TERRAIN CLASSIFICATION USING AIRBORNE LIDAR DATA AND AERIAL IMAGERY

O. Brattberg[*], G. Tolt

Div. of Sensor Systems, Swedish Defence Research Agency (FOI), Linköping, Sweden
(oskar.brattberg, gustav.tolt)@foi.se

Commission III, WG III/3

KEY WORDS: Airborne Laser Scanning, LIDAR, Aerial Imagery, Bare-Earth Estimation, Classification, Spectral Analysis

ABSTRACT:

Updated geographic information and 3-D environment models are becoming increasingly important for both military and civilian applications such as simulation, mission planning, visualization, landscaping, etc. In order to reduce the time and efforts needed to produce accurate and high-resolution models of an area of interest, rapid and highly automatic methods for extracting geographic information from sensor data are required. In this paper, we present recent results from the development of methods for processing airborne lidar data and aerial imagery, that aim towards rapid and automatic classification of important terrain structures: bare-earth, buildings and vegetation. First, an approach for bare-earth estimation is presented that is based on combining two bare-earth estimation techniques. Second, a decision-level fusion approach for classification of lidar data is presented. Third, an approach for updating an existing classifcation result using RGB images is described.

1. INTRODUCTION

Realistic high-resolution 3-D environment models are desirable in many applications, civilian as well as military, as they can be used for visualization, sensor simulation, mission planning and rehearsal, etc. In order to produce the models, detailed and up-to-date geospatial information of the area of interest is needed. In many situations, especially in international operations, the existence of such information is far from certain. Hence, the ability to automatically and accurately extract important geospatial information about the terrain from sensor data is a valuable asset. In applications within the scope of the Swedish Armed Forces, the goal is to be able to perform simulations using a high-resolution 3-D model of an area of operation within 48 hours from data acquisition. In order to meet this goal, fast and accurate data processing techniques are needed, so that the amount of tedious manual work, such as verification and correction of results, can be decreased.

During the last decade, lidar systems have been extensively used for the purpose of extracting 3-D geospatial information, as they provide direct and accurate range measurements also from partially occluded structures, such as the ground underneath tree canopies. One of the most important pieces of geospatial information is the bare-earth level, or Digital Terrain Model (DTM). Consequently, a number of techniques for bare-earth estimation using airborne lidar data have been proposed. Refer to (Sithole and Vosselman, 2004) for an overview. The *off-ground* data points are typically classified into a number of classes, of which buildings and vegetation are two of the most common ones. In order to achieve this, lidar-based classification often relies on a number of measures designed as to capture spatial variation among the data points: maximum slope (Maas, 1999), Laplace operator, local curvature (Vögtle and Steinle, 2003), multiple echoes, etc. See (Rottensteiner *et al.*, 2005) for an overview on lidar features and (Pfeifer *et al.*,

2007) for a survey of techniques for building detection using airborne lidar data.

Although lidar-based classification often provides a good way of distinguishing between buildings and trees, it is error-prone when the spatial properties of vegetation and buildings become very similar. This kind of problem typically occurs with dense, trimmed hedges, in sparsely sampled regions, with buildings having a weirdly shaped roof, etc. In order to be able to classify those objects correctly, more information may have to be used. In this paper we consider RGB images acquired from an airborne platform as this second source of information, since cameras producing such images are very common, relatively inexpensive and can be mounted on UAVs, thus desirable to use for military purposes. See (Schenk and Csathó, 2002) for a discussion on strengths and limitations of lidar data and aerial imagery, respectively.

Assuming that some *a priori* information about the terrain exists, possibly as a result of a previous classification step (e.g. lidar-based) or in the form of an ordinary map, it is desirable to be able to update the results using the available spectral (RGB) information together with the existing information about this area. The spectral signature in the RGB band is light reflected of the objects in the scene, in our case the light originates from the sun. For every wavelength the reflection strength varies depending on the material of the reflecting object. This gives us the possibility to distinguishing different materials from each other. This phenomenon has been used together with multi- and hyperspectral cameras to discriminate between materials created to mimic the real forest (camouflage nets) with good results. Refer to (Ahlberg and Renhorn, 2004) for more information. The same techniques in a less complex manner using the simpler RGB-camera can be used to discriminate between materials. By doing so, possible errors in the existing geospatial information can be corrected and more accurate and

---

[*] Corresponding author

up-to-date knowledge about the terrain is thus obtained. Refer to (Schowengerdt, 1997) for an overview of spectral-based classification.

