adaptive enhancement technique is used (Kamaruzaman and Haszuliana 1996; Ismail and Jusoff 2004; Cai et al. 2010). However, when the study area is bound with a compact plantation because of spectral similarity, small urban and clearance areas are difficult to separate from tree species and mixed agriculture crops (Ismail and Jusoff 2004). Huang (2007) demonstrated that the Landsat ETM + imagery of spread urban trees is less coarse than other medium-resolution imageries.

High-resolution satellites have been developed to overcome the limitation of moderate-resolution imageries, such as low spatial resolution. Hence, high-resolution imageries, such as those from QuickBird (Tooke et al. 2009; Hashiba et al. 2004; Ardila et al. 2012), IKONOS (Green-
berg et al. 2009; Ma and Ju 2011; Pu and Landry 2012), SPOT (Kong and Nakagoshi 2005), RapidEye (Tigges et al. 2013), FormoSat-2 (Sun et al. 2007), and WorldView-1 and WorldView-2 (Immitzer et al. 2012; Latif et al. 2012; Nouri et al. 2014; Rapinel et al. 2014) have been increasingly used to detect and monitor urban forests.

QuickBird and IKONOS are common high-resolution satellite imaging techniques in urban forest studies, and both have panchromatic and four multispectral bands (i.e., red, green, blue, and near infrared) with high spatial resolution (HSR) (Pan: 0.6 to MS: 2.44 m). As red and near-infrared bands are sensitive to vegetation and contain approximately 90% of the vegetation information, they can be used to detect vegetation (Li et al. 2010; Puissant et al. 2014). Nonetheless, more bands may be required to extract information on vegetation and trees from different land-cover types because of the complex environment of urban areas (Ouma and Tateishi 2008). As the resolution of the multispectral image is 2.4 m, objects smaller than 6 m present mixed pixels, whose spectral characteristics refer to mixed objects, such as roads and trees. To avoid this error, researchers have used different techniques. For instance, Hong et al. (2009) applied the grey-level co-occurrence matrix (GLCM) mask and hierarchical classification to improve the accuracy of the information, but these methods remain incapable of performing high-accuracy extraction in urban areas (Hong et al. 2009). Several high-resolution imaging techniques, such as IKONOS and QuickBird, are insufficient in classifying urban vegetation into different species because of the limited number of spectral bands.

The WorldView-2 satellite was launched in 2009 as an improved version of high-resolution satellites. This high-resolution imagery system employs eight spectral bands, including those sensitive to vegetation. These bands consist of four old bands, namely, blue, green, red, and near-infrared, and four new bands, namely, coastal (to detect chlorophyll content), yellow (to detect yellowness), red edge (to detect plant diseases and vegetation species), and near-infrared 2 (to study biomass) (Immitzer et al. 2012). Immitzer et al. (2012) demonstrated that WorldView-2 can possibly detect urban forests because of its four new bands (Pu 2009). Nevertheless, some misclassifications were observed in the classification of tree species because of spectral overlaps, complex structure of the area, and small tree crown, resulting in mixed pixels. Airborne hyperspectral sensor is an excellent sensor that can be used to overcome spectral limitations. Ghiyamat and Shafri (2010) demonstrated that hyperspectral imagery provides adequate data to distinguish homogeneous and heterogeneous forest biodiversity, but urban areas present different environment characteristics and should be evaluated separately.

Hyperspectral data present characteristics, such as narrow band, multi-channel, and continuous spectrum information, which can be used to detect urban vegetation (Hao et al. 2011). Several studies on urban forests have been conducted using hyperspectral data (Wania and Weber 2007; Hao et al. 2011; Cho et al. 2012; Zhang and Qiu 2012; Adeline et al. 2013; Forzieri et al. 2013). Most researchers have demonstrated the effectiveness of hyperspectral data to accurately detect vegetation and even tree species. Nevertheless, hyperspectral data exhibit specific limitations, such as limited coverage, high volume, and high cost (Shafri et al. 2012), thereby compelling researchers to use high-resolution satellite imageries.

With technological developments in remote sensing, active sensors have been used to detect urban forests. Many traditional satellites can accurately detect tree species, but they can only delineate urban features and tree crowns in 2D by using reflected solar radiation. By contrast, active sensors, such as synthetic aperture radar (SAR) and light detection and ranging (LiDAR), can extract the shape of the tree crown and urban features in 3D even in the shadows and at night (Yao and Wei 2013; Zhou 2013; Maksymiuk et al. 2014). Therefore, the use of active sensors has improved the monitoring of urban forests.

Although the benefits of RADAR sensors, such as SAR, to detect forests have been noted (Perko et al. 2010), related studies are mostly limited to large-scale natural forest classification and tree root evaluation, whereas those on urban forests are rare and new. Maksymiuk et al. (2014) utilized SAR data with morphological attribute filters and detected single urban trees. Accordingly, further studies should be performed to evaluate the potential of SAR data to
detect urban trees at a large scale and differentiate different tree species from one another. In contrast to RADAR data, LiDAR data on urban forest have been widely used and adopted (Forzieri et al. 2009; Nicholas et al. 2012; Oshio et al. 2012; Zhang and Qiu 2012; Zhu et al. 2012; Adeline et al. 2013; Nieneyer et al. 2013; Oshio et al. 2013; Zhou 2013).

