SEGMENTATION AND THEMATIC CLASSIFICATION OF COLOR ORTHOPHOTOS OVER NON-COMPRESSED AND JPEG 2000 COMPRESSED IMAGES

A. Zabala [a,*], C. Cea [a], X. Pons [a, b]

[a] Dept. of Geography. Autonomous University of Barcelona, 08193 Cerdanyola del Vallès, Spain - (alaitz.zabala, cristina.cea, xavier.pons) @uab.cat
[b] Centre for Ecological Research and Forestry Applications (CREAF), 08193 Cerdanyola del Vallès, Spain

KEY WORDS: Segmentation**,** Mapping, Compression, Orthophotos, Wavelets, JPEG 2000

ABSTRACT:

Lossy compression has been increasingly used because of the enormous amount of images gathered by airborne and satellite sensors. Nevertheless, the implications of these compression procedures have been scarcely assessed. Moreover, segmentation previous to digital image classification is also an increasingly used technique. This paper presents an object-oriented application for image analysis using color orthophotos but using different compression levels in order to study the effects of the data loss on the segmentation-based classification results. A set of 4 color orthophotos having 1 m spatial resolution and covering an area of about 1200x1200 m^2 (144 ha) was chosen for the experiment. Those scenes were compressed at 8 compression ratios (between 5:1 and 1000:1) using the JPEG 2000 standard.
Among the different segmentation algorithms tested, best results were obtained using the Definiens Multiresolution segmentation algorithm on the first level and the Spectral difference segmentation on the second level.
Aimed thematic, final categories were 7, namely dense vegetation, herbaceous, bare land, road and asphalt areas, buildings, swimming pools and river (if necessary). Best category classification was obtained using a hierarchical classification algorithm over the second segmentation level. The same segmentation and classification methods were applied in order to establish an automatic or semi-automatic technique for all 36 images.
To estimate the overall accuracy, a confusion matrix was calculated using independent test areas. The mean accuracy over non-compressed images was 81%, while the overall kappa index of agreement was 0.7. First and second compression levels results were similar to the reference ones. Differences in the third to fifth levels were moderate to important (accuracies 77%-62%), while more compressed images obtained the worst results (accuracies lower than 65% and kappa lower than 0.5).

1. INTRODUCTION AND OBJECTIVE

The total amount of data generated by remote sensing (RS) platforms is increasing every day. Even thought the commercial providers have an important structure to store all this information, users do not often have availability to this structure and thus data management problems arise. On the other hand, some softwares have significant difficulties to work with large images. Moreover, the new paradigm of Spatial Data Infrastructures (SDIs), promotes the establishment of web data services, which usually need compressed images to transfer them to environments with restricted bandwidth. In a disaster or emergency management environments, and especially in a crisis situation, image compression techniques are crucial for efficient and fast dissemination of this critical data to the reaction teams that normally use laptop computers and mobile devices.
In recent years a set of lossy compression techniques based on different algorithms has been appeared (discrete cosine transform, DCT, discrete wavelet transform, DWT,…). One of the first format appeared was the JPEG, but later wavelet based formats, such as SID, ECW and JPEG 2000, appeared and became particularly popular within the RS community (Zabala *et al*., 2006). JPEG 2000 became an ISO standard in 2000 and was revised in 2004 (ISO/IEC 2004). This widely used compression standard is studied in this paper, which aims to evaluate the implication of image lossy compression on object based classifications, because there are few quantitative analysis of this approach to data reduction. It is important to bear in mind that in every case we are dealing with lossy compression algorithms, which sacrifice part of the data in order to achieve a higher compression ratio, but sometimes with no appreciable loss of image quality (Kiema, 2000; Zabala *et al.*, 2007).
Traditionally, image classification has employed pixel-based methods based on statistical measurements on individual pixels values to classify each pixel through supervised or unsupervised classifiers (Xiaoxia *et al*., 2005; West *et al*., 2009).
Recently, with the improvement on image spatial resolution, (such as aerial or high resolution satellite sensors), the number of object–based techniques-related articles have increased (Blaschke, 2010). In this case, image classification needs groups of pixels selected as objects. So, s segmentation process is needed to obtain this objects, being the first step to classify (Conchedda *et al*., 2008).
There are many different approaches to segmentation and classification, in this case Definiens algorithms (Definiens, 2005) have been used in this study. In this particular paper a set of 4 color orthophotos of different areas have been used. These scenes were first processed without any compression, and then applying eight levels of increasing compression (between 5:1 and 1000:1) were used as well, obtaining a total amount of 32 compressed and 4 uncompressed images.

