PERFORMANCE ASSESSMENT OF AUTOMATED FEATURE EXTRACTION TOOLS ON HIGH RESOLUTION IMAGERY

Daniel A. Lavigne, Defence Scientist
Defence Research and Development Canada – Valcartier
2459, Pie-XI Blvd. North
Quebec (Quebec), CANADA G3J 1X5
daniel.lavigne@drdc-rddc.gc.ca

Gang Hong
Yun Zhang
Department of Geodesy and Geomatics Engineering
University of New Brunswick
15 Dineen Dr., P.O. Box 4400
Fredericton (New-Brunswick), CANADA E3B 5A3
gang.hong@unb.ca
yunzhang@unb.ca

ABSTRACT

The enormous increase in the volume of remotely sensed data that is being acquired by an ever-growing number of earth observation sensors, led to the situation where the image analysts are at risk to be overwhelmed by the number of images to process. A way to cope with this situation is to use automated feature extraction (AFE) tools to help the image analysts in the extraction of geospatial features, and address specific tasks such as object detection for change detection purposes. However, it is not clear which AFE tools perform better against others, and on what scenarios. This paper presents an assessment of the performance of three AFE tools to extract man-made objects for change detection on high resolution airborne and spaceborne electro-optic imagery: Feature Analyst (Visual Learning Systems Inc.), eCognition (Definiens AG), and Genie Pro (U.S. Los Alamos National Laboratory). These tools all intend to assist the image analysts in extracting relevant image features using dissimilar feature extraction algorithms. Feature Analyst employs machine learning techniques to extract user-defined features; eCognition is an object-oriented classification package using a fuzzy logic scheme; Genie Pro is an evolutionary computation software system assembling image processing tools from a collection of image operators by means of genetic algorithms. The performance of these automated feature extraction software packages in extracting relevant geospatial features on high resolution remotely sensed images is assessed, together with their use for change detection applications.

INTRODUCTION

An increasing number of commercial airborne and spaceborne sensor systems are now generating large volumes of remotely sensed data, at high resolutions and with more spectral channels than ever before. Traditionally, image analysts achieve the exploitation of these data manually. However, the sheer volume of data accumulating from a growing variety of remote sensing sources and feature extraction techniques defies human analysis. Indeed, while the creation and development of task-specific feature extraction algorithms is mandatory, the resulting extraction process can be extremely expensive and thus, often requires a significant investment of time and effort so specialized that new algorithms must be defined or generated for each new image content to address. Making sense of these data within a reasonable time frame, mandates the development of automated image exploitation tools that can quickly and reliably extract meaningful geospatial information for numerous scenarios (O'Brien, 2004). Furthermore, the exploitation of these data on a large scale shall also benefits from the use of such tools to extract specific features of interest from the imagery. By using automated feature extraction (AFE) technology, extracted models over a small training set can then be applied to larger areas, reducing the extraction time required by several orders of magnitude.

Among the algorithms developed in order to deal with these issues, particular areas of interest are pixel-based and object-oriented classification tools that are able to produce overlays for the location, identification, and delineation of precise geospatial features within remotely sensed imagery. Such classification tools can be used by the image analysts for a variety of applications such as the detection of changes of a specific geospatial location, within a set of multi-temporal remotely sensed images.

With the constant improvement of spatial and spectral resolutions of commercially available airborne and spaceborne sensors, the use of pixel-based and object-based change detection algorithms coping with man-made objects becomes clever. However, it is not clear which AFE tools perform better, and on what scenarios. This paper presents a comparative performance analysis of three AFE tools employed for the extraction of man-made objects in high-resolution airborne and spaceborne electro-optic imagery, and used for change detection.

AUTOMATED FEATURE EXTRACTION TOOLS

Three AFE software tools, which have high potential of producing cost-effective geospatial extraction results within a change detection framework, have been selected for this evaluation: Feature Analyst (Visual Learning Systems Inc.), eCognition (Definiens AG), and GENIE Pro (U.S. Los Alamos National Laboratory). All these AFE tools embedded unique machine learning technologies that intend to assist the image analysts, in a supervised manner, by extracting the characteristics of relevant geospatial features using dissimilar feature extraction techniques.

Feature Analyst

Feature Analyst employs machine learning algorithms to extract geospatial features. Available as extensions to ArcGIS™, ERDAS IMAGINE™, GeoMedia™, SOCET SET™, and RemoteView™ soon, it uses temporal, spatial, spectral, and ancillary information (such as size, shape, color, pattern, texture, shadow, and spatial association) to model the feature extraction process (O'Brien, 2003a). As a pixel-based AFE tool, its classification scheme incorporates a contextual classifier, set by the user accordingly to the spatial distribution of the image features to be extracted. Small and simple set of user-defined training examples is employed within a hierarchical learning procedure, embedded to iteratively improve the classification results. Classified geospatial features are available for change detection purposes using a post-classification comparison technique.

eCognition

eCognition is a stand-alone object-oriented classification package for the classification of imagery (Benz, 2004), based on attributes of the image objects and their mutual relations, rather than on the attribute of individual pixels. Within a multi-scale image analysis framework, and while segmenting the image into spectrally homogenous objects, eCognition uses geometric characteristics and topological properties to perform sample-based or knowledge-based supervised image classification (Benz, 2001). A multi-scale automated segmentation technique is incorporated to turn raster information into meaningful objects. Rule-based classification is employed where the manifold parameter set of objects characterizes the classes and features by means of fuzzy logic and geospatial relations. Once defined by the user in an assisted fashion, classification rules can then be applied to further images, enabling automation of the image analysis process. Sample-based object-oriented fuzzy classification method can be used, together with a post-classification algorithm, to realize change detection.