Sung (2012) applied LiDAR data to assess the mean canopy height and percent canopy cover of an urban forest. According to the land development ordinance, the landowner should secure a permit to remove trees larger than 41 cm in diameter at breast height. Sung also utilized LiDAR data to establish the canopy height model (CHM) by calculating the difference between the digital elevation model (DEM) and the digital surface model (DSM) (only on tree canopy). The cells with values less than 1 m were not included in the analysis. Although the results demonstrated that LiDAR data are highly applicable and are recommended for use to detect the tree structure and evaluate the tree canopy height, the method mentioned (subtracting DEM from DSM) is not applicable in urban areas because the highest surfaces can sometimes be manmade materials, such as building roofs. When the uppermost surfaces are not tree canopies, the difference between DSM and DEM cannot be used to solve for CHM.

Many studies have been employed on different aspects of urban forests, including tree crown shape and structure (Oshio et al. 2012; Sung 2012; Oshio et al. 2013), tree detection and urban vegetation mapping (Hofle et al. 2012; MacFaden et al. 2012; Yao and Wei 2013; Zhou 2013), tree position and plant density (Forzieri et al. 2009), and individual tree species detection (Nicholas et al. 2012; Zhang and Qiu 2012). Zhang and Qiu (2012) solved the limitation of LiDAR data by using hyperspectral imaging to distinguish more than 10 tree species, and LiDAR data to detect tree crowns only; thus, the accuracy of tree species detection was directly related to the resolution of the images. In another study, Nicholas et al. (2012) applied discrete Fourier transform on LiDAR data to distinguish five individual tree species. In other studies, the discrete point of LiDAR data has been shown to be more accurate than that of airborne waveform Lidar data when performing tree classification (Reitberger et al. 2008; Hollaus et al. 2009; Heinzl and Koch 2011). Nicholas et al. (2012), who were the first to investigate and compare the accuracy of tree species detection by using the discrete point of LiDAR data and waveform LiDAR data, reported that the overall accuracy increases by about 6.2% when waveform information was used (overall accuracy was about 79.2% when the discrete point data were used but increased to 85.4% when the waveform LiDAR data were used). Although the use of airborne remote-sensing data, such as LiDAR data, has many advantages, it also presents several limitations for urban land-cover classification; such limitations include processing and interpolating of point clouds into raster layers, which are time consuming and prone to misclassification (Zhou 2013). Appendix I shows the satellite and airborne sensors for urban forest studies.

In conclusion, although the most suitable data that can be used to detect urban forests and distinguish urban tree species are hyperspectral data, they are not recommended in urban forest studies because of their limitations, such as limited coverage, high volume, and cost. Thus, other satellite imaging techniques should be used. Moderate-resolution imaging techniques is unsuitable for urban forest detection because of its low spatial resolution, which leads to mixed pixels. By contrast, high-resolution imaging techniques can be used to detect urban forests but cannot distinguish urban tree species unless an imaging technique with a higher resolution, such as WorldView-2, is used. The result of urban tree species classification using the traditional and new bands from the WorldView-2 imaging technique has demonstrated the effectiveness of this method to distinguish urban tree species from one another, although some misclassification might occur because of spectral similarity. This limitation can be addressed by using miscellaneous information, such as spatial, texture, and color. The last data, which are almost used as an ancillary data, are LiDAR data. Although these data can extract urban features even in the shadows and at night, the processing and interpolating of point clouds into raster layers are time consuming and may lead to some misclassifications. Finally, this review has shown that the use of data from WorldView-2 provides the most efficient way to detect urban forests if cost, time, and accuracy are considered as research factors.
URBAN TREE DETECTION

Tree detection by using remote-sensing images is the process of recognizing and classifying trees that lead to urban tree canopy and greenspace mapping (Lang et al. 2008; Walton et al. 2008; Johnson and Xie 2013). Mapping is conducted via urban forestry monitoring methods, which can be classified into three groups, namely, visual interpretation and pixel- and object-based methods (Li et al. 2010).

