SUPER-RESOLUTION MAPPING FOR EXTRACTION OF URBAN TREE CROWN OBJECTS FROM VHR SATELLITE IMAGES

Valentyn A. Tolpekin, Juan Pablo Ardila, Wietske Bijker

Faculty of Geo-Information Science and Earth Observation (ITC) University of Twente 7500 AA Enschede, The Netherlands tolpekin@itc.nl, ardila14710@itc.nl, bijker@itc.nl http://www.boomenbeeld.nl

KEY WORDS: Super resolution mapping, tree crown identification, markov random field, object detection, urban trees

ABSTRACT:

Extraction of individual tree crown objects in urban areas using available VHR space-borne imaging systems remains a challenging image processing task mainly because of two factors: the limited spectral information offered by the sensors and their limited spatial resolution. The spectral information of available VHR satellite sensors is not sufficient to discriminate tree crowns from other land cover classes such as grassland and shrubs using spectral pixel-based classification techniques. This can be solved by using a contextual classification approach, while the limitation of spatial resolution can be reduced by using Super resolution mapping (SRM). In this paper we extend the contextual Markov Random Field (MRF) based SRM method developed earlier for multispectral images to include the panchromatic band with higher spatial resolution. We apply this method for extraction of tree crown objects from Quickbird images, setting the spatial resolution of super-resolved map to that of the panchromatic band, 0.6 m. We apply the proposed method for identification of tree crown objects in a residential area in the Netherlands. We find that incorporation of panchromatic band leads to improved object identification. Our object based accuracy assessment indicate that the proposed method identified 73% of the trees in the study area, with some commission errors in tree areas with understory vegetation.

1 INTRODUCTION

With the increasing availability of very high resolution (VHR) visible-infrared satellite imagery, extent of green spaces in urban areas can be mapped cost-effectively with existing image classification methods. Some users, e.g. municipalities, need a product with finer spatial detail: the location of individual trees within the urban area. This information is the basis for their planning activities and allocation of resources related to conservation, planting and maintenance of urban trees (McHale et al., 2007). Currently, determination of location and crown size of individual trees is only possible with the expensive combination of optical imagery and elevation data derived from laser scanning or overlapping photographs, or by GPS measurements on the ground. Automated extraction of locations of individual trees from VHR imagery remains a challenge even for homogeneous environments like forest plantations (Gougeon and Leckie, 2006, Daliakopoulos et al., 2009, Hirschmugl et al., 2007).

Three fundamental factors prevent the effective identification of tree crown objects from modern VHR satellite imagery. Firstly, a comparison between the highest spatial resolution of infrared satellite imagery and the range of tree crown size in urban areas, reveals that we are dealing with a transition zone between the L and H-resolution scene models (Woodcock and Strahler, 1987), where the object is comprised by a single pixel in the former and few pixels in the latter. We show this in Table 1, where we estimate the number of image pixels representing a single tree for two VHR sensors. From the scene models, it follows that the classification accuracy is seriously affected in the L-resolution case. Secondly, the spectral separability between tree crown and other background surfaces such as shrubs and grasslands constitute serious limitations for spectral classifiers.

Crown diameter (m)	Ikonos	Quickbird
Small (3)	< 1	1
Medium (10)	5	14
Large (20)	19	54

Table 1: Number of pixels representing a tree crown object in multispectral modes of Ikonos and Quickbird images

Moreover, the spectral response of tree crowns in images is modified by sun illumination and shadows projected by tall buildings. Thirdly, as newer VHR sensors have a larger dynamic range capturing complex scene models, the within class spectral variance increases (Woodcock and Strahler, 1987), resulting in pixel-based classification outputs that lack spatial coherence and do not explicitly represent the objects of interest.