Genie Pro

Genie Pro is an evolutionary computation software produced by the ISIS group at Los Alamos National Laboratory. Using training inputs provided by the user, it derives automatic pixel classification algorithms for remotely sensed imagery, assigning labels to pixels (Harvey, 2002). Genie Pro employs genetic algorithms to assemble image processing tools from a collection of image operators, and is used primarily for deriving vector overlays and semantically meaningful maps. It integrates spectral information and spatial cues such as texture, local morphology and large-scale shape information in a sophisticated way (Perkins, 2005). It was design at first for detecting complex spatio-spectral terrain features in multispectral imagery (O'Brien, 2003b). Post-classification comparison algorithm can be employed for the computation of change detection.

METHODOLOGY

Selected AFE tools are evaluated for their performance in extracting man-made features in airborne and spaceborne imagery. Extracted features of a pair of multi-temporal imagery are then used to identify changes to the landscape caused by human activities. It is anticipated that the performance of each AFE tool in the delineation of man-made objects, using distinct machine learning techniques, will directly condition the final change detection results.

The objects of interest considered in this research include changes involving aircrafts, buildings, and vehicles. For each change image, a change detection error matrix (Macleod, 1998) is computed from two multi-temporal images of the same geospatial location in the following manner:

1. Two multi-temporal images are registered to a common reference system.
2. The registered images are then classified with the AFE tool, according to selected objects of interest.
3. Change detection is performed on the two classified images.
4. Sample points are collected from the change image, and then evaluated.
5. A change detection error matrix is computed using the results from the evaluation of sample points.

The change detection error matrix provides six possible combinations of Change/No Change (table 1).

Table 1. Change/No Change combinations.

Case #	Reference Status	Classification Status	Error Type
1	True no change	Correctly classified no change class	No error
2	True change	Correctly classified change class	No error
3	True no change	Incorrectly classified no change class	Classification error
4	True no change	Incorrectly classified change class	Commission error
5	True change	Incorrectly classified no change class	Omission error
6	True change	Incorrectly classified change class	Classification error

Four accuracy values are determined and used as metrics to assess the performance of the change detection process, namely: overall accuracy, kappa statistic, producer's accuracy, and user's accuracy:

Producer's accuracy – indicates the probability of correctly detected pixels compared to reference pixels and is a measure of omission error.

User's accuracy – is the probability of correctly detected pixels compared to all pixels detected in the same category and is a measure of commission error.

Overall accuracy – is computed by dividing the total correctly classified pixels by the total number of pixels being checked in the error matrix.

Kappa statistics – incorporates the off-diagonal elements as a product of a row and column marginal and is a measure of how well the remotely sensed changes based on classification agrees with the changes based on reference data.

Producer's accuracy and user's accuracy are computed and are also expressed in terms of omission error and commission error, respectively. Overall accuracy (Congalton, 1999) is used to quantify the change detection performance of each AFE tool.

EXPERIMENTS

Image Dataset

A dataset of six remotely sensed images, segmented along three scenarios (two for each scenario), is used to assess the change detection performance of the AFE tools. Figure 1 illustrates a sample image for each of the three scenarios tested, while table 2 presents their characteristics. The first image pair involves airborne images of Montreal Trudeau airport (0.056 meter), acquired within a span of a few seconds. The objects of interest to be extracted are aircrafts, where two of them have moved between the acquisition times. The second image pair fused an IKONOS image (1 meter) with a QuickBird one (0.64 meter) from CFB Valcartier, acquired three years apart. The objects of interest to be extracted from this image pair are buildings. The last image pair is a set of two airborne images of a parking lot, located at the Port of Quebec (0.112 meter). Few seconds separate the two image acquisitions. The objects of interest to be extracted are vehicles.

(a) (b) (c)

Figure 1. Imagery samples of the dataset. (a) Airborne imagery of Montreal Trudeau airport, (b) QuickBird imagery of CFB Valcartier, and c) airborne imagery of Port of Quebec.

