Combining Hyperspectral and LiDAR Data for Building Extraction using Machine Learning Technique

DR SEYED YOUSEF SADJADI, SAEID PARSIAN
Geomatics Department, School of Engineering
Tafresh University
Tafresh
IRAN
sysadjadi@hotmail.com, parsian_saeid@yahoo.com

Abstract: - In this study, the fusion of hyperspectral and LiDAR data was used to propose a new method to detect buildings using the machine learning algorithm. The data sets provided by the National Science Foundation (NSF) - funded by the Centre for Airborne Laser Mapping (NCALM)- over the University of Houston campus and the neighbouring urban area, were used. The objectives of this study were: 1) automatic buildings extraction using the hyperspectral and LiDAR fused data (automation), 2) detection of the maximum number of listed buildings in the study area (completeness), and 3) achieving high accuracy in building detection throughout the classification procedure (accuracy and precision). After classification of the buildings, a comparison was made between the results obtained by the proposed method and the reference method in this field (Debes et al., 2014). Our proposed method showed a better accuracy for buildings detection in a much shorter time compared to the reference method. The accuracy of the classification was assessed by four parameters of Precision, Completeness, Overall Accuracy and Kappa Coefficient, and the values of 96%, 100%, 99%, and 0.94 were obtained, respectively.

Key-Words: - Building detection, Hyperspectral, LiDAR, Machine Learning, Decision Tree

1 Introduction

Since more than 50 percent of the population lives in urban areas, mapping of different objects in urban areas, and updating these maps to use them in different applications, such as environment monitoring, telecommunication, and urban planning are important [1]. Classification of urban areas using the traditional methods is costly, time consuming, and is very operator demanding. However, remote sensing methods lead to lower cost, time and human errors, by applying various aerial and satellite images, in this regard [2] [3]. In recent years, automatic extraction of urban objects, such as buildings from aerial and satellite images, have gained more attention, therefore different methods have been investigated and proposed.

Aerial and satellite sensors provide different types of images with various characteristics which can be used in various applications. Suitable data set should be selected regarding the application and study area. For instance, two types of remote sensing data have been widely used to extract building boundaries. The first is the hyperspectral data with hundreds of spectral bands to separate the objects with different spectral characteristics. The second is the LiDAR data that contains accurate height information to separate objects with similar spectral characteristics.

Since the buildings can be distinguished by the material of their roofs, the electromagnetic spectrum reflected or emitted from the roofs could be applied to detect the boundaries of buildings. In this regard, hyperspectral data provides rich information about the spectral reflectance of different objects, which can provide a significant role for detecting of buildings, and distinguishing them from the surrounding trees. It should be considered that despite the fact that hyperspectral images provide high spectral information, they cannot be used for separating the objects made up of the same materials, such as the roofs of some buildings, roads, and parking lots. In this case, the objects with the same spectral characteristics might be classified in one class, even if differ in their height. Moreover, since optic sensors are passive sensors, they have some limitations in extraction of the areas covered by clouds.

Unlike hyperspectral sensors, LiDAR which is an active remote sensor can acquire highly accurate data under almost any meteorological conditions, and without an external source of illumination. Moreover, LiDAR data has been widely applied to detect the objects for which the most important characteristic is the height. For instance, objects, such as buildings and parking lots that have the same spectral characteristics can be easily separated by this type of data.

For many applications, the information provided by a specific sensor is not complete. Therefore, the integration of different types of data acquired by different sensors, provides more complete information [4] [5]. In the case of building detection, it is expected to achieve better results by combining the LiDAR and hyperspectral data [6].

The combination of different types of data can be applied on the pixel, signal, feature or decision levels [5] [7]. In signal level fusion, signals from multiple sensors are combined together to create a new signal with a better signal-to-noise ratio than the input signals. In pixel level fusion, the information from different images on pixel level is merged to improve the detection of objects in different tasks, such as segmentation. Feature level fusion consists of integrating features extracted from different images. In this level of fusion, the extracted features from different sensors are combined to create a feature vector for classification using different types of classifiers. In decision level fusion, different datasets are combined at a higher level of integration. In this level of fusion, first the data acquired by each single sensor is classified, and then fusion process is performed [8] [9] [10].

