Automatic Extraction and Classification of Vegetation Areas from High Resolution Images in Urban Areas

Corina Iovan[1,2], Didier Boldo[1], Matthieu Cord[2], and Mats Erikson[3]

[1] Institut Géographique National, Laboratoire MATIS
2/4 avenue Pasteur, 94165 Saint-Mandé Cedex, France
{Corina.Iovan, Didier.Boldo}@ign.fr
[2] Université Pierre et Marie Curie, Laboratoire d'Informatique de Paris VI
8 rue du Capitaine Scott, 75015 Paris, France
Matthieu.Cord@lip6.fr
[3] ARIANA Research Group, CNRS/INRIA/UNSA
2004, route des Lucioles - BP 93, 06902 Sophia Antipolis Cedex, France
Mats.Eriksson@sophia.inria.fr

Abstract. This paper presents a complete high resolution aerial-images processing workflow to detect and characterize vegetation structures in high density urban areas. We present a hierarchical strategy to extract, analyze and delineate vegetation areas according to their height. To detect urban vegetation areas, we develop two methods, one using spectral indices and the second one based on a Support Vector Machines (SVM) classifier. Once vegetation areas detected, we differentiate lawns from treed areas by computing a texture operator on the Digital Surface Model (DSM). A robust region growing method based on the DSM is proposed for an accurate delineation of tree crowns. Delineation results are compared to results obtained by a Random Walk region growing technique for tree crown delineation. We evaluate the accuracy of the tree crown delineation results to a reference manual delineation. Results obtained are discussed and the influential factors are put forward.

1 Introduction

Automatic 3D reconstruction of urban areas is an active research topic in distinct application areas and an issue of primary importance in fields such as urban planning, disaster management or telecommunications planning. Significant progress has been made in recent years concerning the automatic reconstruction of man-made objects or environments from multiple aerial images [1]. Yet, a lot of challenge concerning the modelling of other objects present on the terrain surface, such as trees, shrubs, hedges or lawns still exists. An accurate automatic reconstruction of such types of vegetation areas is a real challenge due to their complex nature and to their intricate distribution between man-made objects in dense urban areas. Many researches deal with automatic tree crown delineation from aerial or satellite images. We can divide them into two classes: methods applied to forest stands and methods applied to urban environments.

B.K. Ersbøll and K.S. Pedersen (Eds.): SCIA 2007, LNCS 4522, pp. 858–867, 2007.

Several algorithms have been proposed for the segmentation of individual trees in forest stands. A first class uses local maxima information to estimate tree top position and the number of trunks [2][3]. A second class of methods exploits the shadows around the tree crowns to delineate their contour [4], such as valley-following algorithms [5] or region growing methods [6]. Other contour based methods use multi-scale analysis [7] or active contours [8] to delineate tree crowns. A third class of methods are object-based methods [9][10][11], modelling synthetic tree crown templates to find tree top positions.

Algorithms developed for the automatic extraction of tree crowns in urban environments firstly detect vegetation areas followed by a finer analysis of objects present therein. Depending on the input data, vegetation areas are detected either using vegetation responses in color infrared (CIR) images [12] or by computing surface roughness indicators on the DSM [13]. A finer analysis of treed areas is then performed and its goal ranges from simple tasks, such as estimating tree position [12][14], to more complex tasks such as tree species classification [6].

This study presents our approach for vegetation detection and segmentation in urban areas. A linear-kernel SVM classifier using a four dimensional radiometric feature vector is used to identify vegetation areas. Texture features computed on the DSM separate lawns from treed areas. A robust algorithm for tree crown delineation taking into account the trees height and shape characteristics is proposed to accurately delineate individual tree crowns.

The accuracy of the segmentation results is evaluated against a reference delineation and they are also compared to results obtained by a random walk tree crown delineation algorithm [6]. Results obtained using the proposed method are very promising and show their potential by improving delineation results obtained by the second method.

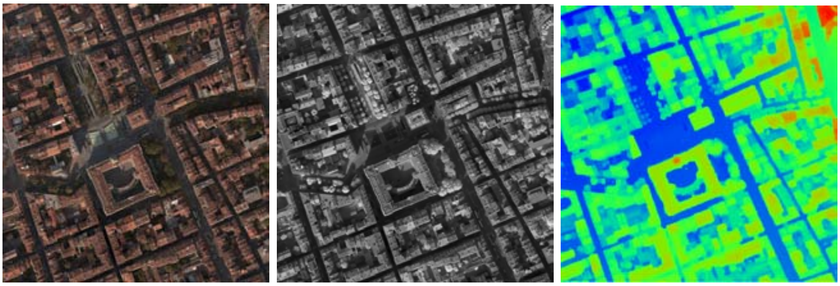
2 Study Area and Data

2.1 Study Area

The study area is located in the city of Marseille, situated in the south-east of France. Marseille's climate is Mediterranean, with a great variety of vegetation species. It's a dense urban area, with many greened and treed resting places, highly intermingled with buildings.

