Estimating the number of planting microsites using UAV and computer vision

Houssem-Eddine Guezbiri, Wassim Bouachir, Rafik Gouiaa, Ikram Djafri, Nicolas Bélanger
TÉLUQ University, Montréal, Canada

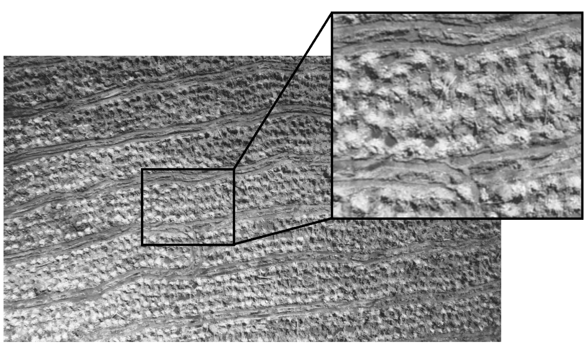
For optimal survival and early yields of tree plantations, forest managers employ strategies that tend to promote a fast development of the root system, which favors water and nutrient uptake [1]. However, certain conditions such as high soil bulk density due to compaction by logging machinery, waterlogging, and plant competition for resources can severely impede root establishment of the planted trees [2]. Under such conditions, mechanical site preparation is generally useful because the induced soil disturbance can improve the quality of planting microsites by increasing soil temperature and water retention capacity, decreasing soil density and plant competition [3]. The redistribution of some nutrient-rich soil layers at depths that coincide to the rooting depth of tree seedlings can also have significant benefits on juvenile growth [4]. Mechanical site preparation by mounding (see Figure 1) is thus quite popular within the forest industry in Quebec and is recommended by the Quebec Government [5] for the intensive culture of coniferous and deciduous tree species. However, since the number of mounds that are prepared with the machinery can vary depending on site characteristics, planting operations are usually planned following a manual count of the mounds. This operation requires several people, is time consuming and costly, and is subject to error. In this context, our research aims at developing a novel computer vision method using UAV imagery for fast and accurate estimation of the number of mounds on a mechanically prepared site.

Figure 1. Examples of mechanically prepared mounds in the boreal forest.

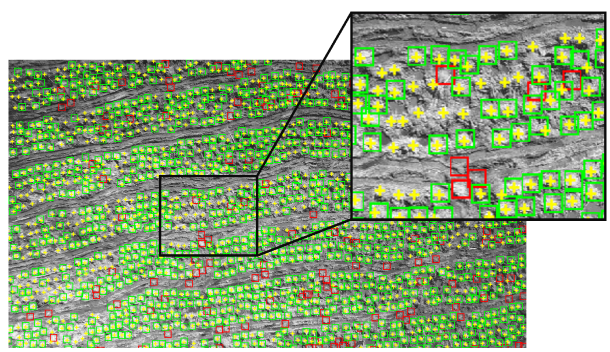
Most existing work on automatic object counting can be categorized into three approaches: (1) the object detection approach that consists in applying an object detector to locate different object instances on the image [6]. Object detection is also an unresolved problem because of multiple real-world difficulties such as overlapping objects, occlusion and scene clutter; (2) the density estimation approach which avoids the above mentioned disturbing factors. This approach estimates a continuous density function whose integral on a region of the image gives the number of objects [7]; and (3) the segmentation approach using a hybrid model to combine detection and density estimation for object segmentation [8]. In our application context, automatic counting need to be performed on aerial images where the objects of interest are separated by approximately two to three meters. In this case, overlap and occlusion situations are not considered. Moreover, it is not necessary to determine precise mound segmentation. We thus adopted the object detection approach to develop an object detection and counting method based on the cascade detector framework [9]. We propose a supervised learning approach for mound detection and counting. Our object detector was firstly trained on local binary pattern (LBP) features [10] extracted from annotated images in the Near InfraRed (NIR) spectrum. For this purpose, we manually annotated a total of 18,151 mounds on aerial images taken at different altitudes ranging from 50 m to 125 m, to be used as positive learning examples. To extract negative examples, we used the location of annotated mounds to generate a distance map from which negative examples are sampled at local maximum distances. This ensured sampling negative examples in between narrow neighbouring mounds. Once the cascade classifier was trained offline, object detection was then achieved using a sliding window on new aerial images.

