Color Based Clustering for Trunk Segmentation

László Czúni, András Kürtösi, Karim Ben Alaya
Department of Information Systems and Electrical Engineering, University of Pannonia, Egyetem u. 10., Veszprém, Hungary
czuni@almos.vein.hu

***Abstract*—While image segmentation is a fundamental and general task of image processing and computer vision, the segmentation of trunks in wood can be considered as a specific, ill posed, and hard problem. However, besides the lots of possible applications of it, the problems to be faced are general and can be found in other areas of open-air image analysis. In our paper we discuss these problems and propose a color clustering based approach as a feature generation step for the segmentation of trunks. The results of classification can be used as features for further segmentation techniques such as MRF. The possibility to apply for angle count based volume estimation is also discussed. The performance is shown through several tests with different types of wood.**

***Index Terms*—Forest inventory, Image segmentation; Clustering; Logistic Model Tree, k-means, Angle count sampling**

I. Introduction

There are lots of possible applications of computer vision in agriculture and forestry from product quality monitoring, classification, and sorting to automated harvesting [1]. The detection, recognition, and classification of trees is an often discussed problem either in remote sensing [2] or in short distance forest inventory [3].

In our paper we focus on the detection of trunks in forests. The possible application of a reliable algorithm can be volume estimation, forest surveying, collision avoidance of unmanned ground/aerial vehicles. For example the so called ”survey runs” are very frequent and important duties in forest inventory. The goal of survey runs is to determine the total harvestable wood volume in a given area. A surveyor must first measure the height of the trees and classify the diameter at breast height, then there are methods to extrapolate this data to estimate the total harvestable wood volume. Besides the traditional manual tools to support this survey there are several possibilities such as monocular digital cameras, stereo cameras, TOF (Time of Flight) cameras, lidars, ultrasonic sensors.

There are several issues that make the segmentation of trunks difficult: the large variety of object color and texture features, shapes, the different appearances due to different lighting conditions, weather and seasonal effects. Please see Fig. 1 for illustrations. On the left of Fig. 1 trunks with different orientation, color and occlusion are visible; in the middle the contrast is much lower and there is a very high density of trees; on the right there is serious occlusion by leaves and there are bright and dark areas of the same trunk.

In our paper we do not deal with model based segmentation considering shape, size or texture. Our purpose is to achieve a good clustering of colors and classification of clusters serving information for other segmentation techniques such as Markov Random Fields, Mean Shift or model based techniques [4]. The output of clustering will be a binary map, where trunk and background pixels will be given different labels. The estimation of tree volume, based on the diameter of trunks is also a good utilization of the proposed classification as will be shown in Section IV.

Fig. 1. Trunks with different orientations, colors and occlusions (left). Trees with low contrast, gray like color and very high density (middle). Serious occlusion by leaves and partially highlighted and shadow areas (right).

II. Related Techniques and Papers

Perceptual grouping refers to humans' visual ability to abstract high level information from low level image primitives without any specific knowledge of the image content. There can be several cues which help us in grouping such as similarity, proximity, symmetry, periodicity. Segmentation (dividing images into non-overlapping homogeneous regions) is an old and wide topic in computer vision and image processing. There are several general approaches such as: bottom-up methods; superpixels; interactive methods; object proposals; semantic image parsing; image cosegmentation; convolutional neural networks. The purpose of many of them is to bridge the semantic gap between low level and high level information. A broad overview can be found in [5].

As introduced above, the segmentation of trunks in forest images is a difficult problem. In [3] the task was to determine

978-1-5386-6979-2/18/$31.00 ©2018 European Union

the boundaries of target trees in a forest. The segmentation was challenging since the target could be occluded by other trees. To compensate for this possibility, an algorithm sensitive to small texture changes was developed. Unfortunately, this technique can be applied only for short distances otherwise the detection of the texture of trunks would require unreasonably high camera resolution and amount of pixel information.