So far, different classification methods have been proposed to classify the aerial and satellite images. The ensemble method is one of the most widely used for image classification. In this method, instead of applying one particular classifier, a series of classifiers are used, and then the average of the results is used to vote the label of a pixel [11]. The decision tree is one of the known learning techniques in satellite image classification, which is also used in this research. In this method, there is no need to preliminarily knowledge about the data [12]. The ensemble method improves the accuracy of a classification by using the integration of several decision trees, in comparison to one decision tree [13] [14]. In this method, for training each decision tree, the train data should be divided between the decision trees. Bagging and boosting methods are used for classification of the training data [15]. The Bagging method was used in this paper to classify the training data.

The purpose of this study was extraction of the buildings, accurately and completely, and this process was done automatically. In this research, the hyperspectral and LiDAR data were combined at the pixel level, and then buildings were extracted through applying the ensemble learning method on the combined data. Finally, the accuracy of the proposed method was compared with the reference method in this field [16].

2 Data

2.1 Study Area and Data Preparation

In this study, two types of aerial imagery were used for building detection. The first one was the LiDAR data acquired by NCALM over Houston's dormitory, and a small part of the neighbouring urban area, in June, 2012 (Fig. 1) [16]. This data set included two elements of wave heights and intensity of recursive waves, in which the average of their density was 4.0 points per square meter. As a result, the image had a spatial resolution of approximately 2.5 meters. The second one was the hyperspectral image acquired by CASI sensor from the same area. CASI captures images in 144 spectral bands in the visible and near-infrared spectral range (380 to 1050 nm).

Fig. 1. Study area.

The train and test data were specified by several experts from the GRSS organisation. To do so, they applied high spatial resolution aerial images [16]. The number of train and test data used in this study is given in Table 1.

Class	Train Data	Test Data
Building	2125	387
Other	10073	2445

Table 1. The number of pixels used as train and test data in this study.

Since an aerial hyperspectral image was used in this study, the radiometric and atmospheric corrections were not performed. Also, the geometric correction had been performed by the GRSS organisation.

3 Methodology

As described in section 2.1, the hyperspectral image acquired from the study area had 144 spectral bands. Most of these bands were associated with noise, and there were high correlations between most of the bands. To be more precise, some of the bands of hyperspectral images do not carry useful information and should be eliminated through a pre-processing procedure [17]. Several methods, such as the Principal Components Analysis (PCA) [18], Independent Component Analysis (ICA) [19], and Maximum Noise Fraction (MNF) [20] have been proposed to reduce the dimensions of hyperspectral images. In this paper, the accuracy of these three methods was first assessed to select the best technique for our study. Then, the LiDAR and hyperspectral images were combined, and an image was produced with 11 layers (7 bands of the hyperspectral image, plus 4 bands of the LiDAR image).

Five different decision trees were established based on different train data to classify the image. The train data was divided between each of the decision trees using the Bagging method and then, each decision tree was defined, in which each tree could classify a given pixel in one of two classes: "Building" or "Other". To decide a final label for each pixel, the results of the five decision trees were combined, and by voting between the results of the five trees, the class of each pixel was determined.

To reconstruct the buildings' boundaries that had any disconnection on the boundaries of their roof, a post-processing procedure was applied using a method called the gap filling [21]. Furthermore, several mathematical morphology operators were defined based on the geometrical concepts, such as size, shape, and orientation. Then, they were used to get a more appropriate boundary of each building, in terms of the visual interpretation. Fig. 2 illustrates a flowchart of the method used in this study.