Alternatives to enhance the spatial details of satellite imagery include pan-sharpenning techniques producing a synthetic image combining the spectral characteristics of the multi-spectral bands with the spatial resolution of the panchromatic products. However, most existing pan-sharpening methods modify the spectral information and very often produce artifacts in the image (Du et al., 2007). Soft classifiers and sub-pixel methods have also been used to produce information on the class-fraction composition within individual pixels but the spatial distribution of these fractions within the pixel remains unknown. In this work we propose a super resolution mapping (SRM) algorithm based Markov Random fields (MRF) for identification of tree crown objects in VHR satellite images. Super resolution mapping produces classification maps at a finer resolution than that of the original input image, and is currently a subject of active research with several approaches (Atkinson, 2009). We adopt a contextual SRM method based on MRF as described in Tolpekin and Stein (Tolpekin and Stein, 2009). That method was proved to work well on synthetic normally distributed multispectral images. The novelty of this study is the incorporation of information from a panchromatic band and application of the method for identification of tree crown objects from real Remote Sensing data. We apply the method on a Quickbird image of a residential area in the city of Enschede, The Netherlands. This research was part of the Boom en Beeld (Tree and Image) project.

2 SUPER RESOLUTION MAPPING METHOD

We consider the classification of a multispectral remote sensing image y that consists of K spectral bands with spatial resolution R and pixel locations $b_i \in B$, where B is a $M_1 \times M_2$ pixel matrix. In addition we have a panchromatic image z with finer spatial resolution r < R. The superresolution map (SR map) c is defined on the set of pixel locations A and covers the same extent on the ground as yand z with spatial resolution r. The scale factor S = R/r is an integer for common VHR images. Hence each pixel b_i corresponds to the area on the ground covered by S^2 finer resolution pixels a_{jli} , $j = 1, \ldots, S^2$.

We further assume the existence of a multispectral image x defined on the set of pixels A with K spectral bands and fine spatial resolution r. Image x is not observed directly while images y and z are considered as spatial and spectral degraded observations of x respectively. Furthermore, we assume that each pixel in the image x can be assigned to a unique class: $c(a_{ji}) = \alpha$, with $\alpha \in 1, 2, ..., L$. The relationship between y and x, and z and x are established by a degradation model as:

$$y_k(b_i) = \frac{1}{S^2} \sum_{j}^{S^2} x_k(a_{j|i}), \ k = 1, \dots, K$$
(1)

$$z(a_{j|i}) = \frac{1}{K} \sum_{k}^{K} x_{k}(a_{j|i})$$
(2)

for $b_i \in B$ and $a_{j|i} \in A$. Note that the linear equations (1,2) cannot be solved for x because for K > 1 and S > 1 the number of equations $(M_1M_2 (K + S^2))$ is smaller than the number of unknown values $(M_1M_2KS^2)$. We therefore do not intend to estimate image x; instead we intend to find the SR map c that corresponds to the maximum a posteriori probability P(c|y, z) (MAP) solution for c given observed data y and z. According to the Bayes theorem P(c|y, z) can be computed from prior probability P(c) and C(z|c) as

$$P(c|y,z) \propto P(c)P(y|c)P(z|c)$$
(3)

Here we used an assumption that y and z are conditionally independent. We specify the respective probabilities by introducing energy functions:

$$P(c) = \frac{1}{A_1} \exp\left(-\frac{U(c)}{T}\right) \tag{4}$$

$$P(y|c) = \frac{1}{A_2} \exp\left(-\frac{U(y|c)}{T}\right)$$
(5)

$$P(z|c) = \frac{1}{A_3} \exp\left(-\frac{U(p|c)}{T}\right)$$
(6)

$$P(c|y,z) = \frac{1}{A_4} \exp\left(-\frac{U(c|y,z)}{T}\right)$$
(7)

where *T* is a constant called temperature, A_i , i = 1, ..., 4 are normalization constants independent of *c*; U(c), U(y|c), U(z|c) and U(c|y, z) are the prior, conditional and posterior energy functions, respectively. Rewriting the Bayes formula for energy functions and omitting terms independent of *c* we get

$$U(c|y,z) = \lambda U_c(c) + (1-\lambda) \left(\lambda_p U(z|c) + (1-\lambda_p) U(y|c) \right)$$
(8)

where $0 \le \lambda < 1$, is a parameter of our model which balances the contribution of prior and conditional energy functions in our model (Tolpekin and Stein, 2009)). In a similar way we introduce here a parameter $0 \le \lambda_p \le 1$ which balances contributions of the two conditional energy functions: based on panchromatic and multispectral observations, respectively.