Although high-resolution data are valuable in extracting land-cover information, tree extraction and information collection are difficult in urban areas when traditional pixel-based image classification methods are used. Traditional methods involve supervised and unsupervised classifications. Although unsupervised classification techniques, such as ISODATA and K-mean, are used for thematic mapping (Langley et al. 2001; Xie et al. 2008; Sung 2012), these methods are rarely used for urban tree detection. Supervised classification methods, such as MLC, are often used to perform urban landcover mapping (including vegetation cover mapping) because they can be easily performed and provide accurate results (Ardila et al. 2010; Peijun et al. 2010; Shen et al. 2010; Zhang et al. 2010; Forzieri et al. 2013). The basis of MLC is a statistic classification of all pixels in each band to a specific class even when the threshold is defined. By contrast, MLC may cause some misclassifications in urban areas. For instance, some parts of the grass area are often classified as trees. Thus, filters, such as intra-class uniformity, inter-class contrast, and smoothness of boundaries between classes can be used to increase the contrast of features and obtain high classification accuracy (Ouma and Tateishi 2008). Minimum distance (MD) is another supervised classification for studies on urban forest, and several researchers believe that this method performs more accurate classification than other methods (Jusoff 2009; Latif et al. 2012). Shen et al. (2010) used three classification algorithms (i.e., ML, MD, and DT) for urban forest mapping and comparison of these algorithms showed that MD leads to the lowest classification accuracy and DT exhibits the highest accuracy among the three algorithms. The principle of DT classification differs from that of MLC, by which separation of the complicated decision to several easier decisions is vital to achieve the required classification (Ouma and Tateishi 2008).

Vapnik developed a new method called support vector machine (SVM) in 1996. This method can classify urban areas because it can overcome several limitations, such as insufficient training data and low sensitivity to the sample size (Van Der Lindan et al. 2007; Mountrakis et al. 2011). In this regard, several studies on urban forests have used the SVM algorithm to detect vegetation (Lafarage et al. 2005; Iovan et al. 2008; Tigges et al. 2013; Iovan et al. 2014).

These pixel-based classification algorithms may lead to low classification accuracy because of the high grade of spectral variability within land-cover classes, which are affected by sun angle, gaps in tree canopy, and shadows (Yu et al. 2006; Johnson and Xie 2013). A pixel, which is a small part of the classification object, is the cause of within-class spectral variability in high-resolution images (Huang et al. 2007). Thus, the use of object-based classification is recommended to overcome this limitation.

Object-based approaches can improve classification accuracy compared with visual interpretation and pixel-based methods. The object-based method can combine color, shape, spatial information, and contextual analysis to detect changes in vegetation (Li et al. 2010). The basis of object-oriented methods is image segmentation, which involves splitting the image into spatially continuous and homogeneous regions and leads to a reduction in local spectral variations (Lobo 1997). Li et al. (2010) combined segmentation and fuzzy multi-threshold classification to classify urban land covers, and the accuracy reached up to 93.72%. Fuzzy logic and intelligence techniques, such as artificial neural network (ANN), or integrated methods, such as adaptive Gaussian fuzzy-learning vector quantization (AGFLVQ), can also be used for urban forest detection and tree species identification (Hofle et al. 2012; Zhang and Qiu 2012). These classification methods not only detect urban forest but also distinguish urban tree species, and this ability is explained in the next section.

Comparison of several methods used in urban tree detection showed that object-based classification provides the most accurate classification result. Although pixel-based methods, such as ML and SVM, are easy to operate and provide accurate results for tree mapping, they are not adequate for use in urban areas, particularly in distinguishing urban tree species (as explained in the following section) because the basis of this classification is a pixel.
Pixel-based methods cannot provide high classification accuracy because of the spectral variability in urban areas. This limitation can be resolved by utilizing additional information about objects, such as spatial, textural, and color, in addition to spectral information. Hence, object-based classification is the optimal technique for urban tree detection.

**URBAN TREE SPECIES DETECTION**

Information on tree species is important for urban planning, disaster management, and ecological safety. Accurate, reliable, and expressive measurements on the types of urban vegetation can help urban planners and researchers reach their targets (Iovan et al. 2008; Hao et al. 2011; Gong et al. 2013). The concept of classification of tree species was introduced in forestry when satellite and aerial imaging techniques were used to monitor forests (Gougeon 1995). Numerous studies on the detection of tree species in forests are available (Immitzer et al. 2012), but research on urban areas remains limited.

The limitations of methods involving different satellites or airborne sensors is one of the challenges in studies on urban tree species. For instance, classical methods, such as MLC, can be applied on multispectral imaging, but these methods cannot be applied to hyperspectral data because of small training samples. Hence, other techniques, such as SAM (Wania and Weber 2007; Forzieri et al. 2013), linear spectral unmixing, and spectroscopic library matching are utilized to classify urban tree species by using hyperspectral data (Zhang and Qiu 2012). Forzieri et al. (2013) applied ML, SAM, and spectral information divergence on airborne hyperspectral data (i.e., multispectral infrared visible imaging spectrometer) to detect 10 urban tree species (i.e., herbaceous, heatland, arundo donax, poplar, oak, pine, Cupressus, spruce, willow, and olive), and ML presented the highest accuracy of up to 92.57%. This high accuracy could be attributed to the availability of LiDAR data because other researchers used them as ancillary or main data to improve the accuracy of classifying urban tree species (Voss and Sugumaran 2008; Tooke et al. 2009; Hofle et al. 2012; Zhang and Qiu 2012; Tigges et al. 2013).