2.2 Data

In this study, tests were carried out on digital color infrared aerial images, taken in November 2004, with a ground resolution of 20cm per pixel. A DSM is derived from the stereoscopic aerial images using an algorithm based on a multiresolution implementation [15] of Cox and Roy's image matching algorithm [16] based on graph cuts. Figure 1 presents the input data for one of our study areas.



a) b) c)

Fig. 1. An aerial image of Marseille (France) representing a high density urban area, where 1 pixel corresponds to approximately 20cm a) RGB channels b) IR channel c) DSM for the same area

3 Methods

Our approach is based on a hierarchical strategy containing several steps: detection of vegetation areas, segmentation of vegetation types according to their height followed by individual tree crown delineation. In the following paragraphs we will describe the methods developed for each of these steps.

3.1 Radiometric Corrections

In a first step, the four channels of the raw images are radiometrically corrected. Radiometric corrections address variations in the pixel intensities that are not caused by the object or scene being scanned. Due to the sea vicinity, haze is often perturbing the signal penetration. Atmospheric haze reduction is a simple method that assumes that some image pixels should have a reflectance of zero. Actual values of zero pixels result from atmospheric scattering. Haze correction consists in subtracting the histogram offset from all pixels in a specific band. The result of the atmospheric correction is depicted in Fig. 2.

Fig. 2. RGB image representing downtown of Marseille before (left) and after (right) the atmospheric haze reduction

3.2 Vegetation Detection

Two methods were developed to identify vegetation areas. The first one is an unsupervised classification method based on different spectral indices. The second one is a supervised classification method using a linear-kernel Support Vector Machines (SVM) classifier.

Unsupervised Classification

The unsupervised classification method uses several spectral indices to identify vegetation areas. The first index computed for each pixel in our images is the NDVI (Normalised Difference Vegetation Index)[17]. It allows the creation of a gray-level image, the NDVI image (presented in Fig. 3 b)), by computing for each pixel the NDVI index, according to (1)

$$NDVI = \frac{\varphi_{IR} - \varphi_R}{\varphi_{IR} + \varphi_R} \tag{1}$$

where φ_{IR} and φ_R are the values of the pixels respectively in the infrared and the red band. This index highlights areas with a higher reflectance in the infrared band than in the red band (i.e. vegetation). Applying a threshold on the NDVI image gives a coarse segmentation of the urban scene in vegetation areas and non-vegetation areas. As there are also other materials present in an urban environment with a high reflectance in the infrared band, we refine vegetation classification results using a second spectral index computed for each pixel, according to (2)

$$SI = \frac{\varphi_R - \varphi_B}{\varphi_R + \varphi_B} \tag{2}$$

This is the saturation index (SI) [18] and the gray-level image obtained for this index for each pixel is presented in Fig. 3 c). The images obtained with these two spectral indices are binarized and used together to create the vegetation mask. The result presented in Fig. 3 d) emphasizes all vegetation areas.

Supervised Classification

Although the vegetation detection method based on spectral indices gives satisfying results, it is a method highly dependent on the spectral characteristics of the

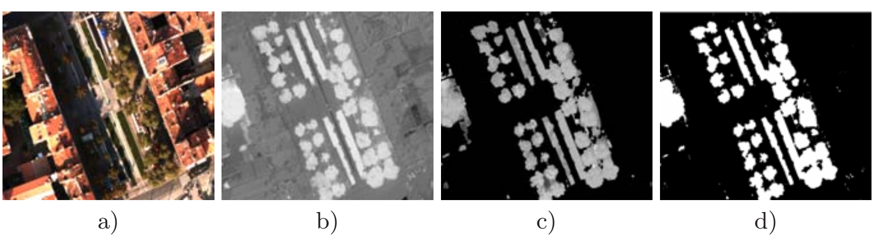
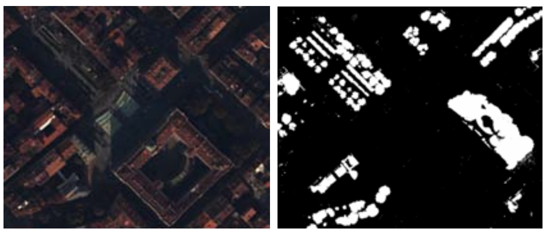