According to [6] volume estimates can be obtained from high-resolution panoramic photographs of forest interiors. Gigapixel images were captured in an Alaskan birch forest, then the stand basal area was measured and the result was compared to estimates made using traditional Bitterlich visual scan. According to the authors: "The image-based method has great precision and acceptable accuracy when tree trunks are not obscured by other trees or shrubs, when stitching is error free, and when adequate depth of field provides well focused images."

3D imaging technologies can overcome some of the problems of 2D imaging easily [7], [8], [9]. During terrestrial laser scanning (TLS) a tripod is placed in the spot and the laser scanner measures individual trees or field plots. TLS produces a dense point cloud from the surrounding trees. Mobile laser scanning (MLS) also produces 3D point clouds but often utilizes global navigation satellite system (GNSS) and inertia measurement units (IMUs) to be able to allow movements during measurements. Similarly, airborne lidar remote sensing is also a powerful tool of forest inventories [10]. The drawback of all of these scanning technologies is the price and the relatively long processing time of the large amount of data.

III. The Proposed Approach

There are a lot of possible segmentation approaches to be used but generally speaking they are based on image features such as color, texture, edges or other local characteristic attributes. Our purpose is to investigate the generation of pixel level color based features to be used for segmentation. Thus we are not building high level models to full the semantic gap rather investigate low level processing to find good color features. Fig. 2 gives a brief overview of the proposed classification method. In our paper we will evaluate the performance of these features with a pixel-wise detection (segmentation) step where trunks and background pixels (soil, leaves, sky, rocks) are to be separated. As it is clearly seen from the sample images of Fig. 1 and Fig. 3 the appearance of the same (types of) objects shows large variability from image to image or even within one image. This means that the clustering of pure RGB values results in inaccurate pixel-wise classification of trunks and background. To handle this problem we propose the following steps (also given in Fig. 2).

A. Preprocessing of RGB Channels

Since the environmental effects and the tree species result in large variations of trunk and background color there is no simple general rule (such as histogram equalization, contrast enhancement, etc.) for the preprocessing of RGB information. To handle this problem we generated, with the help of a forestry engineer, 8 color channels from the input RGB, using a set of about two dozen different images:

- Increasing saturation with 50% in the HSV color space (results in R', G', B'). This enhances the classification of gray like areas.
- Increasing contrast: $(R_m, G_m, B_m) = norm(R-m, G-m, B-m)$ where $m = min(R, G, B)$.
- Estimation of saturation: $I_{Max-Min} = max(R, G, B) - min(R, G, B)$.
- Since a large difference of R and G is typical for different vegetation covered areas we computed: $I_{G-R} = G - R$.

Thus instead of using RGB now we have derived 8 color channels which showed good performance, through a number of subjective visual tests, to discriminate trunks and background.

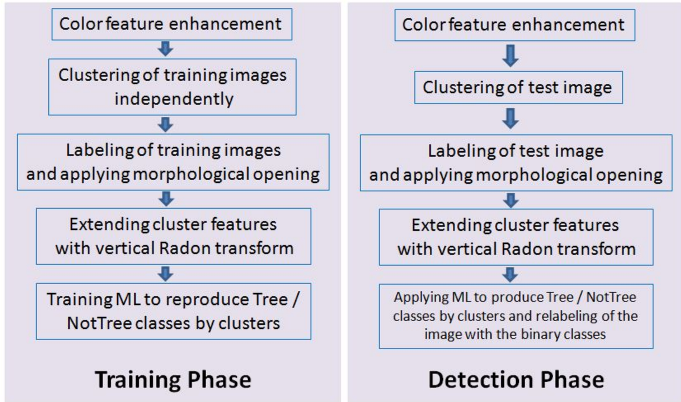

Fig. 2. The overview of the pixel-wise clustering and classification method.

B. Clustering of Data Points

While there are only two classes in our task (foreground - trunk and background - any other object) there are more typical objects in the forest such as leaves, fallen leaves, soil, rocks, sky. On the 8 channel pixel values we run k-means clustering to get 10 clusters with their centroid values. This step is independently applied to all training images since the appearance of the objects may differ from image to image. By collecting these clusters from a larger set of training images we can generate prototype clusters for foreground and background. We can also re-label the input images with these cluster labels for visualization. If any cluster belongs to a tree then its vertical Radon transformation creates high peaks due to the vertically standing nature of the trunks. The highest value of these peaks are attached to each cluster as a new feature resulting in a cluster descriptor vector of dimension 9.