Zhang and Qiu (2012) used LiDAR data to classify urban forest species based on tree crowns (i.e., crown-based species classification) because such data can address the limitation of tree crown-shaded side, small tree crowns (might be seen as one object), and boundary of tree crowns, which may lead to mixed pixels. These authors developed a method based on hyperspectral data by combining the fundamental aspect of the neural network and fuzzy logic. They used the AGFLVQ algorithm to distinguish 20 urban tree species, and the result demonstrated higher classification accuracy (approximately 68.8%) than that obtained using other hyperspectral methods, such as SAM (approximately 39.95%). The classification accuracy is less than the accuracy shown in the study of Forzieri et al. in 2013, although the difference can be due to the number and types of tree species (Forzieri et al.: 10 species, Zhang and Qiu: 20 species). For instance, when evergreen and deciduous trees are considered, the use of the Gaussian fuzzy learning vector quantization (GFLVQ) method is unsuitable because the determination of at least two spectra should be used for deciduous trees. However, the basis of the GFLVQ algorithm is that all species have the same spectral signatures, and one spectral signature is sufficient for evergreen species. This finding reflects the limitation of GFLVQ, which can be solved using ancillary data, such as LiDAR data (Zhang and Qiu 2012).

As multispectral data can be applied to different classification methods, most studies on urban forest species were conducted via multispectral imaging because of the other limitations of hyperspectral data (e.g., high volume, cost, and time required). Appendix II shows a summary of the detection of urban forest species via remote sensing and different classification methods. This review demonstrated that most study areas involve non-tropical areas, which contain abundant evergreen and deciduous species. Distinguishing evergreen from deciduous species by using their spectral signatures in spring or autumn is easier than differentiating tree species in a tropical area. As tree species present different spectral characteristics, spectral signatures are useful in each species. However, pixel-based classifications, such as ML, and MD with multispectral imaging but without any ancillary data, such as LiDAR, demonstrate low accuracy [e.g., Ismail and Jusoff (2004): ML approximately 61%]. By contrast, the complexity of the environment...
leads to a high spectral similarity between vegetation in urban areas. Furthermore, urban areas contain numerous pollutants that can change atmospheric conditions and affect spectral reflectance (Iovan et al. 2014). As a result, spectral signatures in multispectral imaging are insufficient to distinguish urban tree species. Thus, other characteristics of a tree species, such as spatial information, texture, and color, should be utilized to improve the classification of urban tree species.

The use of object-based classification is recommended to overcome the spectral limitation of multispectral imagery. Shouse et al. (2013) compared two classification methods (i.e., pixel- and object-based classification) in two types of multispectral imaging techniques (i.e., aerial and Landsat TM 5) to detect a species called bush honeysuckle (i.e., *Lonicera maackii*). The results showed that the object-based approach presents higher accuracy than the pixel-based approach, and HSR imaging demonstrates high accuracy [aerial (HSR): 94.2% / Landsat (MSR): 74.6%].

Textural information is an effective type of information that can be used to distinguish tree species. Iovan et al. (2008; 2014) used HSR data and SVM to distinguish urban tree species (*Platanus, Sophora, Tilia, Celtis, Pinus*, and *Cupressus*). As spectral information was inappropriate to be used independently, textural information, which involves information regarding the spatial and physical arrangement of objects, was utilized (Tso and Mather 2001). The results demonstrated that both methods of textural measurements (i.e., first- and second-order GLCM) could detect urban tree species and distinguish deciduous trees from coniferous species.

LiDAR data are optimal sources of information on texture or height. LiDAR data have been used in many studies on urban forest species (Voss and Sugumaran 2008; Tooke et al. 2009; Hofle et al. 2012; Zhang and Qiu 2012; Forzieri et al. 2013; Tigges et al. 2013). A robust technique for classification or segmentation is needed when the amount of information increases. Hofle et al. (2012) showed that intelligence algorithms, such as the ANN, are suitable for LiDAR information analysis. He applied two methods according to the object based-approach (ANN and DT) to detect six tree species, namely, *Fagus sylvatica*, *Acer platanoides*, *Platanus acerifolia*, *Tilia cordata*, *platyphyllos*, and *Aesculus hippocastanum*. The result showed that ANN presents a higher accuracy of 95% than DT with an overall accuracy of 72%. Spatial, textural, shape, or height information from LiDAR data can also be used to detect tree species. Zhu et al. (2012) showed that the spectral characteristics from LiDAR data are applicable to distinguishing real leaves from fake ones. Despite these results, trees often have approximately the same height and shape, and the high density of tree species may lead to the misclassification of tree species or small trees may be overlooked (Iovan et al. 2008; Latif et al. 2012).

The challenges of high-resolution imaging for the detection of urban tree species were highlighted when the new HSR imaging technique called WorldView-2 was launched. Pu and Landry (2012) attempted to perform segmentation via two methods (i.e., LDA and regression trees) to detect seven tree species (i.e., sand live oak, laurel oak, live oak, pine, palm, camphor, and magnolia) and demonstrated that the four new bands of WorldView-2 imaging can improve the accuracy by about 16% to 18% (compared with IKONOS imaging).