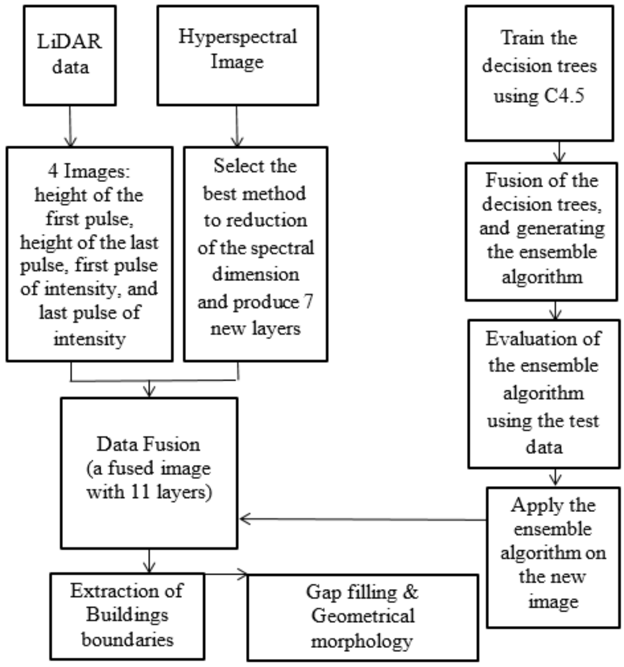

Fig. 2. Flowchart of the proposed method

4 Results and Discussion

This section is provided in 4 subsections to discuss the performance of the proposed method in more details. In subsection 4.1, only hyperspectral image was used to extract the buildings' boundaries. In subsection 4.2, we applied only LiDAR image to detect the building boundaries. In subsection 4.3, both hyperspectral and LiDAR images were combined, and used to achieve better results. Finally, in subsection 4.4, the proposed method is compared to the reference method in this field [16].

4.1 Building Extraction Using Hyperspectral Data

To find the best method for reducing the dimension of the hyperspectral bands, the accuracy of three methods of PCA, ICA, and MNF were evaluated using ensemble supervised machine learning. The results of this evaluation are provided in Table 2. It was concluded that applying 7 layers, obtained using the PCA method, was the most accurate method for building detection using ensemble method in our study. Using these 7 layers, not only reduces noise and correlation, but also had some advantages, such as increasing the speed of the next processes.

Number of layer/Method	PCA	ICA	MNF
3	0.34	0.36	0.40
4	0.52	0.50	0.51
5	0.57	0.54	0.56
6	0.75	0.64	0.70
7	**0.84**	0.71	0.72
8	0.82	0.75	0.74
9	0.82	0.75	0.75
10	0.77	0.73	0.77
11	0.75	0.72	0.78
12	0.73	0.69	0.79
13	0.69	0.64	0.80
14	0.63	0.62	0.80
15	0.60	0.63	0.58

Table 2. Comparison between PCA, ICA and MNF for building detection using ensemble method.

Fig. 3 and Table 3 shows the results of the hyperspectral image classification using 7 layers obtained from PCA method. According to the results, it was concluded that applying only hyperspectral data had some basic limitations for building detection. Hyperspectral images with high spectral resolution can easily be used to classify several objects, such as trees and water area, which are recognised by their reflectance information. However, due to the large variability in the type of buildings and their roofs, any classification based only on spectral information could result in an inaccurate classification. For instance, since buildings and parking lots have almost the same spectral information, they will be classified in the same class, which is not correct. Or, in some parts of the roof of buildings, which are brighter than the other parts, they will be classified into two different classes. Or, if a part of the hyperspectral image is covered by clouds, it will pose some limitations in the classification procedure.

Fig. 3. Classification of the hyperspectral image using machine learning method.

Precision (%)	Completeness (%)	Overall Accuracy (%)	Kappa Coefficient
29	98	67	0.30

Table 3. The accuracy values for building extraction using hyperspectral image.

4.2 Building Extraction Using LiDAR Data

The most stabile feature to describe buildings is their height. Since LiDAR data provides accurate elevation information with high spatial resolution, it is more appropriate than the hyperspectral data for detection of the buildings. Fig. 4 illustrates the results of the classification, obtained by applying the proposed method on LiDAR data. Table 4 provides information about the statistical parameters of the classification results. Comparing the results provided in Table 4 and Table 3, it was concluded that LiDAR data is more appropriate to extract buildings compared to the hyperspectral image. However, to achieve a higher accuracy for extraction of the building boundaries, we needed to use both types of data.

Fig. 4. Classification of the LiDAR image using machine learning method.