2.1 Prior energy function

We assume that SR map is a Markov Random field (MRF) (Li, 2009). Due to Gibbs Random field(GRF)-MRF equivalence we can specify the MRF model by (4) with

$$U(c) = \sum_{i,j} U(c(a_{j|i})) = \sum_{i,j} \sum_{l \in N(a_{j|i})} w(a_l) I(c(a_{j|i}), c(a_l))$$
(9)

where $N(a_{j|i})$ is the neighbourhood system, $U(c(a_{j|i}))$ is the local contribution to the prior energy from pixel $c(a_{j|i})$, $w(a_l)$ is the weight of the contribution from neighbor pixel $a_l \in N(a_{j|i})$ and $I(\alpha, \beta)$ takes value 0 if $\alpha = \beta$ and 1 otherwise. This prior model gives preference to smooth SR map c and penalizes the occurrence of pixels with different class labels in the neighborhood system N. The weights $w(a_l)$ are chosen inversely proportional to the distance $d(a_l)$ between the central pixel $a_{j|i}$ and the pixel a_l :

$$w(a_l) \propto \frac{1}{d(a_{j|i}, a_l)} \tag{10}$$

2.2 Conditional energy functions

The likelihood energies consider the proximity of observed pixel values p and y to each land cover class. We model spectral values x of a class α with the normal distribution. Further, we assume that values x are spatially uncorrelated given their class association. In this case the spectral values in p and y are also normally distributed, see (1) and (2). The conditional term U(y|c) is defined as:

$$U(y|c) = \sum_{i} \frac{1}{2} \left[M(y(b_i), \mu_i, C_i) + \frac{1}{2} \ln |\det C_i| \right],$$
(11)

where $M(y(b_i), \mu_i, C_i)$ is the Mahalanobis distance between the pixel value $y(b_i)$ and the mean vector μ_i with the covariance matrix C_i . The values μ_i and C_i are modeled as linear mixtures of mean vectors and covariance matrices based on proportions of respective land cover classes $c(a_{ji})$ inside the pixel b_i .

Likewise, the panchromatic conditional energy U(p|c) is defined as:

$$U(z|c) = \sum_{i,j} \frac{1}{2} \left[\frac{\left(z(a_{j|i}) - \nu_{\alpha} \right)^2}{\sigma_{\alpha}^2} + \ln \sigma_{\alpha}^2 \right]$$
(12)

where the first term is the Mahalanobis distance between the pixel value $z(a_{ji})$ and the mean value v_{α} with the standard deviation σ_{α} of the respective class $c(a_{ji}) = \alpha$ respectively.

2.3 Energy minimization

The MAP solution for the SR map *c* can be found by minimization of (8) with respect to *c*. This problem is computationally intractable because of large number of possible configurations in *c*. Therefore we adapt simulated annealing for energy minimization (Geman and Geman, 1984). For the initial estimate of *c* we took the downscaled maximum likelihood classification of multispectral image *y*, where all subpixels a_{ji} corresponding to the pixel b_i were assigned to the same land cover class.

2.4 Validation

Accuracy assessment of the classification results was based on the error matrix for the classes tree crown and background. We used commission and omission errors and the kappa coefficient to quantify the pixel based classification accuracy of the obtained results.

In addition to pixel-based accuracy assessment we implemented object-based accuracy assessment, because the result in the form of tree crown objects is more meaningful for the users than the raster output. Adjacent groups of pixels classified to the tree crown class were grouped into tree objects. To establish one-to-one comparison between the identified and the reference objects we first splitted all groups of trees in the classification as well as in the reference data into the individual objects. The splitting was performed with the use of reference data for individual tree positions and aerial triangulation. Our indicators of accuracy are based on polygon overlay between a classification and a reference layer and include false positive errors (Type I) and false negative errors (Type II) for tree identification. Furthermore, we adopt accuracy indicators related to the topological and geometrical difference between the identified and reference tree crown objects (Möller et al., 2007, Clinton et al., 2008, Ragia and Winter, 2000). The topological accuracy quantifies how well the extent of classified objects O_i fit the reference objects R_i , using measures of over- and under identification of the extent of each tree crown object. The geometrical accuracy measures the positional correctness of the extracted objects based on the distance between the centroids of reference and classified tree objects. We describe the over- and under identification of extent of O_i with respect to a reference tree crown object R_i as:

$$Over_id(O_i) = 1 - \frac{\sum_j \operatorname{area}(O_i \cap R_j)}{\operatorname{area}(O_j)}$$
(13)