This review showed that given the complexity of urban areas and spectral similarity between tree species, high-resolution imaging via pixel-based methods is insufficient to discriminate urban tree species; hence, ancillary data such as DEM, spatial information, texture, and color, should be utilized. Although the use of LiDAR data, as ancillary data, presents high classification accuracy, higher classification accuracy cannot be achieved when these data are used separately. Thus, two imaging techniques (HSR and LiDAR) are proposed to obtain high classification accuracy, but these methods are not cost effective. By contrast, the object-based method was used to distinguish species by using information about urban tree species. In this regard, different high-resolution satellite imaging techniques were used, and WorldView-2 showed the highest accuracy over other high-resolution satellite imaging techniques. Accordingly, for urban tree species detection, the object-based techniques should be improved and integrated with WorldView-2 or with new high-resolution imaging techniques, such as WorldView-3.
Counting of trees via remote sensing is crucial in urban forest research because of plantation and urban management. As field measurements for tree counting are time consuming and may be expensive, remote-sensing imaging may be a suitable technique to overcome this limitation (Shafri et al. 2011).

Several studies on tree crown delineation and tree counting have been conducted via remote sensing (Brandtberg and Walter 1998; Andersen et al. 2001; Pouliot et al. 2002; Culvenor 2003; Erikson 2004; Heurich and Weinacker 2004; Karantzalos and Argialas 2004; Mei and Durrieu 2004; Pouliot et al. 2004; Falkowski et al. 2006; Horvath 2007; Wolf and Heipke 2007; Shafri et al. 2011; Katoh and Gougeon 2012; Wu et al. 2012). However, studies on urban areas remain limited and have only been conducted recently (Iovan et al. 2008; Ardila et al. 2010; MacFaden et al. 2011; Ardilla et al. 2012; Latif et al. 2012; Yao and Wei 2013; Iovan et al. 2014).

Remote sensing provides valuable information for urban forests, but conducting automatic tree extraction and counting from images is difficult because of the complexity of urban spaces. Furthermore, the resolution of multispectral images can affect the level of tree crown delineation (Iovan et al. 2008). For example, if one applies the pixel-based classification on high-resolution imaging, the crowns of individual tree species will be detected (Leckie 1990; Beaubien 1994; Erikson 2004). However, the low-spatial resolution imaging can only detect the strands of single species (Gillis and Leckie 1993).

Tree crown delineation should be conducted to count trees. This process presents the following limitations in urban areas: complexity of urban areas, including trees, buildings, roads, and sidewalks; different physical characteristics of trees, such as crown width, crown shape, height, and canopy cover; and different patterns for tree plantation that indicate equally or irregularly distance variations between trees.

Thus, spectral information should be minimized for tree identification and other information regarding trees (e.g., spatial, textural, or color) should be used. Several studies on tree crown delineation have been conducted by using the context and contours of trees. Any information on the object that can characterize its state and situation, which lead to its identification, is called context (Abowd et al. 1999; Oliva and Torralba 2007). Ardilla et al. (2012) performed a study involving most techniques for tree crown delineation by employing geographic object-based image analysis (GEOBIA) technique. Topologic relations between adjacent image objects and segmentation method can be established using GEOBIA. The following are the procedures to reach the tree crown: masking of grassland areas through size and NIR segmentation; identifying tree crown objects through NDVI, which is insensitive to the inside-crown brightness because of sun illumination and tree structure (Liang 2004); identifying individual trees with high background contrast; clustering trees based on size and NDVI segmentation; detecting small trees along roads through shadow, shape, and spectral characteristics (the use of the tree shadow is a common method to delineate a tree contour (Geugeon and Leckie 2001)); joining trees in linear and curvilinear patterns and detecting trees with low background contrast through local maxima filtering (Wulder et al. 2004); and combining the results via GEOBIA to identify tree crowns. Overall, 85% of the tree crown was detected by this method, and the error may be caused by the spatial resolution of the image, because small trees can cover less than two multispectral pixels (Gougeon and Leckie 2006; Ardilla et al. 2012). Nevertheless, most studies on tree crown detection have been conducted using high-resolution imaging (Ardila et al. 2012).

The other methods to delineate tree contours include the region-growing method (Geugeon 1995; Iovan et al. 2008), valley-following algorithms (Erikson 2004), multi-analysis (Brandtberg and Walter 1998; Ardilla et al. 2012), and active contours (Horvath et al. 2006). The seeded-region growing, developed by Iovan et al. (2008), is one example of region growing. This method is based on two steps, namely, seed point detection through DSM to evaluate treetops and the region-growing approach based on geometric criteria (height descent). Finally, the result showed that the accuracy for tree crown delineation by using this method was 78%. Tree crown delineation is based on 3D height information, and the cause of the error may be the low accuracy of the DSM.
Latif et al. (2012) emphasized that the cause for the low accuracy of tree crown delineation and tree counting is height. They used spectral information (i.e., MD and SAM) and segmentation via WorldView-2 imaging for tree crown delineation and tree counting. They demonstrated that each classified region could be considered as one tree after segmentation. The results of the segmentation and tree counting showed low accuracy because tree crown delineation is difficult when trees are clumped together and have similar height. Therefore, the research by Iovan et al. (2008) and Latif et al. (2012) indicated that obtaining the precise height information of trees is essential in improving the accuracy of tree counting.