To evaluate the proposed method, we constructed a dataset by capturing aerial images for several mechanically prepared sites. The sites were overflown using a data acquisition UAV equipped with multispectral and visible cameras at different altitudes. In our experiments, precision, recall, and overall accuracy were calculated by training the method on annotated images from three different sites, while testing was carried out on unknown images from a site that had not been included in the training step. We investigated various choices regarding the acquisition conditions, image types, as well as the most relevant visual features and detection techniques. This investigation also covered the exploration of deep learning techniques by adapting pre-trained CNNs for mound detection using transfer learning.

Preliminary results show that the proposed method allows to reach the highest accuracy rate of 82% among the tested methods (see Figure 2). Our experimental results also suggest that NIR is the optimal choice for the image type. This result can be explained by the temperature difference between a mound and its neighborhood, which makes mounds more distinguishable on NIR images. In fact, the surface of the mounds are formed of mineral soil and are bare of vegetation, whereas between the mounds, there is presence of dead cover (formed by organic matter) and vegetation. Upcoming efforts will therefore be inspired by this observation, and will include the use of a thermal IR camera mounted on the UAV to better exploit the contrast in temperature between the mineral soil of mounds (warmer) and the dead cover between mounds (colder).



(a)



(b)

Figure 2. Example of detection result: (a) NIR image of mechanically prepared site at 100 m of flight altitude, and (b) color coded mound detection results: ground truth center (yellow cross), good detection (green) and false alarm (red).

References

[1]. Grossnickle, S.C. 2005. Importance of root growth in overcoming planting stress. New For. 30: 273–294.

[2]. Wolken, J.M., Landhäusser, S.M., Lieffers, V.J., Dyck, M.F. 2010. Differences in initial root development and soil conditions affect establishment of trembling aspen and balsam poplar seedlings. Botany, 88: 275–285..

[3]. Löf, M., Dey, D.C., Navarro, R.M., Jacobs, D.F. 2012. Mechanical site preparation for forest restoration. New For. 43: 825–848.

[4]. Bilodeau-Gauthier, S., Paré, D., Messier, C., Bélanger, N. 2013. Root production of hybrid poplars and nitrogen mineralization improve following mounding of boreal Podzols. Can. J. For. Res. 43: 1092–1103.

[5]. Prévost, M., Thiffault, N. 2013. La préparation de terrain. Dans Le guide sylvicole du Québec. Tome 2. Les concepts et l'application de la sylviculture. Larouche, C., Guillemette, F., Raymond, P., Saucier, J.-P., Éditeurs. Les Publications du Québec, Québec, Québec, pages 134–157.

[6]. Barinova, O., Lempitsky, V., & Kholi, P. (2012). On detection of multiple object instances using hough transforms. IEEE Trans. PAMI, 34(9), 1773-1784.

[7]. Lempitsky, V., & Zisserman, A. (2010). Learning to count objects in images. In Advances in Neural Information Processing Systems (pp. 1324-1332).

[8]. Ryan, D., Denman, S., Fookes, C., & Sridharan, S. (2009, December). Crowd counting using multiple local features. In IEEE Digital Image Computing: Techniques and Applications, 2009. DICTA'09. (pp. 81-88).

[9]. Viola, P., and M. J. Jones. "Rapid Object Detection using a Boosted Cascade of Simple Features." Proceedings of the 2001 IEEE Computer Society Conference. Volume 1, 15 April 2001, pp. I-511–I-518.

[10]. Ojala, T., M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-scale and Rotation Invariant Texture Classification With Local Binary Patterns". IEEE Trans. on PAMI, Vol. 24, No. 7 July 2002, pp. 971–987.