Under
$$id(O_i) = 1 - \frac{\sum_j area(O_i \cap R_j)}{area(R_i)}$$
 (14)

Where values of Over $id(O_i)$ and Under $id(O_i)$ close to zero represent a good match between classified and reference objects and values close to 1 represent a large difference in extent between classification and reference. The geometric quality of the identified trees was also measured by the Total_error index, that combines over and under identification of each object in a distance measurement as:

$$\text{Total_error}(O_i) = \sqrt{\text{Over_id}(O_i)^2 + \text{Under_id}(O_i)^2} \quad (15)$$

where total error is in $[0, \sqrt{2}]$ with values close to zero defining a good topological match between reference- and classified objects.

3 EXPERIMENTS

We applied the developed method on an image from residential area of Enschede, The Netherlands. The study area is shown in Figure 1. It has an extent of 4.2 ha and contains 104 trees. In the study area, tree crown diameter varies from 2 to 25 m and trees are planted along the roads, in green areas, and in private gardens. Depending on their location, surfaces surrounding the trees are: pavement, grassland, and bare soil. As for distance between trees, the area contains isolated trees, tree pairs and clusters of interconnected and adjacent trees. The variation in tree location, distance between trees and crown size is representative for many residential areas in Dutch cities.

Tree identification was implemented in a Quickbird image captured on September 21, 2006 in multi-spectral and panchromatic mode with a spatial resolution of 2.4 m and 0.6 m respectively, giving the scale factor value S = 4. For Quickbird images, the panchromatic mode captures surface reflectance between 0.4-0.9 μ m, while the multi-spectral mode provides four spectral bands in the blue (0.4-0.5 μ m), green (0.5-0.6 μ m), red (0.6-0.7 μ m) and near infrared (0.7-0.9 μ m) parts of the electromagnetic spectrum. The proposed SRM method requires accurate geometric registration between pixels of the panchromatic and the multi-spectral datasets (see Equations (1) and (equ:eq2)). To avoid unwanted geometric and radiometric distortions we did not resample the image.

Although only tree crowns are of interest, spectral classes representing other land cover classes had to be defined to properly handle spectral mixing in the image *y*. Thus, five spectral classes were defined for SRM classification. Based on visual interpretation and analysis of the feature space, spectral classes characteristic for tree crown, grassland, shadow and impervious areas were sampled. For each of these classes a training set was defined for estimation of



Figure 1: Quick-bird scene of study area. (Left) multispectral image y, false color composite RGB(4,3,2). (Right) Panchromatic image z

Method	Omission(%)	Comission(%)	Kappa
MLC	56.18	30.29	0.46
SRM	28.23	34.53	0.61

Table 2: Pixel based classification errors for tree crown identification using maximum likelihood classification and SRM

class mean μ and covariance *C* values for the image γ and mean v and standard deviation σ values for the image z. To avoid overfitting of the method to the data of the study area we selected training samples from a larger Quickbird scene outside the study area. The separability of tree crown class from the other classes was evaluated by computing transformed divergence (TD). The obtained TD values were found 1.57, 1.99, 2.00 and 1.27 for the classes grass, shadow, impervious and shrubs, respectively. This indicates that the tree crown class may be spectrally confused with the classes grass and shrubs. We adopted a 7×7 pixel square window for the neighbourhood system *N*. The simulated annealing temperature was decreased as $T_{l+1} = 0.99T_l$ where *l* is the iteration number. The starting temperature value $T_0 = 1$ was found optimal. The optimization converged within 100 iterations. Optimal values for the parameters λ and λ_p were selected by trial and error with respect to classification accuracy for a selected validation set. As a result of this procedure, values $\lambda = 0.8$ and $\lambda_p = 0.05$ were selected as optimal. Note that the optimal value for λ agrees with the findings of Tolpekin and Stein (Tolpekin and Stein, 2009) for similar TD values. In addition, we note that for $\lambda_v = 0$, which mean ignoring information from the panchromatic image z we got suboptimal classification results.