In this paper we will outline some work performed at the Swedish Defence Research Agency (FOI) on the development of methods for processing data from airborne sensors in the context of rapid 3-D environment modelling. First, work on lidar-based bare-earth estimation and classification is presented. Then some initial work on updating classification results using RGB image data is reported.

2. LIDAR DATA PROCESSING

In this section, we present an approach for lidar data processing aiming towards a highly automated processing framework, from raw data to extracted geospatial information. The lidar data used in this work have been obtained with the Topeye Mk II system (www.blomasa.com/sweden/se/topeye) and the point density in the data sets corresponds to about 8-10 points/m².

2.1 Pre-processing

The lidar data points from the area under study are first organized as tiles (squares) and rasterized onto a grid where each cell corresponds to 0.25m, in order to speed up the some of the following computations. Each tile represents a region of a typical size of 200mx200m plus an additional boundary, i.e. overlap between neighbouring tiles, to reduce the boundary effects. In each tile, outliers are removed, using the assumption that there are less than k% outliers within a small neighbourhood (a typical value of k used in this work was 2). The data points are sorted according to elevation value and the points among the outermost k% that are too far from the rest of the points are simply discarded.

2.2 Bare-earth estimation

The next step is to estimate the *bare-earth* level. When concerned with rapid mapping of large regions using high-density lidar data, the computational complexity becomes an issue. On one hand, the overlap between neighbouring tiles should be small in order to speed up computations. On the other hand, the overlap should be large to allow for seamless integration of adjacent tiles. The problem is that having large overlaps leads to significantly increased computational time or, even worse, causes problems to store data into the computer's memory. It has been noticed that an overlap of 20m is often a reasonable trade-off between computational requirements and result quality.

2.2.1 Bare-earth estimation techniques: Available to us are two techniques developed at FOI, each with its own merits and disadvantages. One is computationally attractive as it is based on some standard image processing algorithms such as watershed segmentation and region growing (Landgård, 2005, Tolt *et al.*, 2006). It is particularly useful in urban areas, where large connected regions of the ground are clearly visible and the ground itself is relatively flat. However, it is typically not as successful in dense forest areas, since the ground there often appears as fragments that do not form large watershed segments which in turn makes it difficult to determine ground seed points from which the region growing process starts. The region growing process itself relies on a threshold that inhibits the

ability to climb steep slopes.

The problems with this approach often manifest themselves near the boundary of a tile, typically if there is a slope with lots of trees on it. This is expected, since if there is hill at the centre of tile, the region growing can often find some way to climb it (basically it "attacks" the hill from all sides). At the boundary, it simply has fewer directions to try.

The second technique is based on the idea of active contours (Elmqvist, 2002). It can be seen as surface (a "rubber cloth") that iteratively adapts to the data by optimizing an energy criterion that balances the internal forces in the surface (forces due to deformations in the surface) with the external forces acting on it (attraction forces between the surface and the data points). It generally produces good results in complex environments such as forests, at the expense of being computationally intensive due to the iterative nature of the optimization procedure.

2.2.2 The bare-earth estimation approach: The bare-earth estimation process starts with applying the fast, region growing-based technique to all the tiles individually. In order to ensure that the resulting DTM does not contain any significant errors, the result has to be verified. Performing an inspection of large areas "by hand" is time-consuming and it may also be quite difficult to spot subtle artefacts. For this purpose, a tool has been developed that detects significant errors automatically. Here we have focused on the most common source of error, referred to as *tile mismatch*; tiles in which the region-growing technique has failed can typically be detected by looking for elevation jumps along the boundaries of the tiles (see Figure 1). Here we define the tile mismatch measure simply as the maximum elevation difference around the tile border. After having detected such mismatches automatically, the associated tiles are marked as candidates for further analysis.

Now, whether detected errors in the DTM are severe enough to actually require further refinement depends on the requirements for the application at hand. In fact, even large errors may be perfectly acceptable if they occur underneath a building (that will cover the mismatch anyway in the final 3-D model). Moreover, what specific action that is the most desirable to take when an error is detected may also vary. For example, problems may be most efficiently solved by correcting the bare-earth model manually, by re-running the algorithm with a different set of parameters, by defining a larger tile to work on or by applying a different bare-earth estimation technique. Here, we take advantage of the fact that two bare-earth techniques are available to us. Hence, mismatching tiles from the first run are subject to bare-earth estimation with the other, more time-consuming algorithm. The resulting bare-earth estimation method, achieved through such a combination of the two techniques, is efficient, as the computational resources are spent in the most difficult parts of the terrain.