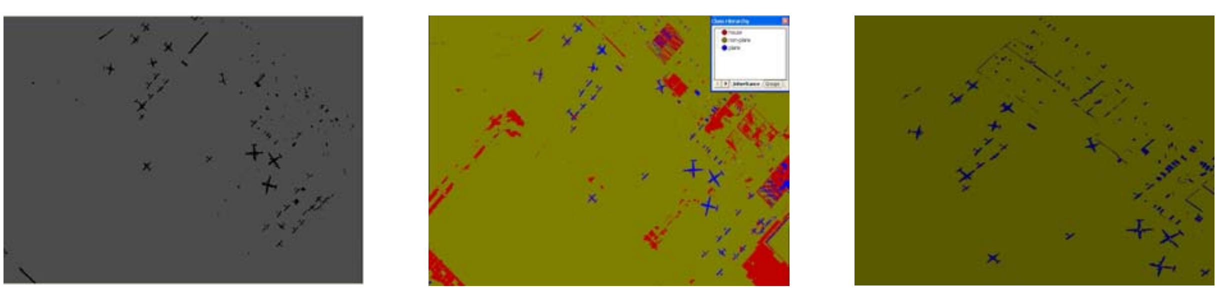
Table 2. Three image pairs used for the evaluation of AFE tools

Image pair #	Objects	Location	Sensors	Image Resolution
1	Aircrafts	Montreal, Quebec	Airborne (t_1) Airborne (t_2)	0.056m 0.056m
2	Buildings	Quebec city, Quebec	IKONOS QuickBird	Pan 1m (Multi 4m) Pan 0.64m (Multi 2.55m)
3	Vehicles	Quebec city, Quebec	Airborne (t_1) Airborne (t_2)	0.112m 0.112m

Aircrafts Change Detection

Image pair #1 of Montreal Trudeau airport (table 2) is used for the evaluation the AFE tools for aircrafts change detection. Figure 2 illustrates the results of aircrafts extracted with the three AFE tools on the first airborne image.

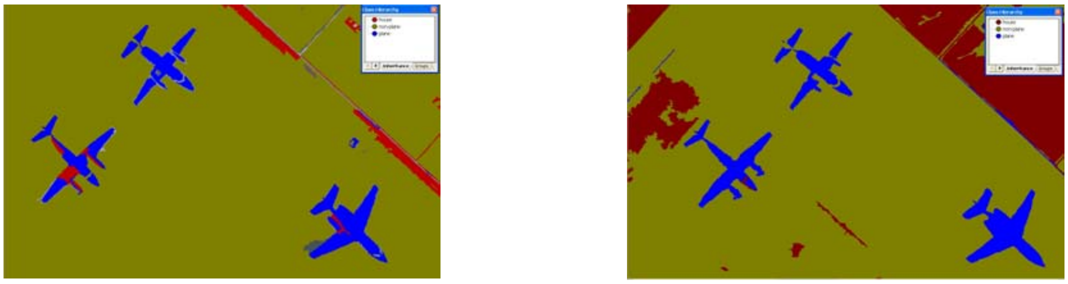
Feature Analyst. Feature Analyst classified the imagery using a two-classes object scheme: *aircraft* and *non-aircraft* (fig. 2a). All aircrafts in the scene have been extracted, while some *non-aircraft* pixels were incorrectly classified as *aircrafs* (false alarms). Shadows of the aircrafts were not extracted. To compute the change detection error matrix, 230 random sample points were generated and stratified so the samples would be distributed among three possible change situations: *aircraft to non-aircraft*, *non-aircraft to aircraft*, and *aircraft to aircraft* (no change). 17 samples represented real change of aircraft status: 10 for the *aircraft to non-aircraft*, and 7 for the *non-aircraft to aircraft*. 11 samples which were non-aircraft were incorrectly labeled as aircraft in both dates. Overall accuracy was 37.4%, with K_{hat} = 19.8%. There exists a 46.5% omission error, and 13.8% commission error.



(a) (b) (c)

Figure 2. Aircrafts detection results using (a) Feature Analyst, (b) eCognition, and (c) Genie Pro.

eCognition. Using eCognition, the airborne images was first segmented using an appropriate scale parameter. Three classes were then created: *aircraft*, *non-aircraft*, and *house* (fig. 2b). Aircraft samples are then selected for the classification of the imagery, using standard nearest neighbor. Classification results showed good extraction performance, where all aircrafts have been extracted. Adding the *house* class helped extract the aircraft objects alone. However, many false alarms exist in the final classification results of the image pair (fig. 3): some pixels were incorrectly classified as *aircraft* and *non-aircraft*. Accuracy assessment was implemented with 235 samples, randomly generated and stratified. 8 samples represented a real change of aircraft status for the two cases *aircraft to non-aircraft* and *non-aircraft to aircraft*. 22 samples which were non-aircraft were incorrectly identified as aircraft in both dates. Overall accuracy was 26.4%, with K_{hat} equals 10.4%. Omission and commission errors are 55.4% and 29.0% respectively.



(a) (b)

Figure 3. Aircraft classification results using eCognition. (a) Classified aircrafts in the first image, (b) and the second image of the image pair.