Based on high-resolution imaging, the information is limited to 2D features and spectral, spatial, and contextual characteristics of trees. LiDAR data present trees in 3D, and the use of these data may lead to an improved accuracy of tree information, such as the height and width of crown base (Yao and Wei 2013). Yao and Wei (2013) demonstrated that the integration of LiDAR and aerial imaging yielded more accurate results in tree crown delineation than when the LiDAR data were used alone (Lin and Hyyppä 2012). Approximately 81% of individual trees were detected, and incorrectly detected trees can be explained by the misclassification of small trees or buildings that block the view.

Although several studies have demonstrated that the use of LiDAR data is necessary for tree crown delineation and counting, the high accuracy of the results from other studies indicated the possibility of urban tree crown delineation without the use of LiDAR data. Nevertheless, all of these researchers believed that in an urban area, the use of only the spectral information may not be feasible; thus, other information such as spatial and textural should be utilized. Moreover, processing the LiDAR data is time consuming and not as economical compared with high-resolution imaging. Thus, new high-resolution imaging techniques, such as the combination of WorldView-2 with an object-based technique, are recommended for use in urban tree crown delineation and tree counting. However, if the cost and processing time are not important factors to consider, the use of the LiDAR data alone is sufficient in increasing the accuracy of urban tree crown delineation.

In summary, tree counting is simplified through tree crown delineation, and field surveying is no longer needed. Studies have reported that most tree crown delineation and detection processes can be conducted automatically, but the counting remains semi-automatic. Particular studies have shown accurate results in semi-automatic tree counting in forests (Shafri et al. 2012), but fully automatic tree counting has neither been performed in a forest nor an urban area. To sum up, this review demonstrated that studies on urban tree crown delineation and tree counting are rare and further research must be performed.

CONCLUSION

This paper presented the status of urban forest monitoring involving remote sensing. First, different remote sensors used to generate urban vegetation maps were evaluated. Second, various classification methods used to extract urban forest information and distinguish urban tree species were assessed. Third, different methods for tree crown delineation and tree counting were discussed. This paper considered the most significant problems and mentioned the solution based on remote-sensing methods through related studies.

Remote-sensing imaging can detect urban forests, but different sensors have their own limitations. For instance, the use of hyperspectral data is optimal in extracting and distinguishing the urban forest based on spectral information, but high volume data, availability, and cost are limitations. The other data, which can detect urban features even in shadow and at night, is LiDAR data. However, the use of this data leads to misclassification and is time-consuming because it involves the conversion of point clouds into raster layers. Thus, other high-resolution imaging techniques, such as WorldView-2, are utilized.

Urban areas are complex environments, and the limitation of spectral information for multispectral imaging, particularly for distinguishing tree species, has compelled researchers to utilize other urban vegetation information during classification; such information includes spatial, textural, and color. Hence, the object-based approach is more applicable than traditional pixel-based classifications, such as MLC and MD, because the former can combine various data.
The last issue in studies on urban forest is related to tree crown delineation and tree counting. One of the main aspects of urban management is to determine the precise number of urban natural and planted trees. Studies have been conducted to perform tree crown delineation and tree counting in urban environments. The common methods used to achieve this purpose include the use of LiDAR data and WorldView-2 imaging. Although both techniques have shown acceptable results, WorldView-2 is more applicable because it is more economical and uses spectral, spatial, textural, and color information. These studies have enabled tree crown delineation to be done automatically, but tree counting is still a manual process. Thus, tree counting has to be performed semi-automatically and manually despite the fact that this process is time-consuming and expensive.

As remote-sensing techniques have been proven to detect and monitor urban forests, studies on remote sensing have gained increasing interest. Nonetheless, the work in this field remains limited, particularly in tropical areas. In conclusion, further studies on remote sensing in urban forests recommend to focus on the following: develop a vegetation index with an urban target, develop a high-accuracy algorithm to distinguish urban tree species automatically via high-resolution multispectral imaging, and develop a method for automatic tree counting.

Acknowledgments. The authors would like to thank the Ministry of Education (MOE) Malaysia and Universiti Putra Malaysia (UPM) for providing research grants through the Fundamental Research Grant Scheme (FRGS) and Research University Grant Scheme (RUGS). In addition, the anonymous reviewer comments helped to improve this manuscript and are highly appreciated.