In Figure 1-Figure 3 the approach is illustrated with an example. Figure 2 shows the Digital Surface Model (DSM) of this area which contains a number of hills with trees on. The resulting DTM, shown in Figure 3, consists of a majority of tiles from the region-growing technique complemented with a few tiles from the iterative approach and now contains no significant mismatches. The only elevation differences left between the tiles are now mainly due to sensor data noise data and natural height variations around the tile borders.
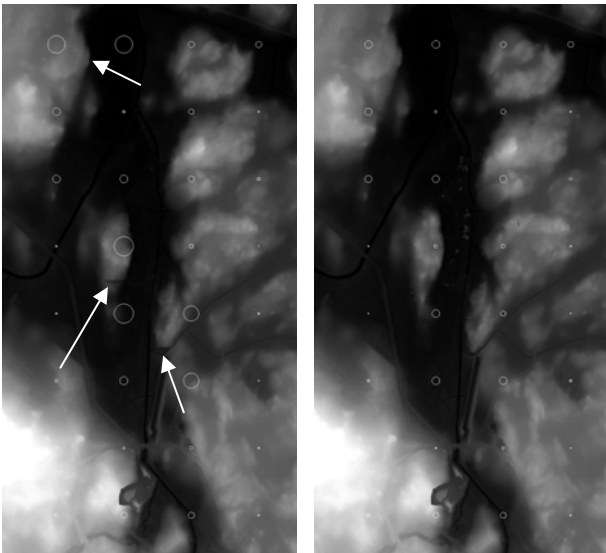
Figure 1. The size of the circles corresponds to the degree of mismatch (defined here as the maximum elevation difference along the tile border) of each tile relative to its neighbours. Left: Bare-earth estimation result using the region growing technique. The arrows indicate the mismatches. Right: Mismatching tiles replaced by tiles from the active contour technique, producing a bare-earth level without significant mismatches.

2.3 Lidar-based object classification

After the bare-earth estimation is finished, the off-ground lidar data points undergo further classification.

2.3.1 Lidar data classification techniques: Some of the main interests at FOI concerning processing of lidar data are detection and reconstruction of buildings. For this purpose, three classification techniques have been developed, each with its own merits and disadvantages. One technique (Brandin and Hamrén, 2003, Tolt *et al.*, 2006) segments the data using multiple echo information and classifies the resulting segments using shape and height variation features. Another technique (Tolt *et al.*, 2006) detects large regions containing no ground hits (using the fact that laser beam often penetrate canopies and hits the ground below) and mark the ones containing flat surfaces (i.e. roofs) as potential buildings. A third technique (Ruhe and Nordin, 2007) performs a connected component segmentation of the data and uses features from Principal Component Analysis to classify objects.

2.3.2 Decision-level fusion of classifier outputs: The above techniques have been all developed using a data from a particular system and for a limited geographic region. Hence, when applying them to new data, possibly from another system and containing previously unseen types of objects, their performance may degrade. However, they generally do a quite good job at detecting buildings, at least after finetuning/re-training of the parameters. It has also been noticed that the classifiers often make different errors, which is actually not unexpected since they are based on different ideas of segmentation and feature extraction.
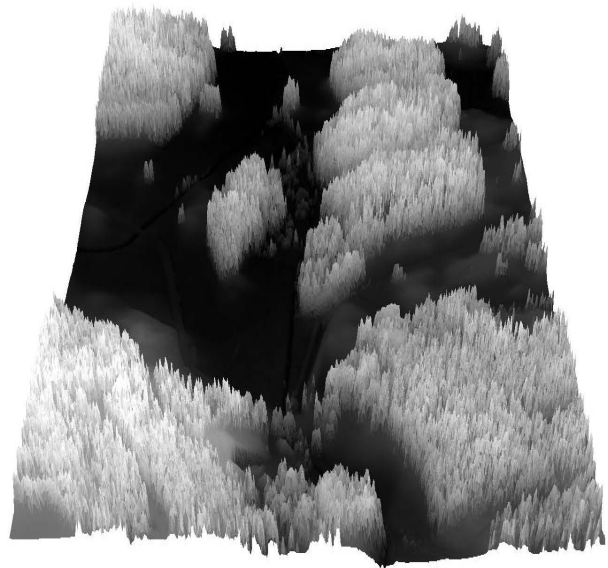