Genie Pro. All aircrafts in the two airborne images were correctly extracted using a two-classes object scheme (Harvey, 2004), similar to the one used with Feature Analyst. However, Genie Pro incorrectly classified many pixels that originate from hangars and vehicles as belonging to the *aircraft* class. For this scenario, Genie Pro was less robust than Feature Analyst to discriminate between *aircraft* and *non-aircraft* pixels. 230 samples were generated and stratified to compute the change detection error matrix, distributed among the three change situations. 10 samples represented real change of aircraft status: 7 for the *aircraft to non-aircraft*, and 3 for the *non-aircraft to aircraft*. 13 *non-aircraft* samples were incorrectly identified as *aircraft* in both dates. Overall accuracy was 31.3%, with K_{hat} = 12.9%. There exists a 55.4% omission error, and 17.3% commission error.

Table 3 presents, for the three AFE tools, the associated computed producer's accuracy and user's accuracy, with their omission error and the commission error, respectively. Feature Analyst performs with a higher overall accuracy than eCognition and Genie Pro, with lowest values of omission and commission errors. Among the three tested scenarios, aircraft change detection was the most successful one, as quantified by the overall accuracy values.

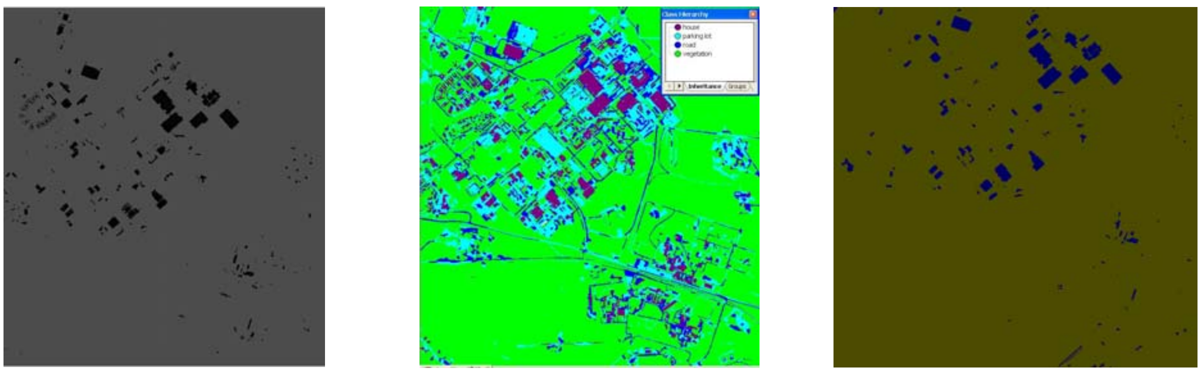
Table 3. Aircraft Change Detection error matrix.

Change		**Feature Analyst**		**eCognition**		**Genie Pro**	
		Producer's accuracy Omission error	User's accuracy Commission error	Producer's accuracy Omission error	User's accuracy Commission error	Producer's accuracy Omission error	User's accuracy Commission error
from	to						
Aircraft	Aircraft	53.5% 46.5%	86.2% 13.8%	44.6% 55.4%	71.0% 29.0%	44.6% 55.4%	82.7% 17.3%
Non-aircraft	Non-aircraft	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%
Aircraft	Non-aircraft	100% 0%	13.2% 86.8%	100% 0%	5.1% 94.9%	100% 0%	9.3% 90.7%
Non-aircraft	Aircraft	100% 0%	9.5% 90.5%	100% 0%	4.9% 95.1%	100% 0%	3.7% 96.3%
Overall accuracy		37.4%		26.4%		31.3%	
K_{hat}		19.8%		10.4%		12.9%	

Buildings Change Detection

The second image pair is an urban scenario where the objects of interest to be extracted are buildings. It consists in a pair of an IKONOS image fused with a QuickBird one. Figure 4 shows the extraction results obtained for each AFE tool.

Feature Analyst. The extraction of buildings is a complicated process when no ancillary data are provided. In this scenario, the spectral signature of the buildings is similar to some extend to other image features such as parking lot and road. Therefore, in a first time, Feature Analyst was employed to extract roads and vegetation in order to generate road and vegetation masks. Then, using these masks, buildings were extracted. Computing the change detection error matrix involved the use of 226 samples that were randomly generated and stratified according to three change categories: *non-building to building*, *building to non-building*, and *building to building* (no change). 3 samples from the change image represented a real building change: 2 samples for the *building to non-building* situation, 1 for the *non-building to building* one. No *non-building* sample was incorrectly identified as *building* in both dates. Overall accuracy was 35.8% with K_{hat} equals 15.1%. There exists a 50.3% omission error, and 0% commission error.