LITERATURE CITED


©2016 International Society of Arboriculture


Résumé. Les forêts urbaines sont essentielles dans les communautés urbaines, car elles purifient l’air, absorbent l’eau et protègent l’environnement des chaleurs intenses. La destruction de la forêt urbaine par l’urbanisation croissante constitue une menace considérable pour l’écosystème. Par conséquent, les planificateurs urbains doivent recueillir et gérer les données sur les forêts urbaines, mais la complexité des aires urbaines a rendu ces tâches difficiles. Avec l’évolution des technologies de télédétection, la surveillance et le dépistage des forêts urbaines peuvent être obtenus sans effectuer de mesures sur le terrain. Pour cette étude, différentes images de télédétection et diverses autres méthodes ont été évaluées pour l’obtention de données sur la forêt urbaine. Cette revue des systèmes démontre que l’imagerie satellite à très haute résolution (VHR) , telle que le WorldView-2, génère les données les plus efficaces pouvant être utilisées pour obtenir des informations sur la forêt urbaine. L’utilisation des données LiDAR (télédétection par laser) en combinaison avec les images THR accentue la précision des informations, notamment sur la délimitation du houppier de l’arbre. Les méthodes de classification traditionnelle en pixels ne sont pas applicables de manière efficace afin d’obtenir des informations sur les arbres urbains du fait de la variabilité spectrale significative dans les zones urbaines. La technique de classification fondée sur les objets utilisant les informations spatiales, texturales et de couleur peut constituer une méthode plus efficace afin de surmonter les limites des données d’une forêt urbaine.

