Research Note / Note de recherche

Riparian vegetation classification from airborne laser scanning data with an emphasis on cottonwood trees

A. Farid, D. Rautenkranz, D.C. Goodrich, S.E. Marsh, and S. Sorooshian

Abstract. The high point density of airborne laser mapping systems enables achieving a detailed description of geographic objects and the terrain. Growing experience indicates, however, that extracting useful information directly from the data can be difficult. In this study, small-footprint lidar data were used to differentiate between young, mature, and old cottonwood trees in the San Pedro River Basin near Benson, Arizona, USA. The lidar data were acquired in June 2003, using the Optech Incorporated ALTM 1233 (Optech Incorporated, Toronto, Ont.), during flyovers conducted at an altitude of 750 m. The lidar data were preprocessed to create a two-band image of the study site: a high-accuracy canopy altitude model band, and a near-infrared intensity band. These lidar-derived images provided the basis for supervised classification of cottonwood age categories, using a maximum likelihood algorithm. The results of classification illustrate the potential of airborne lidar data to differentiate age classes of cottonwood trees for riparian areas quickly and accurately.

Résumé. La densité élevée de points des systèmes de cartographie laser aéroportés permet la description détaillée des objets géographiques et du terrain. L'expertise croissante dans ce domaine indique toutefois qu'il peut être difficile d'extraire l'information utile directement des données. Dans cette étude, des données lidar à petite empreinte ont été utilisées pour différencier les peupliers deltoïdes jeunes, matures et âgés dans le bassin de la rivière San Pedro, près de Benson, Arizona, États-Unis. Les données lidar ont été acquises en juin 2003, à l'aide d'un appareil Optech ALTM 1233 (Optech Incorporated, Toronto, Ont.), au cours de survols réalisés à une altitude de 750 m. Les données lidar ont été prétraitées pour créer une image à deux bandes du site d'étude : une bande haute précision du modèle d'altitude du couvert et une bande d'intensité dans le proche infrarouge. Ces images dérivées des données lidar ont servi de base pour la classification dirigée des catégories d'âge du peuplier deltoïde utilisant l'algorithme du maximum de vraisemblance. Les résultats de la classification montrent le potentiel des données lidar aéroportées dans la différenciation, de façon rapide et précise, des classes d'âge du peuplier deltoïde dans les zones riveraines.
[Traduit par la Rédaction]

Introduction

Lidar has emerged in recent years as a leading technology for the extraction of information about physical surfaces. The ever-increasing point density of current airborne systems allows a detailed description of the surveyed surfaces to be achieved and provides a wealth of information on physical objects and the terrain.

Previous work has used lidar data in two principal ways, by classifying the data (*i*) as terrain and nonterrain points, and (*ii*) as features such as trees or buildings. Examples of the first approach include the work of Kraus and Pfeifer (1998), who used an iterative linear prediction scheme for removing vegetation points in forested areas, and that of Vosselman (2000), who used gradient-based techniques to separate building points from terrain points. Using the second approach, Axelsson (1999) presented algorithms for filtering and classification of data points into terrain, buildings, and electrical power lines using multiple returns of lidar data and the intensity return. Song et al. (2002) focused on assessing separation of different land cover types, such as trees, grass,

roads, and roofs, based on interpolated intensity data, using three different interpolation techniques. Filin (2004) proposed a surface clustering technique for identifying regions in lidar data that exhibit homogeneity in a certain feature space, using attributes of position, tangent plane, and relative height difference for every point. The surfaces were categorized as high vegetation, low vegetation, smooth surfaces, and planar surfaces. Most of these previous studies involving

A. Farid.[1] Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA.

D. Rautenkranz and S.E. Marsh. Arizona Remote Sensing Center, Office of Arid Land Studies, University of Arizona, Tucson, AZ 85719, USA.

D.C. Goodrich. Southwest Watershed Research Center, Agricultural Research Service, US Department of Agriculture (USDA–ARS–SWRC), Tucson, AZ 85719, USA.

S. Sorooshian. Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA.

[1]Corresponding author (e-mail: farid@hwr.arizona.edu).

classification of lidar data have focused on the use of unsupervised algorithms in a small number of classes, resulting in generalized classifications.

This research represents the first evaluation of the use of lidar data for age discrimination of riparian vegetation. We employed a supervised classification technique, the maximum likelihood algorithm, for differentiating age classes of cottonwood trees by using small-footprint lidar data for riparian areas. Riparian cottonwood trees use water in proportion to their age and are especially large users of water in floodplains along rivers in semiarid environments. More accurate quantification of riparian water use is required to manage basin water resources to maintain the economic, social, and ecological viability of these areas and ensure water for the growing human population in the basin.