(a) (b) (c)

Figure 4. Buildings extraction (QuickBird data) with (a) Feature Analyst, (b) eCognition, and (c) Genie Pro.

eCognition. The classification process used with eCognition to extract buildings was similar to the one employed previously for the aircraft classification. Different scale parameters were selected due to the different resolutions of the IKONOS and QuickBird images. The images were classified into four classes: *building*, *parking lot*, *road*, and *vegetation* (fig. 4b). These additional classes were added because many buildings were incorrectly classified as non-buildings, and vice versa. Once buildings were identified, all three other classes were merged together. Then, images were reclassified according to two categories: *building* and *non-building*. Without elevation information, some buildings were correctly differentiated from other spectrally similar objects, and some non-buildings and buildings were incorrectly classified. 226 samples were randomly generated and stratified according to three change categories: *building to non-building*, *non-building to building*, *building to building* (no change). One sample represented a real building change (fig. 5). Overall accuracy was 31.4% with K_{hat} equals 16.1%. There exists a 44.9% omission error, and 4.1% commission error.

(a) (b)

Figure 5. Real building change between the multi-date (a) IKONOS and (b) QuickBird imagery.

Genie Pro. Using a two-classes classification scheme, less building pixels were detected with Genie Pro than the two other packages (fig. 4c). For the accuracy assessment, 226 samples were randomly generated and stratified according to three change categories: *non-building to building*, *building to non-building*, and *building to building* (no change). Only 1 sample from the change image represented a real building change (*non-building to building*). Overall accuracy was 21.7% with K_{hat} equals 5.0%. There exists a 57.9% omission error, and 2.0% commission error.

Table 4 presents a summary of the computed values for the building change detection error matrix. Feature Analyst showed the best overall accuracy performance amid the three AFE tools. Indeed, No *non-building* sample was incorrectly identified as *building* in both dates. Using four classes, eCognition performed significantly less then did Feature Analyst. The overall accuracy value of Genie Pro was the lowest one; it produced the greater omission error than Feature Analyst and eCognition. Many buildings located in the spaceborne imagery were not extracted (fig. 4).

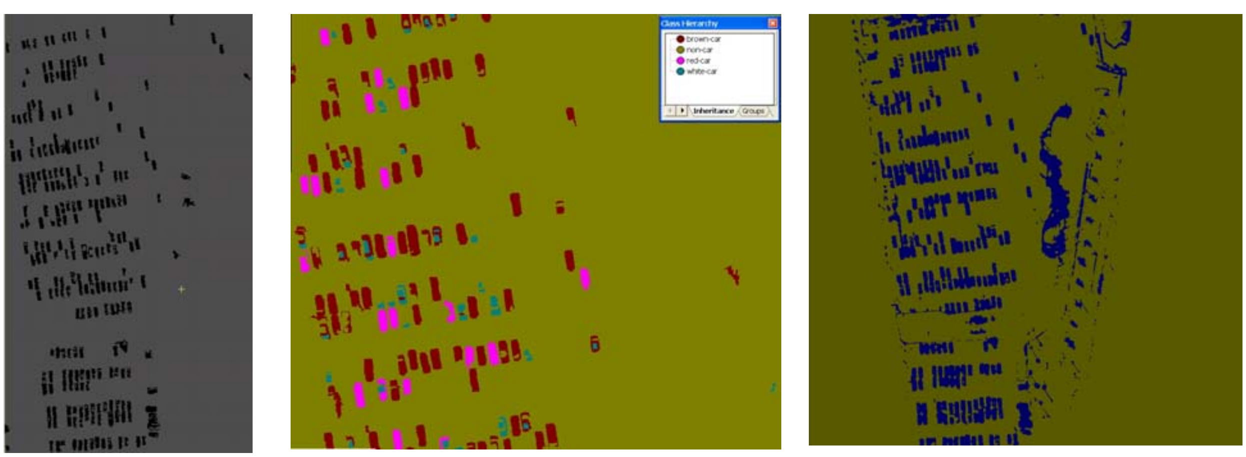
Table 4. Building Change Detection error matrix.

Change		**Feature Analyst**		**eCognition**		**Genie Pro**	
		Producer's accuracy Omission error	User's accuracy Commission error	Producer's accuracy Omission error	User's accuracy Commission error	Producer's accuracy Omission error	User's accuracy Commission error
from	to						
Building	Building	49.7% 50.3%	100% 0%	55.1% 44.9%	95.9% 4.1%	42.1% 57.9%	98.0% 2.0%
Non-building	Non-building	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%
Building	Non-building	100% 0%	2.7% 97.3%	0% 100%	0% 100%	0% 100%	0% 100%
Non-building	Building	100% 0%	1.4% 98.6%	100% 0%	1.5% 98.5%	100% 0%	1.1% 98.9%
Overall accuracy		35.8%		31.4%		21.7%	
K_{hat}		15.1%		16.1%		5.0%	

Vehicles Change Detection

The third set is an airborne image pair of a parking lot, acquired within a span of a few seconds. Objects of interest to be extracted are vehicles; results for each AFE tools are showed in figure 6.