The validation set was prepared by manually digitizing tree crown objects from a VHR aerial image captured by the DKLN program in Summer of 2006 supported by oblique imagery available for the study area. The digitized tree crown outlines and 104 tree locations used as ground truth are shown in 4 overlaid on the used DKLN image.

4 RESULTS

The SRM result obtained for the study area is shown in Figure 2. The result of maximum likelihood classification (MLC) obtained with the same class parameters are shown in Figure 3. The results for pixel based classification accuracy assessment for tree crown identification with SRM and MLC are reported in Table 2.

From Figure 2 we observe that the SRM classification output produces compact tree crown regions. Such regions can be



Figure 2: SRM classification results



Figure 3: Maximum likelihood classification results



Figure 4: False color composite of VHR aerial image used to delineate the reference tree crown objects (shown in yellow)

Туре	Identification	Error Type I	Error Type II
Individual	76 (73%)	33 (37%)	28 (27%)

Table 3: Positive identification, and error types I (commission) and type II (omission) for SRM tree identification

easily transferred to an object domain with a vectorization operation.

By comparing the SRM crown objects with the reference data we found that the SRM classification of the QuickBird image permitted to identify 76 out of 104 trees existing in the area, with an error rate Type-I ate of 33% (commission), and an error rate Type-II of 27% (omission) as shown in Table 3.

Figure 5 presents the error map for the SRM tree crown identification obtained from object based accuracy assessment. Figure 6 present the distribution of over-identification and under-identification errors for tree crown objects obtained from SRM.

5 DISCUSSION AND CONCLUSIONS

The SRM method proposed here for tree crown identification produced superior results to a conventional maximum likelihood classifier according to our pixel based accuracy assessment. In particular, the SRM map presents spatially compact regions that resemble tree crown objects and can easily be combined with other vector data in a GIS. Object based accuracy assessment revealed that the SRM identified 73% of the trees in the area and most of the missing objects have a tree crown radius smaller than the resolution cell of the QuickBird image. Most of the medium and large trees (crown diameter > 3 m) could be identified, which are most important for the municipality in terms of maintenance and replacement costs, as well as being protected by law. Trees in groups, with intersecting crowns, could not be separated and were mapped as one object. Regarding the quality of delineation of identified tree crown objects,



Figure 5: Accuracy map for the SRM tree crown identification with indication of total error, over-identification and under-identification of matching objects between reference and classification. Missing trees and false identifications represented as errors type-II and type-I respectively



Figure 6: Scatter plot of under and over identification errors of tree crown objects in SRM

we found acceptable results given the image resolution, with over-identification in areas of poor spectral contrast between tree crown and background surfaces, such as trees surrounded by (high) shrubs. In a number of cases it was even difficult in the field to accurately map the extent of a tree crown object because of the large heterogeneity of the scene. The omission errors were found mainly associated with trees in strong shadow areas. Such trees are characterized by different spectral values and require separate treatment which is possible with our method.

The SRM classification as implemented here, was used for tree crowns identification even though other spectral classes were also mapped. This study demonstrates a feasible application of an SRM method, which so far has remained mainly experimental. The performance of the method was evaluated on Quickbird image. The method is, however, applicable for tree crown identification from images of other sensors including visible, near-infrared and panchromatic bands.

In this study we reported the SRM result as a crisp classification. Real tree canopy observed from above is in fact a fuzzy object: it is partially transparent, its boundary is difficult to precisely define and the canopy density is decreasing from the center to the boundary. We note that the proposed method is based on posterior probability modelling and therefore allows to compute the posterior probabilities for all land cover classes. Analysis of the posterior probability values allows to convert the classification results into fuzzy memberships as demonstrated in (Bijker et al., 2010). In this way trees can be identified as fuzzy objects, which is a more suitable way to describe real trees compared to a raster thematic map. This issue did not receive sufficient attention in geo-information community and requires further research.