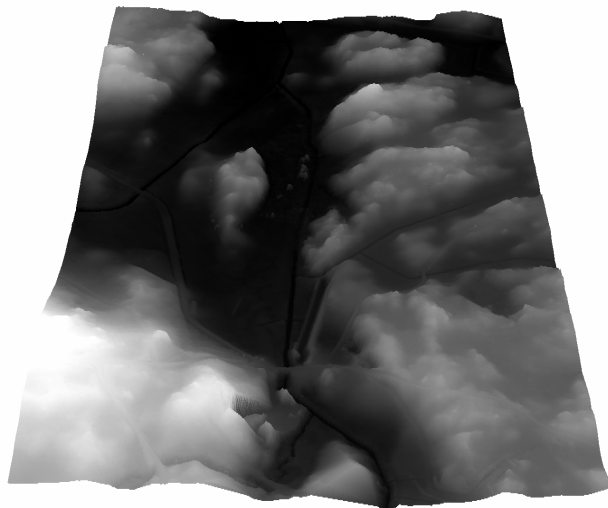Figure 2. Digital Surface Model (DSM) of the region under study.

Figure 3. 3-D visualization of the final bare-earth surface model (DTM) corresponding to the region in Figure 2.

These observations have led to the idea of fusing the results of the individual classifiers to improve the final classification result. In this work, a straightforward majority voting technique was used, which lead to the cancellation of many of the errors (see an example in Figure 4 and Figure 5).

3. REFINING CLASSIFICATION USING AERIAL IMAGES

In this section we present some initial work on updating an existing classification result, here obtained through analysis of sensor data as discussed above but potentially also in the form of a conventional map, using aerial RGB images. In this paper we focus on some quite common problems that we encountered in the lidar-based classification: to correctly classify flat, trimmed vegetation (that may often be mistaken for buildings) and buildings with "weird-looking" roofs (whose spatial properties may make them escape detection).
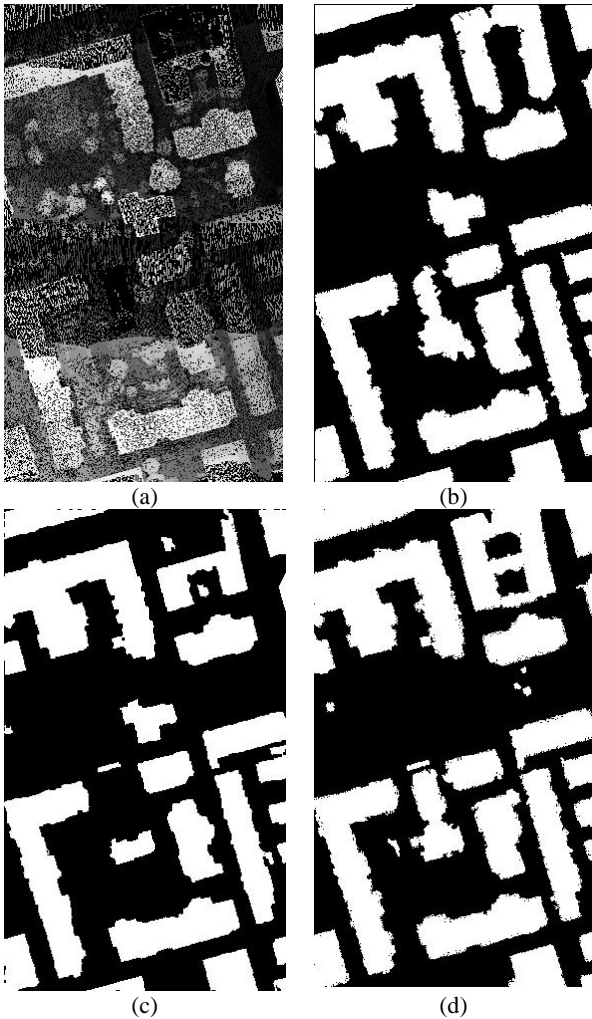
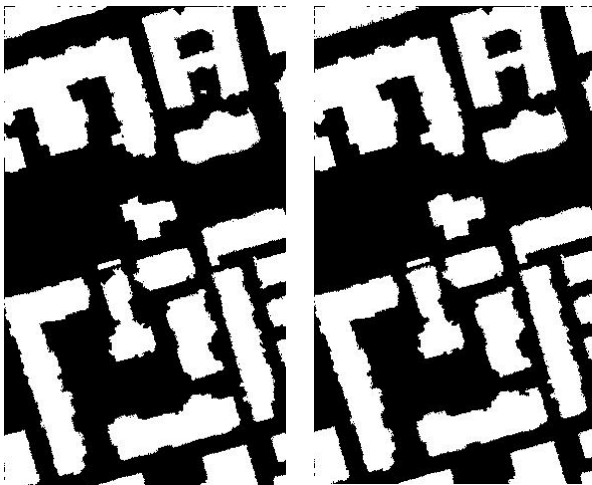Figure 4. (a) Elevation image. (b)-(d) The respective outputs of the three classifiers for the same area.