Feature Analyst. Feature Analyst was able to extract most of the vehicles in the imagery, using only a two-classes classification procedure (fig. 6a). Using 246 samples generated and stratified for the accuracy assessment, 6 samples represented a real vehicle change (*non-vehicle to vehicle* situation). 12 samples which were non-vehicles were incorrectly identified as vehicles in both dates. Overall accuracy was 28.1% with K_{hat} equals 12.5%. There exists a 55.9% omission error, and 16.0% commission error.



(a) (b) (c)

Figure 6. Vehicles detection with (a) Feature Analyst, (b) eCognition, and (c) Genie Pro.

eCognition. Similarly to the previous aircraft classification process, the airborne image pair was initially classified into two classes: *vehicle* and *non-vehicle*. However, since the classification results were not acceptable, the images were then reclassified using four classes: *brown-vehicle*, *white-vehicle*, *red-vehicles*, and *non-vehicles* (fig. 6b). Final result is then classified into two categories: *vehicle* and *non-vehicle*. White vehicles presented some problems related to the spectral characteristics of the vehicle body and window parts that are so different that the segmentation algorithms considered them as separate objects. Most of detected changes were false changes. The classification of vehicles was more difficult than for aircrafts and buildings. 244 samples were generated and stratified to build the change detection error matrix. 2 samples represented real change (*non-vehicle to vehicle* situation). Overall accuracy was 25.8% with K_{hat} equals 12.3%. There exists a 53.8% omission error, and 10.2% commission error.

Genie Pro. Genie Pro was able to detect the vehicles using a two-classes classification procedure (fig. 6c). However, many false alarms were produced. Using 246 samples generated and stratified for the accuracy assessment, 2 samples represented a real vehicle change (*non-vehicle to vehicle* situation). 3 samples which were *non-vehicle* were incorrectly identified as *vehicle* in both dates. Overall accuracy was 27.6% with K_{hat} equals 13.1%. There exists a 54.2% omission error, and 4.4% commission error.

Table 5 summarizes the vehicle change detection error matrix for the three AFE tools, including the producers' accuracy, user's accuracy, commission error, and omission error. All three AFE tools showed similar low overall accuracy values; none of the three AFE tools was able to perform significantly better than the other ones.

Table 5. Vehicle Change Detection error matrix.

Change		Feature Analyst		eCognition		Genie Pro	
		Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy	Producer's accuracy	User's accuracy
from	to	Omission error	Commission error	Omission error	Commission error	Omission error	Commission error
Vehicle	Vehicle	44.1% 55.9%	84.0% 16.0%	46.2% 53.8%	89.7% 10.2%	45.8% 54.2%	95.6% 4.4%
Non-vehicle	Non-vehicle	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%
Vehicle	Non-vehicle	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%	0% 100%
Non-vehicle	Vehicle	100% 0%	7.5% 92.5%	100% 0%	2.1% 97.9%	100% 0%	2.3% 97.7%
Overall accuracy		28.1%		25.8%		27.6%	
K_{hat}		12.5%		12.3%		13.1%	

DISCUSSION

Objective of the Research

The objective of this research was to perform a comparative analysis of three AFE tools, for the delineation of man-made objects to be used for change detection purposes. All three AFE tools extracted geospatial features in a supervised fashion, using different machine learning techniques. Feature Analyst used machine learning algorithms to classify pixels, according to the spatial distribution of the image features to be extracted. eCognition employed object-oriented fuzzy logic algorithms to classify image objects. Genie Pro exploited genetic algorithms to assign labels to pixels. Every AFE tool has been evaluated for its capability to be used for change detection applications. For each tested scenario, the accuracy assessment was realized using change detection error matrices. The assessment was expressed in terms of the overall accuracy, producer's accuracy, user's accuracy, commission error, and omission error, employing techniques proposed by (Congalton, 1999). Overall accuracy was used to compare the overall performance of each AFE tool; it indicates the likelihood of any information category being correctly classified. Producer's accuracy and user's accuracy are ways of reporting individual category accuracies. Figure 7 shows the overall accuracy computed for each AFE tools on the three tested scenarios.

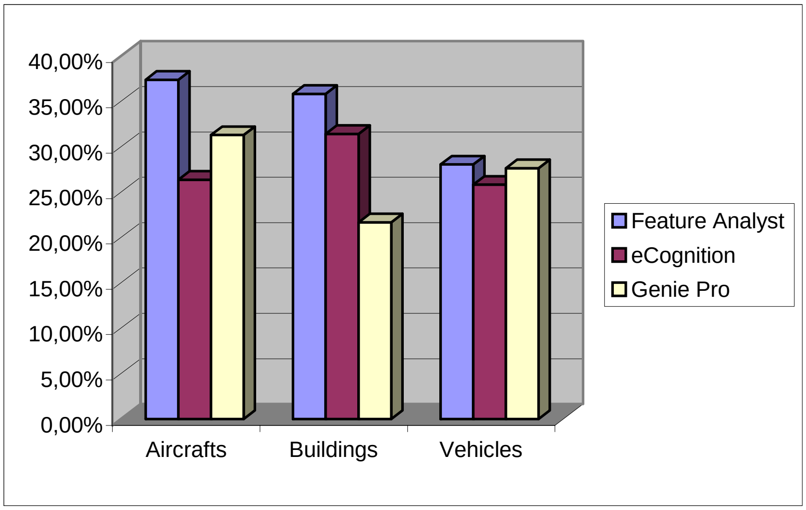


Figure 7. Computed overall accuracy values for each AFE tools on three tested scenarios.