The method described in this study is the extension of the method of Tolpekin and Stein (Tolpekin and Stein, 2009). The novelty is in the incorporation of information from panchromatic band. The contribution of this band in the total energy function was found to be small as expressed by the low value of the parameter $\lambda_p = 0.05$. Nevertheless, this extension led to improved classification results compared to classification ignoring the panchromatic information, i.e. $\lambda_p = 0$. For application of the method to other images and land cover classes the two parameters, λ and λ_p have to be estimated. The optimal value of the parameter λ was found consistent with the automatic procedure proposed by Tolpekin and Stein (Tolpekin and Stein, 2009). Automatic estimation of parameter λ_p requires further research.

The method requires definition of several land cover classes even if the user is interested in only one class. Urban areas are challenging areas for classification, analysis of another city with our method might require definition of more land cover classes. We consider the possibility to provide fuzzy classification output as a way to partially solve this issue and extend the applicability of the method.

Finally we conclude that the method introduced in this study reduces the limitation of insufficient spatial resolution of modern VHR images for tree identification in urban areas and can be recommended for urban tree inventories.

ACKNOWLEDGEMENTS

We thank Alfred Stein for constructive comments on the manuscript. This study has been made possible by funding of the Netherlands Space Office (NSO) through their program Pre-qualification ESA Programs (PEP) for the project Tree and Image (Boom en Beeld). Enschede municipality is kindly acknowledged for their contribution and supplying tree location data, topographic maps, cadastral data and aerial imagery. The Faculty of Geo-Information Science and Earth Observation (ITC) of the University of Twente is acknowledged for making the height data available.

REFERENCES

Atkinson, P. M., 2009. Issues of uncertainty in superresolution mapping and their implications for the design of an inter-comparison study. International Journal of Remote Sensing 30(20), pp. 5293–5308.

Bijker, W., Tolpekin, V. and Ardila, J., 2010. Change detection and uncertainty in fuzzy tree crown objects in an urban environment. In: The international Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. Vol.XXXVIII-4/C7.

Clinton, N., Holt, A., Gong, P. and Yan, L., 2008. An accuracy assessment measure for object based image segmentation. In: ISPRS (ed.), ISPRS Congress Beijing 2008, Vol. XXXVII, Beijing.

Daliakopoulos, I. N., Grillakis, E. G., Koutroulis, A. G. and Tsanis, I. K., 2009. Tree crown detection on multispectral vhr satellite imagery. Photogrammetric Engineering and Remote Sensing 75(10), pp. 1201–1211.

Du, Q., Younan, N. H., King, R. and Shah, V. P., 2007. On the performance evaluation of pan-sharpening techniques. Ieee Geoscience and Remote Sensing Letters 4, pp. 518–522.

Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6, pp. 721–741.

Gougeon, F. A. and Leckie, D. G., 2006. The individual tree crown approach applied to ikonos images of a coniferous plantation area. Photogrammetric Engineering and Remote Sensing 72(11), pp. 1287–1297.

Hirschmugl, M., Ofner, M., Raggam, J. and Schardt, M., 2007. Single tree detection in very high resolution remote sensing data. Remote Sensing of Environment 110(4), pp. 533–544.

Li, S. Z., 2009. Markov random field modeling in image analysis. Springer, New York.

McHale, M. R., Gregory McPherson, E. and Burke, I. C., 2007. The potential of urban tree plantings to be cost effective in carbon credit markets. Urban Forestry & Urban Greening 6(1), pp. 49–60.

Möller, M., Lymburner, L. and Volk, M., 2007. The comparison index: A tool for assessing the accuracy of image segmentation. International Journal of Applied Earth Observation and Geoinformation 9(3), pp. 311–321.

Ragia, L. and Winter, S., 2000. Contributions to a quality description of areal objects in spatial data sets. ISPRS Journal of Photogrammetry and Remote Sensing 55(3), pp. 201–213.

Tolpekin, V. A. and Stein, A., 2009. Quantification of the effects of land-cover-class spectral separability on the accuracy of markov-random-field-based superresolution mapping. IEEE Transactions on Geoscience and Remote Sensing 47(9), pp. 3283–3297.

Woodcock, C. E. and Strahler, A. H., 1987. The factor of scale in remote sensing. Remote Sensing of Environment 21(3), pp. 21.