* Corresponding author

The conclusions of this study will also discuss the feasibility of such approach (segmentation over compressed images) in an emergency management environment.

2. METHODS

2.1 Study area and material

A set of 4 aerial RGB orthophotos have been selected: 3 of them over Catalonia (NE of Spain) and the other one over Navarre (N of Spain). More specifically, on Catalonia we have selected areas in Sant Cugat del Vallès, Vallvidrera and Olot. In Navarre, we have selected areas in Zizur Mayor. These 4 orthophotos have been selected be cause of their similar urban characteristics, characterised for low density urban areas with a significant number of swimming pools, gardens, forested areas, and field crops. However, some differences can be detected among the scenes: in two of them (Vallvidrera and Sant Cugat) the predominant urban structure is detached houses while in the other ones more semi-detached and multi-family houses are present.
A working spatial resolution of 1 meter was selected. In some cases, the original spatial resolution was 0.5 m, and in those cases, the pixel size it was changed to 1m using a mean algorithm during the interpolation. A sub-scene of approx. 1200 x 1200 pixels were selected over each area to speed up all the process.

2.2 Image compression

The compression and decompression application used was BOÍ, an implementation of JPEG 2000 standard Part 1, developed by the GICI team of the Autonomous University of Barcelona (Aulí-Llinàs *et al*. 2005).
The 3 RGB bands of the color orthophoto of each area have been compressed to a single JPEG 2000 compressed file. This jointly compression benefits of preserving as much information from the three bands together, allowing to keep more information of any of the color components if necessary. A lossy compression has been computed over original images, at eight compression ratios (CR): 5:1, 10:1, 20:1, 40:1, 100:1, 200:1, 400:1 and 1000:1%, where the first number on the ratio defines the size of the original image related to a compressed file size equal to 1.
Compressed images are then decompressed to feed the segmentation and classification process. A total amount of 36 images have been used in this study (1 non-compressed and 8 compressed for each of the 4 study areas).

2.3 Image segmentation

Definiens professional 5 has 4 segmentation algorithms. An important part of the literature suggested Multiresolution Algorithm (Baatz *et al.*, 2007) as first step to obtain a highly homogeneous image objects in the first level of segmentation, when used over aerial orthophotos and satellite images.
This algorithm allows defining several parameters. The first one is the weight of the different layers; the same weight is chosen for each of the RGB components. Another one is the scale parameter, which defines the desired object size that in turn also will depend on the size of the objects found in the image. Finally, the parameters of shape/colour and compactness/ smoothness require more attention because of the difficulty to apply an appropriate value.
Different tests have been performed using different values of shape/colour and compactness/smoothness, and they have been chosen empirically based on visual inspection. The default values for shape/colour offered by the software were used as the first try. However, giving most of the weight to the digital values is not the best option when it involves only 3 visible bands in the process. The shape of objects is also significant for the segmentation and the digital values are not enough to define the categories. In a second try, a weight of 50% has been selected. In the case of the compactness/smoothness, default values are 50% in both cases, however it has been checked that the highest values of compactness, the better discrimination of the edges of different objects. Little objects as private pools have been identified as individual objects with these selected values, as well as isolated trees and some building roofs.
After establishing an initial segmentation level, a second segmentation has been carried out to collect some objects belonging to the same class. The Difference Spectral Segmentation algorithm has been used. This algorithm creates a new level from an existing level since the objective is to join neighbouring objects according to the mean DN of the different layers (Definiens, 2005). The most important parameter is the total Spectral difference among the objects generated in the previous segmentation. If the spectral difference between neighbours is less than the selected value, neighbour objects join together. Different values have been tested. With a spectral value of 5 no important changes occurred between both segmentation levels; nevertheless with a value of 15, objects of different classes were confused, for example bare soil and asphalt. The optimal selected value was 10.
Different segmentation levels are important in order tot extract boundaries of the objects occurring at the corresponding scales.
The same parameters have been applied to all original and compressed orthophotos, as a rule.