Figure 5. Left: Result after decision-level fusion of the classifier outputs in Figure 4. Right: Manually corrected classification, showing all buildings in the area. By fusing the result of the individual classifiers, all buildings have been detected and only one small misclassified object remains.

As with the lidar data, the RGB images used in this work have all been acquired with the Topeye Mk II system. The resolution of the images is 10 cm.

3.1 Spectral classification

By using the RGB spectrum in each pixel we want to discriminate between different objects. In this first approach we only consider the spectral information within the image. The measured spectrum will not only depend on the material but also on the current sunlight, the atmosphere and the camera used at the moment. Our considered techniques need spectral models to compare the measure spectrum with. Considering the many parameters affecting the spectrum collected by the camera, nearly all of them are unknown and thus we need a simplified model to estimate the spectrum.

In order to create a spectral model we need some initial spectral information which will make it possible to approximate the model. All these parameters are rarely available so we have to create a local model and consider them all parameters (except the material parameter) fixed. The training data needed can be found using the lidar classification of trees and buildings. As these data generally contain only a few errors, it will give us a good model of the light reflected by the objects.

If a spectrum deviates from our forest spectral model it can be considered an anomaly/outlier to the forest class. The spectra from buildings and other human-made objects will generally deviate from the forest model and will thus be considered an anomaly, if the spectral model of the forest is correct.

Many different models can be considered and have been evaluated before. In this work we have focused on the models consisting of a number of Gaussian components. Each Gaussian component consists of a mean vector and a covariance matrix. The distance from the model is measured as the Mahalanobis distance function. Refer to (Mahalanobis, 1936) for further reading.

3.2 Anomaly detection

First we initialise the model using the result from the previous lidar-based classification to select spectra from the areas classified as trees. In the same way we can create a spectral model buildings using areas classified as buildings. Note that by using the lidar-based classification results, spectra corresponding to measurements of the ground can be removed, as the bare-earth level has been estimated. In this way, spectra corresponding to the ground materials (such as grass and asphalt), do not influence this spectral classification process.

The training data may contain a small amount of faulty spectrums; the outliers are the erroneous classifications. These outliers are later identified and reclassified. Since only areas illuminated by the sun can reflect a spectrum, we only consider pixels with intensity above a threshold. The intersections of illuminated and classified pixels create the training sets needed.

Using the training spectrum we now estimate the two models, for trees and buildings, respectively. Since all training data contains several different materials, each with it own spectrum, we have to use a number of Gaussian components to represent each complete model. The number of Gaussians used depends on the complexity of the training data; buildings usually require a model for each type of roofing material. The training

algorithm can be used to choose the number of components needed.