Fig. 3. Spectral indices used for vegetation detection a) RGB input image b) NDVI image c) SI image d) Vegetation mask obtained with the NDVI and SI spectral indices

data present in the study area. Our goal was to develop a method which performs well in any case. Therefore, we used a reliable supervised classification method based on SVM's. For all pixels in the training dataset, the feature vector contains four characteristics, namely, the reflectance values of each pixel in the infrared, red, green and blue bands. The choice of a linear-kernel for the classifier was motivated by the fact that the spectral indices we used in the first method are linear combinations of the image's channels and they perform well for distinguishing between vegetation and non vegetation areas. Therefore, instead of deciding where the separator between the two classes is, by combining different spectral indices, we decided to leave this task to the SVM and thus exploit its capacities in finding the optimal linear separator. Figure 4 b) presents the vegetation mask obtained by the SVM classifier for the test area presented in Fig. 4 a).



a) b)

Fig. 4. Vegetation detection based on linear-kernel SVM classification a)Input image b)Vegetation mask

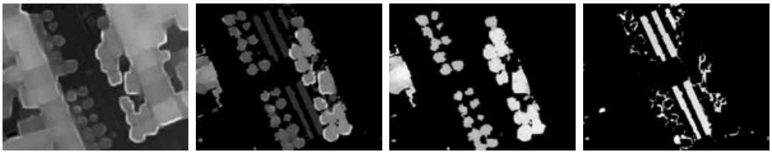
Both methods presented in this section give pertinent results for vegetation detection but in the following of our study we use the vegetation mask obtained by the unsupervised classification.

3.3 Grass/Tree Segmentation

Once the regions of interest identified, i.e. vegetation areas, we proceed to a finer level of analysis of these vegetation structures, by performing texture analysis on the corresponding areas of the DSM. The goal of this second step is to differentiate lawns from trees. In a CIR image, grasses are characterized by ranges of coloration and texture. In the DSM, treed areas are characterized by a higher gray level variance compared to lawn areas. The method we developed to separate grass from trees takes into account this property by computing the local variance on the DSM. The resulting image is thresholded to obtain masks for grass and treed areas. Figure 5 shows the results obtained for grass/treed area separation for the test area depicted in Fig. 3 a).

3.4 Tree Crown Delineation

To separate tree crowns from each other, we developed a region growing method taking into account the treed areas previously detected. All region growing (RG)



a) b) c) d)

Fig. 5. Differentiation between grass and treed areas a)Local variance computed on the DSM b)Vegetation areas on the DSM c)Treed areas d)Lawns

methods need seed points for each region before growing. The performance of these algorithms is highly dependent on the number of seed points initialising each region. The ideal case is to have one seed point per region.

Seed points. We aimed at achieving this goal: having one seed point for each tree. This information could be the information concerning the top of each tree. We use the DSM to estimate tree tops. To reduce the number of possible candidates for a tree top, we use a Gaussian filter as a smoothing filter for the DSM, with an empirical determined mask, approaching the average size of the trees in the image. To determine tree tops, we evaluate the maximum height of the trees present in the DSM and we consider all points having the same height as tree tops. In the first iteration we obtain points corresponding to the highest trees in the stand. Therefore, we iteratively decrease the analysis altitude, h. At each step, we analyse all points at higher heights than h and detect a new seed when a new region appears and it doesn′t touch pixels previously labelized as seeds. A graphical ilustration of this algorithm is presented in Fig. 6.

Region growing. Starting from the previously determined tree tops, tree crown borders are obtained by a region growing approach based on geometric criteria of the trees. We used the DSM to evaluate the height of all points neighbour to a seed point. All pixels corresponding to a lower height point are aggregated to the region corresponding to the closest (in terms of height) tree top. The results

a) b) c)

Fig. 6. Detecting tree tops from the DSM a)3D view of the DSM: all points higher than the analysis altitude h are evaluated for tree top estimation b)2D view of the 30th iteration c)Seed points detected after the final iteration: we can notice that we obtain one seed region for each tree