**APPENDIX I. Satellite and airborne sensors for urban forest studies.**

<table>
<thead>
<tr>
<th>Spatial resolution</th>
<th>Band</th>
<th>Dynamic range</th>
<th>Methods</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperspectral</td>
<td>1–5 m</td>
<td>VIS-NIR</td>
<td>ICARE (3D atmospheric correction code), NDVI-LAI relationship, DSM (0.25 resolution) Decision tree (DT) classifier LAI</td>
<td>Adeline et al. 2013</td>
</tr>
<tr>
<td></td>
<td>0.61 m</td>
<td>panchromatic</td>
<td>SRM based on Markov random fields (MRF), maximum likelihood classifier (MLC) GLCM, VTI building, and NDVI</td>
<td>Zhang and Qiu 2012 Hao et al. 2011</td>
</tr>
<tr>
<td>QuickBird</td>
<td>2.44 m</td>
<td>multispectral</td>
<td>Super-resolution mapping (SRM) based on Markov random fields (MRF), maximum likelihood classifier (MLC)</td>
<td>Ardila et al. 2010</td>
</tr>
<tr>
<td></td>
<td>0.61 m</td>
<td>panchromatic</td>
<td>Reproducible geographic object-based image analysis (GEOBIA)</td>
<td>Ardila et al. 2011 Ardila et al. 2012</td>
</tr>
<tr>
<td></td>
<td>R,G,B,NIR</td>
<td></td>
<td>Global environment monitoring index (GEMI) and NDVI</td>
<td>Hu 2011 Huang et al. 2007</td>
</tr>
<tr>
<td>MODIS</td>
<td>250</td>
<td>band 1–2</td>
<td>MODIS EVI and DT classifier</td>
<td>Zheng and Qui 2012 Hao et al. 2011</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>band 3–7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>band 8–36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Color infrared (CIR)</td>
<td>10 m</td>
<td>8 bit</td>
<td>DSM (3D), GLCM, SVM Random Forest (RF)</td>
<td>Iovan et al. 2008 Johnson and Xie 2013</td>
</tr>
<tr>
<td>SPOT</td>
<td>20 m</td>
<td></td>
<td></td>
<td>Kong and Nakagoshi 2005 Zhang et al. 2007 Liu et al. 2008</td>
</tr>
<tr>
<td>IKONOS</td>
<td>1 m</td>
<td>Pan.</td>
<td>NDVI, SR, ARVI, SAVI, LAI</td>
<td>Ma and Ju 2011</td>
</tr>
<tr>
<td></td>
<td>4 m</td>
<td>Multi.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAR</td>
<td></td>
<td></td>
<td></td>
<td>Maksymiuketal.2014</td>
</tr>
<tr>
<td></td>
<td>Spatial resolution</td>
<td>Band</td>
<td>Dynamic range</td>
<td>Methods</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------</td>
<td>------</td>
<td>---------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Landsat</td>
<td></td>
<td></td>
<td></td>
<td>NDVI and GEMI</td>
</tr>
<tr>
<td>MSS</td>
<td>60 m</td>
<td></td>
<td></td>
<td>MLC, NDVI</td>
</tr>
<tr>
<td>TM</td>
<td>30 m</td>
<td></td>
<td></td>
<td>MLC, NDVI</td>
</tr>
</tbody>
</table>
| ETM+          | 15–30 m            |      |               | RVI, NDVI, PVI, NDBI     | Peijun et al. 2010                
|               |                    |      |               |                          | Hasmadi and Jusoff 2004           |
|               |                    |      |               |                          | Gong et al. 2013                  |
|               |                    |      |               |                          | Cai et al. 2010                   |
| WorldView-2   | 0.5 m              | Pan. | 11 bit        | (RF) classification, Linear discriminant analysis (LDA) | Immitzer et al. 2012 |
|               |                    |      |               | Minimum Distance (MD) classification, spectral angle mapper (SAM), NDVI | Abd Latif et al. 2012 |
|               |                    | Multi|               |                          | Nouri et al. 2014                 |
## APPENDIX II. Satellites and airborne sensors for urban tree species detection.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Acquisition date</th>
<th>LiDAR</th>
<th>Classification algorithm</th>
<th>Species no.</th>
<th>Overall Acc.</th>
<th>Tree species details</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS40, MIVIS</td>
<td>December 2011</td>
<td>X</td>
<td>ML&gt;SAM, SID</td>
<td>10</td>
<td>92.57</td>
<td>Herbaceous, heatland, pine, Arundo donax, poplar, oak, Cupressus, spruce, willow, olive</td>
<td>Forzieri et al. 2013</td>
</tr>
<tr>
<td>Landsat TM</td>
<td>May 1996</td>
<td></td>
<td>Supervised &gt;</td>
<td>3</td>
<td>61</td>
<td>Oil palm, rubber tree, bush, grass (+6 different land-cover types)</td>
<td>Ismail and Jusoff 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unsupervised</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIR Sensor</td>
<td>2004</td>
<td></td>
<td>SVM (texture measure)</td>
<td>2</td>
<td></td>
<td>Planetree (Platanus Hispanica), lime tree (Tilia)</td>
<td>Iovan et al. 2008</td>
</tr>
<tr>
<td>CIR Sensor</td>
<td>2004</td>
<td></td>
<td>SVM</td>
<td>6</td>
<td></td>
<td>Platanus, Sophora, Tilia, Celtis, Pinus, Cupressus</td>
<td>Iovan et al. 2014</td>
</tr>
<tr>
<td>QuickBird</td>
<td>2007</td>
<td></td>
<td>Segmentation</td>
<td>2</td>
<td>93.72</td>
<td>Forest, grassland, thick grassland</td>
<td>Li et al. 2010</td>
</tr>
<tr>
<td>QuickBird</td>
<td></td>
<td></td>
<td>&amp; fuzzy multi-</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>thresholds classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IKONOS &amp; WV2</td>
<td>IKO: April 2006 /WV2: May 2011</td>
<td></td>
<td>LDA/ regression trees</td>
<td>7</td>
<td>16% to 18%</td>
<td>Sand live oak, laurel oak, live oak, pine, palm, camphor, magnolia</td>
<td>Pu and Landry 2012</td>
</tr>
<tr>
<td>Aerial &amp;</td>
<td>Aerial: 2006 &amp;</td>
<td></td>
<td>OB &gt; PB</td>
<td>1</td>
<td></td>
<td></td>
<td>Shouse et al. 2013</td>
</tr>
<tr>
<td>LandSat TM 5</td>
<td>LandSat: 2005 &amp; 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Bush honeysuckle (Lonicera maackii)</td>
<td></td>
</tr>
<tr>
<td>RapidEye</td>
<td>2009, LiDAR: 2007</td>
<td>X</td>
<td>SVM</td>
<td>8</td>
<td></td>
<td>Pinus, Aesculus, Platanus, Tilia, Acer, Populus, Fagus, Quercus</td>
<td>Tigges et al. 2013</td>
</tr>
<tr>
<td>QuickBird</td>
<td>2008</td>
<td>X</td>
<td>SMA</td>
<td>2</td>
<td></td>
<td>Evergreen and deciduous species</td>
<td>Tooke et al. 2009</td>
</tr>
<tr>
<td>QuickBird</td>
<td>March 2007</td>
<td>X</td>
<td>SMA / DT</td>
<td>2</td>
<td></td>
<td>Evergreen and deciduous species</td>
<td>Tooke et al. 2009</td>
</tr>
<tr>
<td>Sensor</td>
<td>Acquisition date</td>
<td>LiDAR</td>
<td>Classification algorithm</td>
<td>Species no.</td>
<td>Overall Acc.</td>
<td>Tree species</td>
<td>Tree counting</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>------------------</td>
<td>-------</td>
<td>--------------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Hyperspectral, LiDAR (Terra Remote Sensing) TRSI</td>
<td>2008</td>
<td>X</td>
<td>ANN / AGFLVQ / SAM</td>
<td>20</td>
<td>AGFLVQ: 68.8% / SAM: 39.95</td>
<td>American elm, hackberry, pecan, eastern red cedar, Shumard red oak, tree-of-heaven, cedar elm, green ash, red mulberry, chinaberry, gum bumelia, baldcypress, cherry laurel, boxelder, post oak, live oak, bur oak, cottonwood, crapemyrtle, black willow</td>
<td>Individual detection</td>
</tr>
</tbody>
</table>