Assuming a two-class problem involving only trees and buildings we proceed as follows. First we compute the distance from all illuminated pixels to the forest and building models. All pixels that were initially classified as forest but are considered outliers to the forest spectral model are now instead classified as building pixels. And all pixels that were previously classified as building but are considered outliers to the building spectral model *and inliers in the forest spectral model* are now instead considered forest. As a last step, a morphological opening is applied in order to remove spurious "building pixels". The result is an updated and improved classification of the terrain. In Figure 6 and Figure 7 some examples are shown.
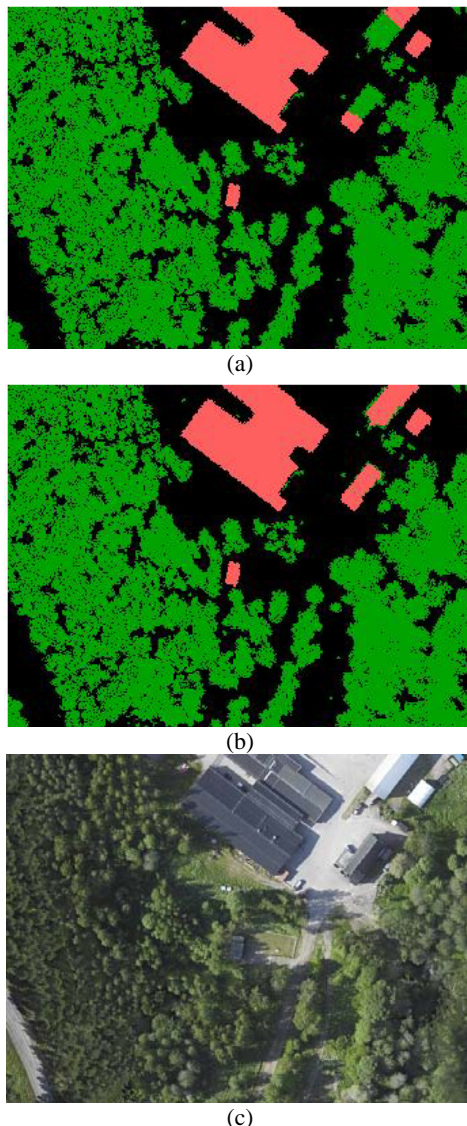

(a)

(b)

(c)

Figure 6. (a) Initial classification result obtained through lidar-based classification. Note the building pixels misclassified as vegetation. (b) Improved classification of the area through spectral analysis. (c) Aerial image.
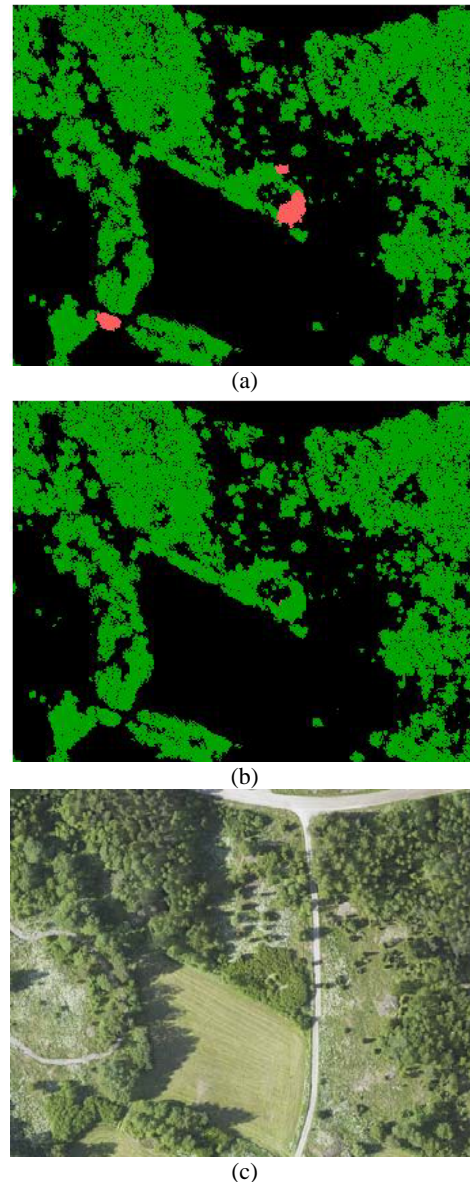

(a)

(b)

(c)

Figure 7. (a) Initial classification result obtained through lidar-based classification. Note the vegetation pixels misclassified as buildings. (b) Improved classification of the area through spectral analysis. (c) Aerial image.

4. CONCLUSIONS

With a goal of producing 3-D models of areas of operation within 48 hours, having fast and automatic data processing methods available is crucial. In this paper, a number of tools for classification of airborne lidar data and aerial images have been presented, starting with bare-earth extraction and classification using lidar data, followed by an approach for classification using RGB images.

First, an approach for bare-earth estimation was presented. It is based on applying a fast bare-earth estimation technique to the data and then check for suspicious-looking artefacts in the resulting bare-earth model. For this purpose, a tool for verifying bare-earth level results was presented, that aims at reducing the amount of time that a user has to spend on inspecting the result and determining which (if any) regions have to be re-processed.