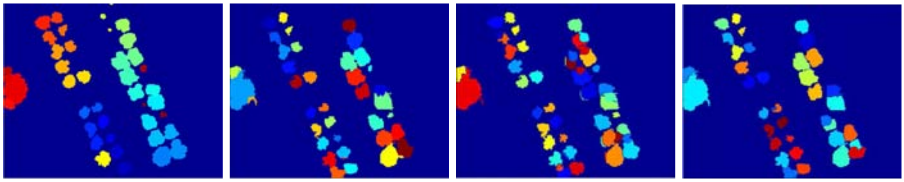

a) b) c) d)

Fig. 7. Tree crown delineation results a)Reference delineation of tree crowns b)Segmentation results for the Height-RG (HRG) method c)Segmentation results for the (RW) region growing method ($RWRG$) d) Tree crown delineation results for the ($RWRG$) method applied to height data ($H - RWRG$)

of this algorithm for tree crown delineation for the test area presented in Fig. 3 a) can be seen in Fig. 7 b).

We compared the method we developed with a random-walk (RW) region growing method, described in [6]. This method, briefly described below, is applied on an artificial image containing for each channel the DSM, instead of simply applying it on the colour image.

Seed points. To find seed points, the first band of the input image (DSM) is thresholded and a distance transform is performed on the resulting image. This distance image is smoothed with a Gaussian filter and local maxima on this image represent the seeds for the region growing algorithm.

Region growing. Each of the previously detected seeds is grown to become a region. A priority queue is established for the order the seeds are processed: seeds which are border pixels to a region are processed sooner than seeds which are not border pixels. The higher the value the border pixel is the sooner it will be connected to a region. This value is taken from a new image, a random walk image which is obtained from the original image by simulating random walks for each seed point. The value of each pixel represents the number of times the simulated particles have reached the pixel. A series of constraints decide on the rapidity of a pixel aggregation to a region. Further details can be found in [6]. Figure 7 c) presents the tree crown delineation results of this method for the same test area as the one presented in Fig. 3 a).

We also combined the two methods by using seed points found by the first method with the region growing method of the second method. The results of tree crown delineation for this case can be seen in Fig. 7 d).

4 Results

This section evaluates the results of the tree crown segmentation methods presented in the first part of this article and depicted in Fig. 7 b) - d). All segmentation results are compared to the reference manual delineation of the trees presented in Fig. 7 a). This manual delineation has been generated by an experienced photo interpreter by means of stereo restitution. It contains, for all trees visible in the CIR images, the exact delineation of tree crowns which will be considered as reference delineation in the following of the evaluation.

The accuracy assessment results are presented in table 1, where *Nt* denotes the number of trees and *Ratio*, the percentage of the total number of trees computed using the total number of trees in the stand.

Table 1. Comparison between the reference delineation of tree crowns and results obtained for tree crown delineation by the three methods

	Height Region Growing Method		RW Region Growing Method		Height-RW Region Growing Method	
	Nt	Ratio	Nt	Ratio	Nt	Ratio
Trees correctly segmented	32	78.0	23	56.1	30	73.1
Trees over-segmented	1	2.4	11	26.8	3	7.3
Trees under-segmented	4	9.7	4	9.7	4	9.7
Trees omitted	4	9.7	4	9.7	4	9.7
Total number of trees in the stand	41		41		41	
Total number of detected trees	37		51		37	

4.1 Evaluation Measures

The approach used for the evaluation is similar to the one presented in [14]. A statistical analysis is first performed taking into consideration the total number of trees in the ground truth and the omission (omitted trees) and commission errors (segments not associated with a tree). We take into consideration the following cases for the spatial analysis of the segmentation: pure segments, over-segmented trees, under-segmented trees. Pure segments correspond to correctly identified trees. We consider that a segment is 100% pure if it corresponds to one and only one segment in the ground truth and vice versa, with an overlap area greater than 80%. Over-segmented trees correspond to the case when more than one segment is associated with the ground truth delineation. Under segmented trees correspond to segments which include a significant part ($> 10\%$) of more than one tree.

4.2 Disscusion

The two methods for vegetation/non−vegetation classification give very good results. Surface classification rates are high for the two methods, from 87.5% for the spectral index based method to 98.5% for the SVM classification method.

Concerning the grass/lawn segmentation, the results were evaluated using a manual delineation and the results are very promising. More than 97% of the grass surface in the reference delineation was correctly classified as lawn.

Regarding the tree crown delineation, we notice that all the three methods previously described have detected most of the large trees. The H−RWRG and the HRG methods detect the same number of trees in the stand, and this is due to the fact that the same seeds are used for the region growing step. Omitted trees have in fact a low height, and due to the gaussian blurring of the DSM before finding seed points, these trees are not present in the DSM when the region growing part starts. The number of correctly segmented trees is higher for the HRG method, and this is due to the way the seeds are grown. The results of the RWRG method are improved by 17% when it uses one seed for region. These results show the good potential of the proposed method to find one seed for each tree.