Precision (%)	Completeness (%)	Overall Accuracy (%)	Kappa Coefficient
78	100	96	0.83

Table 4. The accuracy values for building extraction using LiDAR image.

4.3 Building Extraction Using Fused LiDAR and Hyperspectral Data

The combination of the best characteristics of different types of remote sensing data is a common method in remote sensing science, which often leads to better results. In this study, integration of hyperspectral and LiDAR images were used to benefit from both geometric information of LiDAR data and the spectral characteristics of hyperspectral images to extract the boundaries of buildings. The elevation of a building is higher than the surrounding objects, and the boundary of a building has a sudden elevation change, which can be detected appropriately using LiDAR images. On the other hand, trees can also be classified as buildings, because of their high elevation. However, this problem can be easily solved by using the hyperspectral images. Fig. 5 shows the results of classification by applying the proposed method on combined LiDAR and hyperspectral images.

Fig. 5. Classification of the fused LiDAR and hyperspectral data using machine learning method.

The evaluation parameters (precision, completeness, overall accuracy, and kappa coefficient) obtained by applying the machine learning technique on combined hyperspectral and LiDAR images are given in Table 5. It was concluded that all of the evaluated parameters had considerable improvement compared to cases using either LiDAR or hyperspectral images. As an example, a 67% and an 18% incensement were observed in Precision through using combined data compared to use of only hyperspectral and only LiDAR data, respectively.

Precision (%)	Completeness (%)	Overall Accuracy (%)	Kappa Coefficient
96.56	100	98.66	0.94

Table 5. The accuracy values for building detection using the fusion of LiDAR and hyperspectral images.

It is worth mentioning that all of the steps in our proposed method were automated and none of the steps were provided by an operator (Automation). Moreover, as shown in Table 5, building extraction completeness was 100%, which indicated that the proposed method extracted the maximum numbers of buildings. Classification accuracy and precision were also high.

4.4 Comparison between the Proposed Method and the Reference Method

The best method in IEEE GRSS Data Fusion Contest using both hyperspectral and LiDAR data, which is known as the reference method in this field, was introduced by Reference 16. Since in this study, we used the same data set, a comparison was made between the proposed method and the reference method (Table 6). It is worth noting that Reference 16 classified the combined image in 15 different classes. However, the focus of this research was only on building detection.

Quality Control Parameters	Precision (%)	Completeness (%)	Overall Accuracy (%)	Kappa Coefficient
Proposed method	96.56	100	98.66	0.94
Reference method	94.10	100	96.40	0.85

Table 6. Comparison between the results obtained from the proposed and reference methods.

According to this Table, it was concluded that the method presented in this study for building detection was more accurate than the reference method. All of the evaluated parameters had almost higher values compared to the reference method. Also, it should be noted that since our method applies several simple machine learning techniques, it operates in a shorter time in comparison with the reference method.

5 Conclusion

In this study, a method was developed to extract buildings' boundaries using the combination of two types of aerial imagery (LiDAR and hyperspectral). The most useful characteristics of each data set regarding the building detection were considered to achieve a more accurate classification. LiDAR data contains only the height and intensity of backscatters, and has a high geometric accuracy. On the other hand, hyperspectral images contain hundreds of spectral bands, which can be used to distinguish different types of objects with various spectral characteristics. Therefore, the LiDAR and hyperspectral images can be considered complementary components in remote sensing classification procedures. In this paper, the ensemble learning technique was applied on hyperspectral image, LiDAR data, as well as the combined data of hyperspectral and LiDAR data to extract buildings. The results were as follows: (1) If the proposed method is applied only on the hyperspectral image, many objects such as parking areas will be classified incorrectly. Since the main parameter for building detection is elevation information, using only hyperspectral image will not be resulted in an accurate classification. The Overall Accuracy and Kappa Coefficient were 67% and 0.3, respectively. (2) If the machine learning method is applied on LiDAR data, the results will be better than the first case. However, these results will not be extremely accurate, due to not considering the spectral characteristics of different objects in urban areas. In this case, the values of 96% and 0.83 were achieved for Overall Accuracy and Kappa Coefficient, respectively. (3) If LiDAR and hyperspectral data are combined, and the proposed method is applied on the combined image, the results will be better than the first and second cases. The values of 99% and 0.94 were obtained for Overall Accuracy and Kappa Coefficient, respectively. Furthermore, by comparing the proposed method and the reference method [16] for building detection, it was concluded that our method had a higher accuracy than the reference method in this field. Also, we concluded that if the number of decision trees decrease, the accuracy decreases, as well. On the other hand, if the number of decision trees increases, classification will lead to over fitting because the training data, which is divided between the trees, will not be enough. Finally, to have a more versatile evaluation on our proposed method, a parameter called qualitative accuracy was used to visually compare between the extracted buildings using the proposed method and the reference method [16]. The results showed that in addition to the quantitative accuracy, authenticity and quality of the proposed method were better than the reference method [16].