C. Training with Decision Trees

We used 13 different training images of size 600x800 to go through the above process resulting in 130 clusters. The 13 training and the further test images showed the following species: beech, turkey oak, sessile oak, hornbeam, black alder, black locust. Then each cluster centroid was classified as

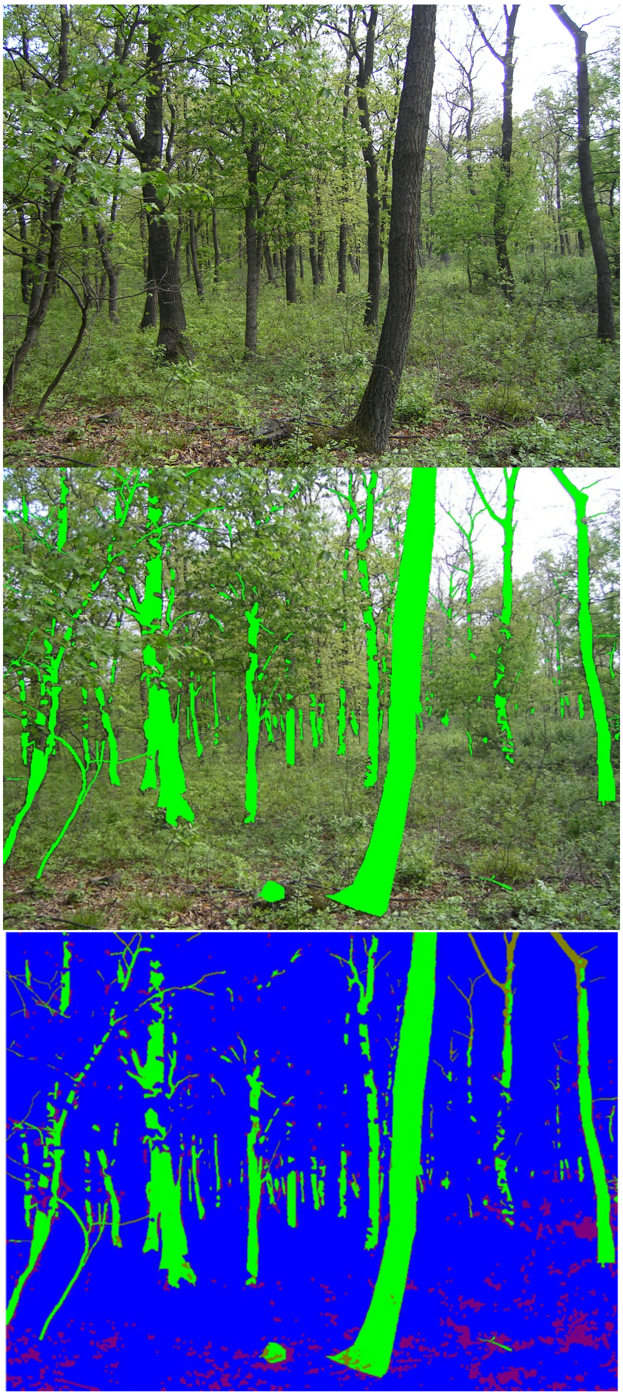

Fig. 3. An example photograph of trees (top), the manually annotated version (middle), and the evaluation of automatic classification (bottom). Green pixels are the correctly classified trunks, blue ones are the correctly classified background, other colors code wrong classification.

trunk or background by a manual annotation process. If one cluster belonged either to foreground and background objects the decision was made based on the majority rule. We tested different machine learning methods to build a binary classifier for the recognition of trunks based on the 9 dimensional feature vectors, now we show the results obtained with Logistic Model Tree (LMT) [11]. To verify the coding efficiency of the 130 clusters and the classification performance of LMT we repeated the whole process on 12 new test images but applied the already trained LMT. Results are given in Table I using the following notations: TPR - True Positive Rate; FPR - False Positive Rate; F1 - F1 Score; MCC - Matthews Correlation Coefficient. All measures show very good performance on each class.