Study area

The study was conducted along a reach of the San Pedro River (Escalante study site; 31°51′N, 110°13′W; 1110 m elevation) within the San Pedro Riparian National Conservation Area (SPRNCA) in southeastern Arizona, USA. The study site is 1.2 km long north to south, 1.4 km wide east to west, and relatively flat. The overstory is dominated by riparian forest vegetation, consisting of cottonwood (*Populus fremontii*) and mesquite (*Prosopis velutina*) as dominant and subdominant overstory species, respectively. The study area is populated by young-to-old dense cottonwood stands. Patches of cottonwood riparian forest are located along the stream channel.

Methods

The Optech Incorporated ALTM 1233 (Optech Incorporated, Toronto, Ont.) was used to survey the study site on 6 June 2003. Characteristics of the ALTM 1233 include a scanning frequency of 28 Hz, a scan angle of ±20°, a collection mode of first and last returns, and intensity of returns from a 1064 nm laser. Cross-track point spacing is 0.9 m, forward point spacing is 2.1 m, and footprint size is approximately 15 cm. The data were preprocessed into first and last returns. The attributes of

any given laser return include not only x, y, and z coordinate data but also an intensity return value. The point data in first-return hits were interpolated using a kriging technique to produce a canopy altitude model (CAM) with a 0.5 m pixel grid. Additionally, the intensity return data and its x and y coordinates for both first and last hits were interpolated to the same 0.5 m regular grid corresponding to the CAM, thereby creating a near-infrared intensity image. **Figure 1** illustrates the CAM and the corresponding intensity image for the study site. We performed a supervised classification of the two-band lidar image (altitude and intensity images) using the maximum likelihood algorithm. Supervised classification is particularly applicable to a study site such as ours, in which there are a limited number of land cover types, the researchers are familiar with the area, the geographic extent is reasonably small, and ground-truth data can be obtained. Ground-validation data were collected from July 2004 to March 2005. Three different ages of cottonwood trees were included in the field sampling: young cottonwoods (less than 15 years), mature cottonwoods (16–50 years), and old cottonwoods (greater than 50 years). Stem diameters at breast height (dbh; diameter measured 1.37 m above the ground) were measured with a diameter tape and recorded to the nearest millimetre to discriminate between young, mature, and old cottonwood patches, based on river-specific equations that relate dbh to tree age (Stromberg, 1998). We selected training sites representative of the main land cover types, from which class signatures were generated. Signatures were generated for different age classes of cottonwood trees, mesquite, saltcedar, dry stream channel, and open ground categories. Training sites were chosen by visual inspection of the lidar image, using expert knowledge of the study site, to identify a region of signal purity (pixel uniformity) for each category. Training sites for cottonwood are sets of trees visually selected in the image from the three age classes. Selection criteria for trees used for signature generation are (*i*) unambiguous identification in the lidar images, and (*ii*) clear separation of signatures between classes. To validate the image classification, we performed an accuracy assessment, in which actual land cover, as determined by field identification, was compared with classes for the corresponding areas assigned by

Figure 1. Spatial patterns of (a) canopy altitude model and (b) near-infrared (1064 nm) intensity for the study site.
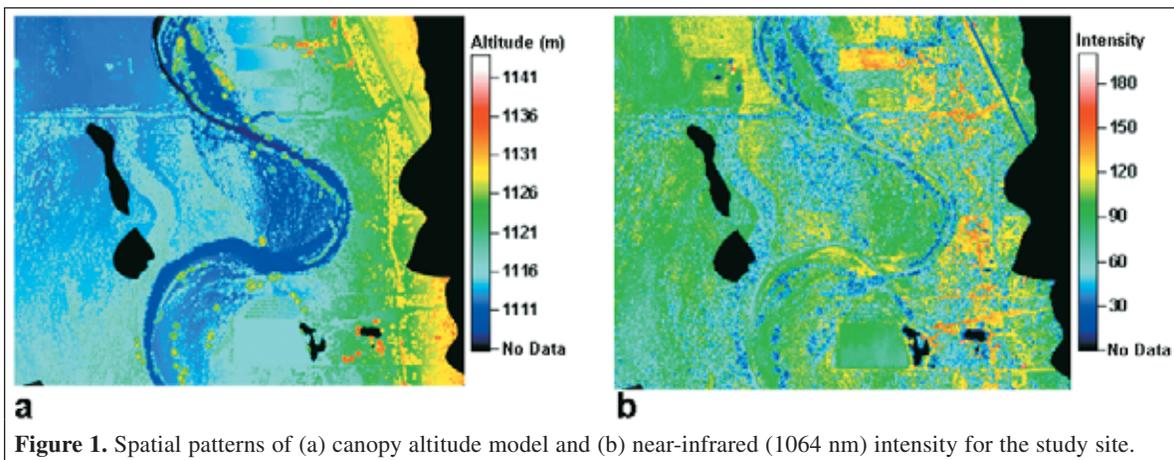
the maximum likelihood classification. Two different methods were used for field ground truth: (*i*) four differentially corrected global positioning system (GPS) points were acquired at the corners of a square centered on each cottonwood, from which we identified each cottonwood in the classified image; and (*ii*) for non-cottonwood classes (stream channel, open ground, saltcedar, and mesquite), we generated a random sample of 20 points for each class. Ground-truth data were obtained at the GPS coordinates for each random point. We calculated the classification error matrix using these ground-truth results as our reference data. Because there are low numbers of young and old cottonwoods, all field-identified young and old trees were used both in signature generation and in the accuracy assessment. Twenty of the 40 mature cottonwoods were used in signature generation. All 40 mature trees were used in the accuracy assessment. For the other categories there is no overlap in regions used for signature generation and accuracy assessment. Users' and producers' accuracy and Cohen's kappa were computed.