5 Conclusion

Three region growing methods for tree crown delineation have been evaluated and show the capacity of having a realistic geometric description of tree crowns in urban areas. Our ongoing research deals with the improvement of the height−based tree crown delineation method by including information from the CIR images in the segmentation step. Extra information (tree crown diameter, height) can be extracted for each tree and it can be used for a 3D modelisation of trees. We will also consider evaluating the performances of the proposed method for tree crown delineation on laser DSM′s.

Acknowledgments. The fourth author thanks the Arc Mode De Vie Project at INRIA for financial support.

References

1. Taillandier, F., Vallet, B.: Fitting constrained 3d models in multiple aerial images. In: British Machine Vision Conference, Oxford, U.K. (August 2005)
2. Pinz, A.: Tree isolation and species classification. In: Hill, D., Leckie, D., eds.: Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, British Columbia, Canada pp. 127–139 (February 1998)
3. Wulder, M., White, J., Niemann, K., Nelson, T.: Comparison of airborne and satellite high spatial resolution data for the identifcation of individual trees with local maxima filtering. International Journal of Remote Sensing 25(11), 2225–2232 (2004)
4. Gougeon, F., Leckie, D.: Individual tree crown image analysis - a step towards precision forestry. In: Proc. of the First International Precision Forestry Symposium, Seattle, U.S (June 2001)
5. Gougeon, F.: A system for individual tree crown classification of conifer stands at high spatial resolution. In: Proc. of the 17th Cananadian Symposium on Remote Sensing, Saskatchewan, Canada pp. 635–642 (June 1995)
6. Erikson, M.: Segmentation and Classification of Individual Tree Crowns. PhD thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden (2004)
7. Brandtberg, T., Walter, F.: Automatic delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis. Machine Vision and Applications 11(2), 64–73 (1998)

8. Horvath, P., Jermyn, I., Kato, Z., Zerubia, J.: A higher-order active contour model for tree detection. In: Proc. International Conference on Pattern Recognition (ICPR), Hong Kong (August (2006)
9. Pollock, R.: The Automatic Recognition of Individual Trees in Aerial Images of Forests based upon a Synthetic Tree Crown Image Model. PhD thesis, University of BC, Department of Computer Science, Vancouver, Canada (1996)
10. Larsen, M.: Crown modelling to find tree top positions in aerial photographs. In: Proc. of the Third Internationl Airborne Remote Sensing Conference and Exhibition. vol. 2, pp.428–435 (1997)
11. Perrin, G., Descombes, X., Zerubia, J.: 2d and 3d vegetation resource parameters assessment using marked point processes. In: Proc. International Conference on Pattern Recognition (ICPR). pp. 1–4 (2006)
12. Straub, B.M.: Automatic extraction of trees from aerial images and surface models. In: ISPRS Conference Photogrammetric Image Analysis (PIA). Volume XXXIV, Part 3/W834., Munich, Germany (September 2003)
13. Rottensteiner, F., Trinder, J., Clode, S., Kubik, K.: Using the dempster-shafer method for the fusion of lidar data and multi-spectral images for building detection. Information Fusion 6(4), 283–300 (2005)
14. Mei, C., Durrieu, S.: Tree crown delineation from digital elevation models and high resolution imagery. In: Proc. of the ISPRS Workshop 'Laser scanners for Forest and Landscape Assement'. vol. 36., Fribourg, Germany (October 2004)
15. Pierrot-Deseilligny, M., Paparoditis, N.: A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery. In: Proc. of the ISPRS Conference Topographic Mapping From Space (With Special Emphasis on Small Satellites), Ankara, Turkey, ISPRS (February 2006)
16. Roy, S., Cox, I.: A maximum-flow formulation of the n-camera stereo correspondence problem. In: Proc. of the IEEE International Conference on Computer Vision, Bombay, India pp. 492–499 (January 1998)
17. Gong, P., Pu, R., Biging, G.S., Larrieu, M.R.: Estimation of forest leaf area index using vegetation indices derived from hyperion hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 41(6), 1355–1362 (2003)
18. Mathieu, R., Pouget, M., Cervelle, B., Escadafal, R.: Relationships between satellite based radiometric indices simulated using laboratory reflectance data and typic soil colour of an arid environment. Remote Sensing of Environment 66, 17–28 (1998)