TABLE I
EVALUATION OF TRAINED CLUSTERS AND TRAINED LMT ON 12 TEST IMAGES.

	Classification measures			
Class	TPR	FPR	F1	MCC
tree	0.964	0.022	0.947	0.931
bg.	0.978	0.036	0.984	0.931

D. Testing

The purpose of our method is the clustering of colors and we do not expect precise segmentation: for this purpose spatial and pixel neighborhood information should also be included to get homogeneous regions and to rule out homogeneous larger background regions. Nevertheless, to evaluate the performance as pure pixel level segmentation we used three test images manually annotated. The segmentation is carried out by the above described steps (clustering into 10 classes, labeling image with class labels, morpholigical opening, extending cluster features by vertical Radon transformation, recognition by LMT) also given in Fig. 2. While Table II contains the average performance of these tests, see Fig. 3 for illustrations of input, output, and evaluation of one of the test images.

TABLE II
EVALUATION OF PIXEL-WISE DETECTION OF TRUNK PIXELS AS AN AVERAGE OF 1.44M PIXELS OF 3 TEST IMAGES.

	Classification measures			
	TPR	FPR	F1	MCC
average for 3 images	0.869	0.17	0.693	0.65

IV. APPLICATION: ANGLE COUNT SAMPLING

In practice still the most common technique to make survey runs for volume estimation is based on the manual survey by human observers. The *angle count sampling method* (also called *prism cruising* or *Bitterlich sampling*) was developed by Walter Bitterlich (1948) an Austrian forester: from a sample point the neighboring trees are selected strictly proportional to their basal area. The only device needed to do this can be anything that produces a predefined viewing angle. That can be a dendrometer, a relascope, a wedge prism. While standing on a sample point and aiming over the device to the diameter, at breast height, of the surrounding trees; all trees that appear larger than the opening must be

978-1-5386-6979-2/18/$31.00 ©2018 European Union

counted in a sweep around 360^o [12].

While in [6] digital images are used for angle count sampling, itself the measurement was done manually. iBitterlich (http://www.taakkumn.com/taakkumnRoom/iBitterlich.html), Bitterlich relaskope (http://www.deskis.ee/relaskoop/), and MOTI (http://www.moti.ch) are also manual tools for counting trees with the help of the screen of a mobile phone.

To use a camera to get the horizontal viewing angle of the trunks the camera should be calibrated with standard methods. The difference between the visible diameter and the real diameter can be neglected in case of large distances.

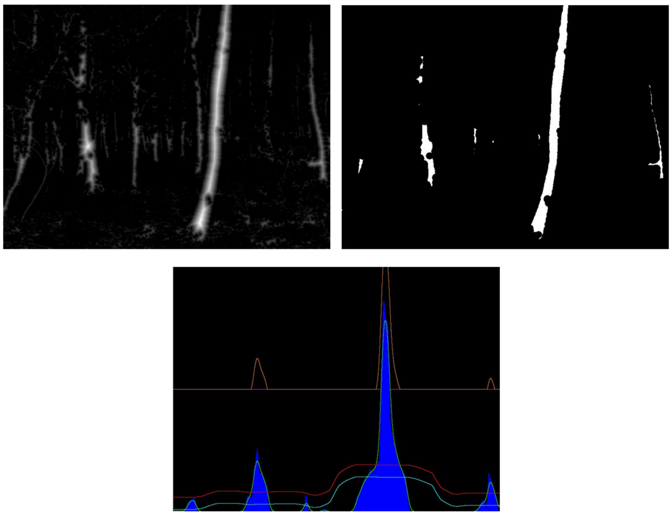

Fig. 4. Distance transform (DT) of automatically detected pixels of Fig 3 (Top left). Thresholded DT (Top right). Vertical Radon transform (RT, blue) of Top right image with strong smoothing (light blue line) and 3 peaks detected (upper part) as thresholded difference of the strongly smoothed and slightly smoothed RT signal (Bottom image).