2.4 Image classification

The segmented image has been classified using a user-defined fuzzy classification. First step is assigning classes, the semantic meaning of image objects in the Definiens network (Definiens, 2005). Each class was selected based on their visual appearance in the true color image. To create the classification scheme, six dominant object classes (land-cover types) were identified, namely bare land, dense vegetation, building areas, road/asphalt areas and herbaceous (see table 1).

	Orthophotos			
Classes	**Olot**	**Sant Cugat**	**Vallvidrera**	**Zizur Mayor**
Dense vegetation	Yes	Yes	Yes	Yes
Herbaceous	Yes	Yes	Yes	Yes
Bare land	No	Yes	Yes	Yes
Road and asphalt areas	Yes	Yes	Yes	Yes
Buildings	Yes	Yes	Yes	Yes
Swimming pools	Yes	Yes	Yes	Yes
River	Yes	No	No	No

Table 1. Classes defined for each orthophoto.

Some classes were divided into subclasses due to the spectral and textural differences among them. This division helped the classifier and improved the classification results because each subclass has their own features. Thus, bare land was then divided into two sub-classes as fallow field and abandoned agriculture land. Buildings were separated into different sub-

classes depending on the roof colour, light and dark. Herbaceous was divided into garden and agriculture. Although swimming-pools could be integrated as buildings areas or asphalt areas, we decided to maintain then as one class because there were an important number of them in the orthophotos.
To perform the classification we have tried the methods offered by the software. First try was to classify separately for each class and subclass using simple conditions or using feature description classes. Second try, was to classify simultaneously the set of classes and subclasses in the class hierarchy from their feature description. And finally, last try was an automatic classification based on the nearest neighbour algorithm, defining training areas.
Better results were obtained with the classification based on class hierarchy and fuzzy functions defined for selected features, because it allows more control over the classification process and can be more easily adapted to other orthophotos (Kressler, 2005). Using only simple variables conditions produced confusion between classes. Using Nearest Neighbour algorithm (NN) was not appropriate because training areas in each image must be defined and, thus, the classification can not be automated.
Table 2 shows the features used for each class. The values of these features have been adapted, changing this value or activating or disabling expressions in class description, to each orthophoto and each compression level.

Class	Features
Dense vegetation	Ratio green, Brightness, Max. Difference
Herbaceous	Brightness, Standard Deviation Green
Bare land	Brightness, Area, Mean red Ratio red
Road and asphalt areas	Max. Difference, length, brightness
Buildings	Brightness, Ratio RGB, Area
Swimming pools	Area, Mean blue, Distance to edifications,
River	Topographic map mask

Table 2. Features used for classification

It is important to note that working with only 3 bands from the visible spectrum makes difficult to obtain an optimal classification because the limitation on the input information.

2.5 Validation/accuracy assessment

A set of independent test areas have been selected to carry out the accuracy assessment of all classifications by generating a confusion matrix. The accuracy reports overall accuracy, describing the percentage of the well classified pixels and the total amount of classified pixels, and overall *kappa* index (representing a measure of agreement among classes) for each classification (Campbell, 2002). User's and producer's accuracy for each class have been obtained as well.