Results and discussion

Our initial classification included young, mature, and old cottonwood classes; mesquite; saltcedar; stream channel; and open ground. The objective of our study was to determine the feasibility of identifying cottonwood age classes from lidar data. The final classified image is shown in **Figure 2**. Accuracy assessment indicated that 78% of the two-band lidar image was correctly classified. The classification resulted in a Cohen's kappa of 0.73. If only cottonwood classes are considered, accuracy assessment gives a Cohen's kappa of 0.44 and an overall accuracy of 68%. The error matrix used for accuracy assessment including all categories is given in **Table 1**. Accuracy assessment results are presented in **Table 2**.

Although the overall classification accuracy was reasonably good, it can be seen from **Table 1** that discrimination of young cottonwoods from mature cottonwoods failed for almost half the young trees. Overlap in height between young and mature cottonwoods may be a contributing factor in the confusion between young and mature cottonwoods. Another confounding factor is the absence of a significant difference in lidar mean intensity between young and mature trees. Leaves and branches of young and mature cottonwoods have almost the same color and brightness and have very similar surface reflectance in the near-infrared, giving rise to similar lidar intensity returns. All of the trees classified in the image as young cottonwoods corresponded to trees identified by fieldwork as young cottonwoods. We failed to achieve accurate separation of mature from old cottonwoods in the image for approximately one third of the mature trees because of overlap between the crowns of mature and old cottonwoods; 73% of mature cottonwoods identified as such by fieldwork were classified as mature in the image. Successful separation of old from young and mature cottonwoods was facilitated by the strong differences in crown shape. Old cottonwoods have conical and flat-topped crowns, whereas young cottonwoods have narrow and upright crowns. The crowns of these old cottonwoods are isolated from each other, and also there is a large difference between the elevations of each tree crown and the surrounding understory vegetation, allowing the maximum likelihood algorithm to classify them more accurately. Of the non-cottonwood classes, saltcedar proved the most difficult to classify correctly. Only 55% of saltcedar in the field was classified as such in the image. The rest were classified as either mesquite (3 of 20 trees) or open ground (6 of 20 trees). The confusion between saltcedar trees and open ground may result from saltcedar with sparse foliage, allowing penetration through the canopy to the ground by lidar pulses. Overall, we