In case of calibrated cameras the viewing angle of trunks, needed for the Bitterlich sampling, can be estimated with the processing of classified pixels obtained using our proposed method through the following steps:

1) Distance transform (DT): Distance transform is applied on the image containing only the trunk pixels.
2) Thresholding: Naturally the image after DT contains several noisy areas but for Bitterlich sampling we need only trunks which have diameter above a threshold. Thus thresholding the DT image with the half of the threshold diameter we get rid off those areas.
3) Radon transformation: Due to possible overlapping tilted trunks or heavy occlusion by leaves these thresholded images can be still unsatisfactory thus we applied vertical Radon transformation for the detection of vertical structures.
4) Detection of peaks: The 1D vector of vertical Radon transformation should be smoothed and high peaks are to be detected to get the number of trees above a given viewing angle required for volume estimation.

These above steps are illustrated in Fig. 4.

V. Conclusion and Future Work

We discussed pixel level clustering and classification of forest images for trunk segmentation. There are several applications of trunk segmentation from wood volume estimation to collision avoidance of unmanned vehicles. Color features themselves, without higher level shape or area models, are not suitable for accurate trunk segmentation. However, with the proposed techniques we can cluster image pixels used in further shape analysis or segmentation (with MRF, Mean Shift, etc.). As a direct utilization of the proposed clustering we showed the main steps of automatic angle count sampling. In future we plan to apply uncalibrated stereo to generate rough disparity maps as further features to get better pixel level classification of trunks.

Acknowledgment

This work was partially supported by the NKFI-6 fund through project K120369 (OTKA) and by Széchenyi 2020 under EFOP-3.6.1-16-2016-00015.

References

[1] L. Saxena and L. Armstrong, “A survey of image processing techniques for agriculture,” 2014.

[2] T. Brandtberg, T. A. Warner, R. E. Landenberger, and J. B. McGraw, “Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America,” *Remote sensing of Environment*, vol. 85, no. 3, pp. 290–303, 2003.

[3] J. Byrne and S. Singh, *Precise image segmentation for forest inventory*. Carnegie Mellon University, The Robotics Institute, 1998.

[4] C. Benedek, X. Descombes, and J. Zerubia, “Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 33–50, 2012.

[5] H. Zhu, F. Meng, J. Cai, and S. Lu, “Beyond pixels: A comprehensive survey from bottom-up to semantic image segmentation and cosegmentation,” *Journal of Visual Communication and Image Representation*, vol. 34, pp. 12–27, 2016.

[6] C. L. Fastie, “Estimating stand basal area from forest panoramas,” in *Proceedings of the Fine International Conference on Gigapixel Imaging for Science*, 2010.

[7] M. Vastaranta, T. Melkas, M. Holopainen, H. Kaartinen, J. Hyyppä, and H. Hyyppä, “Laser-based field measurements in tree-level forest data acquisition,” *Photogramm. J. Finl*, vol. 21, pp. 51–61, 2009.

[8] X. Liang, P. Litkey, J. Hyyppa, H. Kaartinen, M. Vastaranta, and M. Holopainen, “Automatic stem mapping using single-scan terrestrial laser scanning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 661–670, 2012.

[9] X. Liang, J. Hyyppä, A. Kukko, H. Kaartinen, A. Jaakkola, and X. Yu, “The use of a mobile laser scanning system for mapping large forest plots,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 9, pp. 1504–1508, 2014.

[10] L. Wallace, A. Lucieer, C. Watson, and D. Turner, “Development of a uav-lidar system with application to forest inventory,” *Remote Sensing*, vol. 4, no. 6, pp. 1519–1543, 2012.

[11] N. Landwehr, M. Hall, and E. Frank, “Logistic model trees,” *Machine learning*, vol. 59, no. 1-2, pp. 161–205, 2005.

[12] C. Kleinn, *Lecture notes for the teaching module forest inventory*. Department of Forest Inventory and Remote Sensing. Faculty of Forest Science and Forest Ecology, Georg-August-Universität Göttingen, 2007.

978-1-5386-6979-2/18/$31.00 ©2018 European Union