Discussion of the Evaluation

The aircraft scenario was the one presenting the most acceptable results in term of overall accuracy. Feature Analyst and Genie Pro performed well in the extraction of aircrafts, even thought Genie Pro produced more false alarms than Feature Analyst. eCognition showed a much lower overall accuracy for the aircraft scenario than the two other AFE tools. Feature Analyst results were good for building extraction, thus slightly less than for the aircraft scenario. Road extraction and vegetation extraction were needed to generate road and vegetation masks. Using multi-classes segmentation procedure, eCognition showed the best performance for buildings extraction than the other scenarios. Many buildings were not extracted using Genie Pro, with direct repercussion expressed by the low overall accuracy value of the change detection error matrix for the building scenario. The vehicle scenario was obviously the most challenging one, with the lowest computed overall accuracy values: for this scenario, none of the three AFE tools performed significantly better than the others.

Feature Analyst performed relatively well in the extraction of man-made objects such as aircrafts, buildings, and vehicles, with straight consequence on the change detection results obtained in this research. Indeed, the performance of Feature Analyst for aircraft change detection and building change detection was superior than the one obtained by eCognition and Genie Pro. In the case of vehicle change detection, the performance of Feature Analyst was similar to eCognition and Genie Pro. Feature Analyst has several benefits; it holds promise in the classification of high resolution imagery, with relatively few parameters needed to be set. Nonetheless, using a hierarchical learning approach, a good classification solution usually requires several iterations. Moreover, it is not obvious to know when a satisfactory extraction result is obtained from some user-selected samples. The key point is to recognize when an optimum classification result has been achieved, where additional iterations may not improve the result. Finally, bad initialization of samples can only produce unsatisfactory extraction result, so the quality of the final extracted solution is directly related to the relevancy of the training samples provided.

eCognition's original object-oriented design offered a great potentiality for the extraction of man-made objects, using specified criteria such as class spectral and geometrical properties. Indeed, object-oriented algorithms are theoretically promising for the classification of high resolution imagery. Still, it requires the user to have a good knowledge on the spectral and spatial characteristics of the objects to be extracted, in order to differentiate them from other objects. The user must set many parameters, and a lot of iterations are required to achieve an acceptable classification results. Finally, the segmentation plays a crucial role in the whole process, affecting the final object extraction result; an appropriate scale selection will determine the final accuracy of the classified image.

Genie Pro provided a clever process which is relatively straightforward to use. It employs an evolutionary algorithm to explore potential attribute extractors in the neighborhood of local pixels. As for eCognition, the possibility of using multi-classes offered a great potentiality to extract different kind of features and for dealing with imagery of diverse modality. However, it is not easy to know when the training process has reached an optimum in order to produce the best classification results.

Limitations of the Evaluation

Within the scope of this research, the evaluation of AFE software tools for change detection of man-made objects on high resolution electro-optic imagery brings some limitations.

The use of post-classification comparison methods implies some issues related to image co-registration. Indeed, as image co-registration algorithms were required at the first step of the change detection process, introduced misregistration errors may cause spurious change detection results and false alarms. The influence of the co-registration accuracy on the change detection results mainly depends on classification accuracy and co-registration accuracy.

According to our experiences, pixel-based classification methods have shown some limitations when applied to high resolution imagery, because man-made objects become less smoothed and more heterogeneous in such imagery than in coarser resolution ones. Each object may contain components of different spectral signatures, thus producing more noisy results than if homogeneous coarser resolution images are used. While many subjective factors affect the classification accuracy using pixel-based supervised image classification, our results suggest that object-oriented classification on high resolution imagery are also affected by subjective factors such as the levels of multiresolution segmentation, the weight of shape (texture) and pixel value, and the selection of objects as training samples.

The accuracy assessment metric used to compare each AFE tool against the others was the overall accuracy, computed from change detection error matrices. Overall accuracy values were quite low throughout the evaluation, which result from misclassification but also by the way the sampling points were chosen for computing the accuracy assessment. It should be noted that if sample points were randomly generated for the entire imagery – as oppose to only areas in which cover type was equal to objects of interest – the overall accuracy would most likely be much higher. This is because the commission and omission errors for sample points in which cover type was equal to objects not of interest would most likely be much lower. Moreover, since there are no more actual changes based on the reference data but there are more changes based on the classification data, the agreement is not so well. The bigger is the disparity between the reference data and the classification data, the smaller is the K_{hat}. Thus, K_{hat} was low because of the sampling design. For this research, the same accuracy assessment metric has been used to evaluate all three AFE tools. Even though the overall accuracy is low, it provides a similar comparative metric to assess the change detection performance of each AFE software tool against the others.