References:

[1] Jensen, J. R., & Cowen, D. C. (1999). Remote sensing of urban/suburban infrastructure and socio-economic attributes. Photogrammetric engineering and remote sensing, 65, 611-622.

[2] Gruen, A., Baltsavias, E., & Henricsson, O. (Eds.). (2012). Automatic extraction of man-made objects from aerial and space images (II). Birkhäuser.

[3] Mayer, H. (1999). Automatic object extraction from aerial imagery—a survey focusing on buildings. Computer vision and image understanding, 74(2), 138-149.

[4] Simone, G., Farina, A., Morabito, F. C., Serpico, S. B., & Bruzzone, L. (2002). Image fusion techniques for remote sensing applications. Information fusion, 3(1), 3-15.

[5] Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion in remote sensing: concepts, methods and applications. International journal of remote sensing, 19(5), 823-854.

[6] Simental, E., Ragsdale, D., Bosch, E., Dodge Jr, R., & Pazak, R. (2003, May). Hyperspectral dimension reduction and elevation data for supervised image classification. In Proceedings, ASPRS Annual Conference, Anchorage, AK-USA.

[7] Esteban, J., Starr, A., Willetts, R., Hannah, P., & Bryanston-Cross, P. (2005). A review of data fusion models and architectures: towards engineering guidelines. Neural Computing & Applications, 14(4), 273-281.

[8] Du, P., Liu, S., Xia, J., & Zhao, Y. (2013). Information fusion techniques for change detection from multi-temporal remote sensing images. Information Fusion, 14(1), 19-27.

[9] Dong, J., Zhuang, D., Huang, Y., & Fu, J. (2009). Advances in multi-sensor data fusion: algorithms and applications. Sensors, 9(10), 7771-7784.

[10] Zhang, Y. (2004). Understanding image fusion. Photogrammetric engineering and remote sensing, 70(6), 657-661.

[11] Brill, E., & Wu, J. (1998, August). Classifier combination for improved lexical disambiguation. In Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1 (pp. 191-195). Association for Computational Linguistics.

[12] Mitchell, T. M. (1997). Machine Learning (International Edition). Computer Science Series. McGraw-Hill, New York.

[13] Polikar, R. (2006). Ensemble based systems in decision making. Circuits and systems magazine, IEEE, 6(3), 21-45.

[14] Rokach, L. (2010). Ensemble-based classifiers. Artificial Intelligence Review,33(1-2), 1-39.

[15] Polikar, R. (2006). Ensemble based systems in decision making. Circuits and systems magazine, IEEE, 6(3), 21-45.

[16] Debes, C., Merentitis, A., Heremans, R., Hahn, J., Frangiadakis, N., van Kasteren, T., ... & Philips, W. (2014). Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest. Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of, 7(6), 2405-2418.

[17] Lee, C., & Landgrebe, D. (1993). Feature extraction and classification algorithms for high dimensional data.

[18] Jutten, C., & Herault, J. (1991). Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. Signal processing,24(1), 1-10.

[19] Boardman, J. W., & Kruse, F. A. (1994). Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada. In Proceedings of the Thematic Conference on Geologic Remote Sensing(Vol. 1, pp. I-407). Environmental Research Institute of Michigan.

[20] Pitas, I. (2000). Digital image processing algorithms and applications. John Wiley & Sons.