So far, the focus has been on the most frequent type of error, elevation jumps at the border of adjacent tiles, but checks for other errors could also be added, if needed.

Second, a decision-level fusion technique for combining the outputs of several individual classifiers was presented. It is based on the idea that each classifier does a pretty good job in detecting most objects (in this case, building) and that they make different errors. By combining the outputs of the classifiers, many of these errors cancel out. So far, this work has been limited to illustrating the potential the benefits of using such an approach. Plans for future work in this area include assessing classification performance using ground truth data, as well as improving the classification techniques themselves.

Third, a technique for updating existing classification results using aerial imagery was presented. It is based on the assumption that an existing classification map is mostly correct (e.g. after lidar-based classification), but may contain errors. In this paper, the focus has been on some of the most common types of errors found in lidar-based classification, i.e. misclassification of flat, trimmed vegetation into buildings and buildings missed by lidar methods. By using Gaussian models of the simple RGB spectral information is shown that spectral information can affect the confidence of the classification.
Future plans include studying more ways of combining data and information from different sensors (on signal-level, decision-level, etc.), for improving the terrain classification performance further.

ACKNOWLEDGEMENTS

REFERENCES

Ahlberg, J., Renhorn, I., 2004. Multi- and Hyperspectral Target and Anomaly Detection. Scientific report, Swedish Defence Research Agency, FOI-R--1526--SE.

Brandin, M., Hamrén, R., 2003. Classification of Ground Objects Using Laser Radar Data. Master Thesis, Linköping University, Sweden, LITH-ISY-EX-3372-2003.

Elmqvist, M., 2002. Ground surface estimation from airborne laser scanner data using active shape models. In: *Photogrammetric Computer Vision – ISPRS Commission III Symposium*, vol. XXXIV, part A, pp. 114-118.

Haala, N., Brenner, C., 1999. Extraction of buildings and trees in urban environments. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, pp. 130-137.

Landgård, J., 2005. Segmentering och klassificering av Lidar-data. Master Thesis, Linköping University, Sweden, LITH-ISY-EX-YY/3729-SE.

Maas, H.-G., 1999. Fast determination of parametric house models from dense airborne laserscanner data. In: *The International Archives of the Photogrammetry and Remote Sensing*, Bangkok, Thailand, Vol. XXXII, 2/W1, pp. 1–6.

Mahalanobis, P .C., 1936. On the generalized distance in statistics, In: *Proceedings of the National Institute of Science of India 12*, pp 49-55.

Pfeifer, N., Rutzinger, M., Rottensteiner, F., Muecke, W., Hollaus, M., 2007. Extraction of building footprints from airborne laser scanning: Comparison and validaton techniques. In: *Proc. of Urban Remote Sensing Joint Event*, Paris, France, pp. 1- 9.

Rottensteiner, F., Trinder, J., Clode, S., Kubik, K., 2005. Using the Dempster–Shafer method for the fusion of LIDAR data and multi-spectral images for building detection. *Information Fusion*, 6(4), pp. 283-300.

Ruhe, J., Nordin, J., 2007. Classification of Points Acquired by Airborne Laser Systems. Master Thesis, Linköping University, Sweden, LITH-ISY-EX--07/3874--SE.

Schenk, T., Csathó, B. 2002. Fusion of LIDAR data and aerial imagery for a more complete surface description. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXIV 3/W6 pp. 310–317.

Schowengerdt, R. A., 1997. Remote Sensing Models and methods for image processing (Second Edition). Academic Press.

Sithole G., Vosselman, G., 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. In: *ISPRS Journal of Photogrammetry & Remote Sensing* 59, pp. 85 –101

Tolt, G., Persson, Å., Landgård, J., Söderman, U., 2006. Segmentation and classification of airborne laser scanner data for ground and building detection. In: *Proceedings of SPIE Defense and Security Symposium,* Orlando, FL, Vol. 6214, Laser Radar Technology and Applications XI.

Tóvári, D., Vögtle, T., 2004. Object Classification in Laserscanning Data. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Freiburg, Germany, Vol. XXXVI, Part 8/W2, pp. 45-49

Voegtle, T., Steinle, E., 2003. On the quality of object classification and automated building modeling based on laserscanning data. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Dresden, Germany, Vol. XXXIV, 3/W13, pp. 149–155.