Figure 2. Classified lidar image, showing three cottonwood age classes, mesquite, saltcedar, dry stream channel, and open ground.
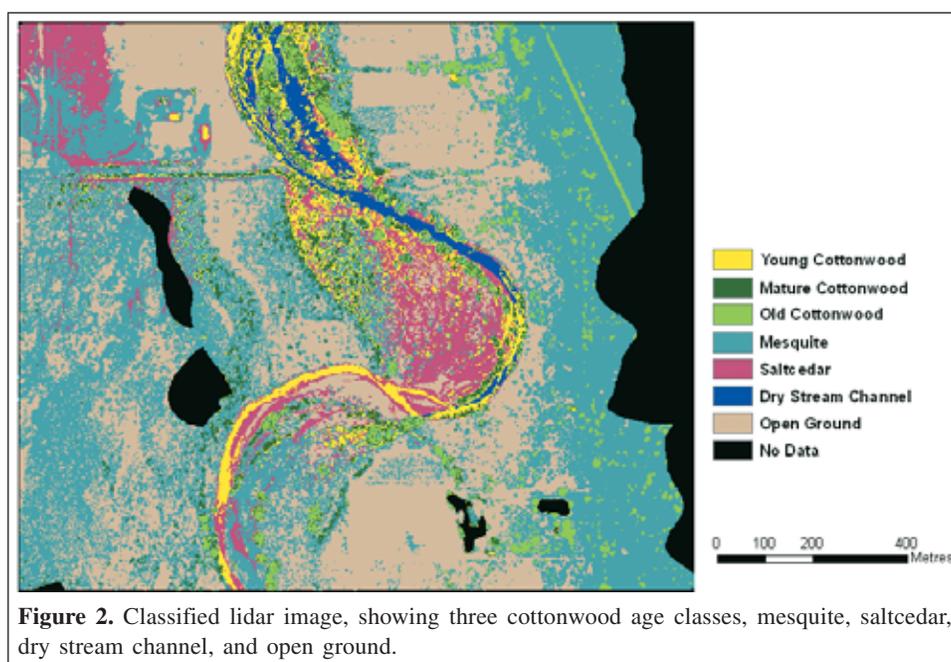
Table 1. Error matrix for maximum likelihood classification of the two-band lidar image.

Classified data	Old cottonwood	Mature cottonwood	Young cottonwood	Mesquite	Saltcedar	Dry stream channel	Open ground	Row total
Old cottonwood	8	11	2	0	0	0	0	21
Mature cottonwood	1	29	7	0	0	0	0	37
Young cottonwood	0	0	7	0	0	0	0	7
Mesquite	0	0	1	57	3	0	1	62
Saltcedar	0	0	0	1	11	0	2	14
Dry stream channel	0	0	0	0	0	16	1	17
Open ground	0	0	0	0	6	4	16	26
Column total	9	40	17	58	20	20	20	184

Table 2. Accuracy assessment result for maximum likelihood classification of the two-band lidar image.

Class name	Reference total	Classified total	Number correct	Producers accuracy (%)	Users accuracy (%)
Old cottonwood	9	21	8	89	38
Mature cottonwood	40	37	29	73	78
Young cottonwood	17	7	7	41	100
Mesquite	58	62	57	98	92
Saltcedar	20	14	11	55	79
Dry stream channel	20	17	16	80	94
Open ground	20	26	16	80	62
Total	184	184	144		

Note: Overall classification accuracy is 78%, and overall kappa statistic is 0.73.

did not find this to be a factor for cottonwoods. A major factor in the classification of non-cottonwoods as cottonwoods, especially as mature cottonwoods, is the topography of the study area. Topographic effects were not taken into account. For example, the assignment of a road northwest of the stream channel to the mature cottonwood class was caused by elevation change between the stream channel and the area to the northwest. Correction for topographic effects might be accomplished by including a digital elevation model based on ground returns as a third image band in the analysis.

Conclusions

This research employed a supervised classification technique, the maximum likelihood algorithm, for differentiating age classes of cottonwood trees in a riparian area using small-footprint airborne lidar data. The accuracy assessment gave an overall classification accuracy of 78% and an overall kappa statistic of 0.73. Overall, classification results illustrate the potential of airborne lidar data to differentiate age classes of cottonwood trees for riparian areas quickly and quantitatively. Thus, identifying cottonwoods in the riparian forest and, more importantly, differentiating cottonwood age classes using lidar provide an important contribution to precision forest inventory and automated data processing for riparian forestry applications. Future research should consider fusing high spatial resolution multispectral or hyperspectral data and lidar data to improve classification results for species identification in riparian areas.

References

Axelsson, P. 1999. Processing of laser scanner data-algorithms and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 54, pp. 138–147.

Filin, S. 2004. Surface classification from airborne laser scanning data. *Computers & Geosciences*, Vol. 30, pp. 1033–1041.

Kraus, K., and Pfeifer, N. 1998. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 53, pp. 193–203.

Song, J.H., Han, S.H., Yu, K.Y., and Kim, Y.I. 2002. Assessing the possibility of land-cover classification using lidar intensity data. In *Photogrammetric Computer Vision PCV'02: Proceedings of the ISPRS Technical Commission III Symposium*, 9–13 September 2002, Graz, Austria. Part B, p. 259.

Stromberg, J.C. 1998. Dynamics of Fremont cottonwood (*Populus fremontii*) and saltcedar (*Tamarix chinensis*) populations along the San Pedro River, Arizona. *Journal of Arid Environments*, Vol. 40, pp. 133–155.

Vosselman, G. 2000. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXIII, Part B3, pp. 935–942.