Only six images (3 pairs) segmented in three scenarios have been employed in this research. Although the preliminary results appear promising, additional experiments with more imagery are needed in order to acknowledge further the performance and limitations of each AFE tool.

Even thought all tested AFE tools use different machine learning techniques, they all share the same sensitivity problem related to the selection of training samples by the image analysts to obtain satisfactory extraction results. Operator's subjectivity determinates the quality and accuracy in the selection of appropriate training samples, conditioning the extraction results, and thus the change detection performance of the AFE tool. Additional experiments of these AFE tools using a plurality of image analysts shall then be carry on.

CONCLUSION

The proliferation of commercially available sensor platforms has led to the situation where the image analysts are at risk to be overwhelmed by the number of images to be analyzed. Any automated image procedures that would assist the image analysts in the extraction of meaningful knowledge from image data, and cue their attention to significant events of interest, are thus welcome (Lavigne, 2006). Automated feature extraction (AFE) software packages benefit from manual digitizing, by providing a way to enhance the extraction of geospatial features. Such tools can be used on high resolution imagery to realize pixel-based and object-based change detection, by extracting pixels and objects of interest in an assisted manner, and then comparing the extracted features for change detecting analysis. As each AFE tool has its own embedded machine learning technique, this paper presented a comparative analysis of three AFE tools, for the extraction of man-made objects for change detection.

A dataset of six remotely sensed images (3 pairs), segmented along three scenarios, was used to assess the change detection performance of these AFE tools. The accuracy assessment of each AFE tool for change detection purposes was quantified using the overall accuracy, computed from change detection error matrices. The performance of Feature Analyst for aircraft change detection and building change detection was better than the ones provided by eCognition and Genie Pro. eCognition showed the best performance for building change detection than for the other scenarios. Genie Pro performed well for aircraft change detection but significantly less than the other AFE tools for building change detection. Among the three tested scenarios, vehicle change detection was obviously the most challenging one, with the lowest computed overall accuracy values, and where no AFE tools performed better than the others. In order to increase the use of AFE tools on high resolution imagery for change detection applications, the extraction capabilities of these AFE software packages must be enhanced. Pixel-based and object-based classification tools are likely an indistinguishable way to pursue. Nevertheless, increase in the global learning performance of the spatial and spectral characteristics of the objects of interest is mandatory and undoubtedly the key toward the use of AFE tools for change detection on high resolution imagery.

REFERENCES

Benz, U.C., Schreier, G., (2001). Definiens Imaging GmbH: Object oriented classification and feature detection. *IEEE Geoscience and Remote Sensing Society Newsletter*, *9*:16-20, September.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, *58*:239-258.

Congalton, R.G. (1999). *Assessing the accuracy of remotely sensed data: Principles and practices*. Lewis Publishers, New York.

Harvey, N.R., Theiler J., Brumby S.P., Perkins, S., Szymanski, J.J., Bloch, J.J., Porter, R.B., Galassi, M., Young, A.C., (2002). Comparaison of GENIE and conventional supervised classifiers for multispectral image feature extraction. *IEEE Transactions on Geoscience and Remote Sensing, 40*(2):393-404.

Harvey, N.R., Theiler J., (2004). Focus-of-attention strategies for finding discrete objects in multispectral imagery. In: *Proc. SPIE Ultraviolet Ground and Space-based Measurements, Models, and Effects IV*, Vol. 5546, pp. 179-189.

Hong, G., Zhang, Y., Lavigne, D.A., (2006). Object-based change detection in high resolution image. *ASPRS 2006 Annual Conference*, Reno (Nevada), 1-5 May 2006.

Lavigne, D.A., (2006). A fully automated image co-registration system. In: *Proc. SPIE Signal Processing, Sensor Fusion, and Target Recognition XV*, Vol. 6235, Orlando (Florida).

Macleod, R., Congalton, R. (1998). A quantitative comparison of change detection algorithms for monitoring eelgrass from remotely sensed data. *Photogrammetric Engineering and Remote Sensing, 64*(3):207-216.

O'Brien, M.A., (2003a). Feature extraction with the VLS Feature Analyst system. *ASPRS 2003 Annual Conference*, Anchorage (Alaska), 5-9 May 2003.

O'Brien, M.A., (2003b). Performance evaluation of the Genie system. *ASPRS 2003 Annual Conference*, Anchorage (Alaska), 5-9 May 2003.

O'Brien, M.A., Irvine, J.M., (2004). Information fusion for feature extraction and the development of geospatial information. *7th International Conference on Information Fusion*, Stockholm (Sweden), June 28th - July 1st 2004.

Perkins, S., Edlund, K., Esch-Mosher, D., Eads, D., Harvey, N., Brumby, S., (2005). Genie Pro: Robust image classification using shape, texture and spectral information. In: *Proc. SPIE Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, Vol. 5806, pp. 139-148.