3. RESULTS

The results are presented in two main sections. The first is a detailed description of the generated segmentation for all compression ratios. The second section presents the classification results for each orthophoto.
The most important categories have been displayed and discussed in sections 2.3 and 2.4. Figure 6 shows accuracies based on the confusion matrices.

3.1 Segmentation

As described in section 2.3, segmentation was performed in two levels for all original and compressed images. In the first step, image was divided into numerous objects as a previous step to merge these into larger ones. Figure 1 shows a detail of the Zizur Mayor uncompressed orthophoto with the first segmentation level overlaid in black. Different land cover are well separated, but larger objects as roofs, public swimming pools, green areas and roads are divided into numerous objects.

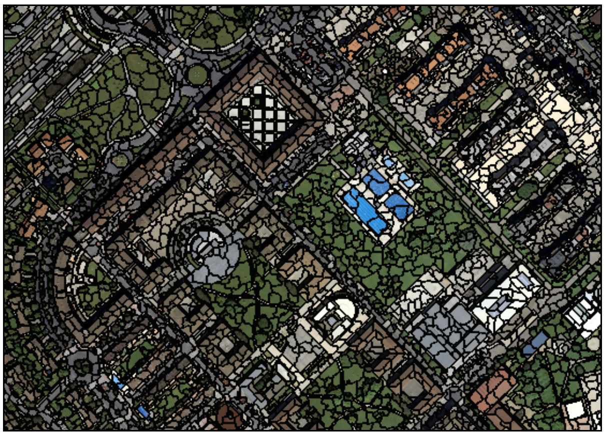

Figure 1. Level 1 segmentation in a detail of the uncompressed Zizur Mayor image.

To reduce the number of objects a second level of segmentation was necessary. Figure 2 shows the result of the second segmentation over the same bounding box. Roofs, green areas and swimming pools were well merged into one or two objects, except roads.

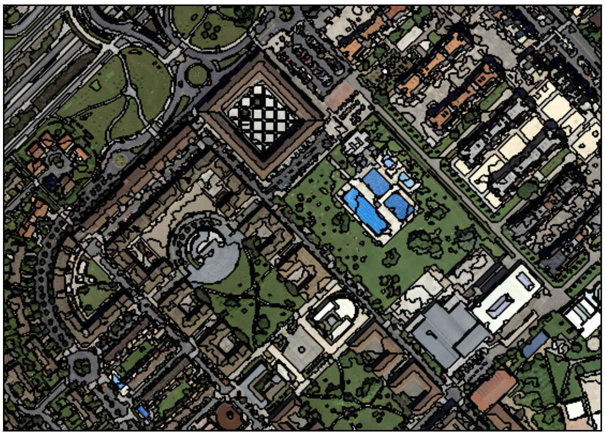

Figure 2. Level 2 segmentation in a detail of the uncompressed Zizur Mayor image.

Figure 3 and 4 show the relation in percentage between the number of objects of each compressed image and the number of objects of the original image (all of them obtained using the same segmentation parameters). The four studied areas show a similar trend. The first four compressed levels have a very similar number of objects with respect to the original images; however after the fourth compression level, the number of objects decreases with the increase of compression. It occurs because at high compression ratios images are homogenised, *i.e.*, small objects (lands covers) tend to merge in to one influenced by close covers.

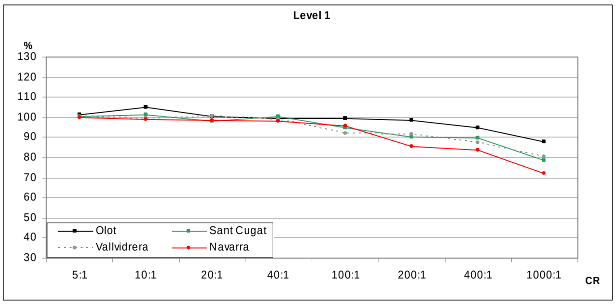

Figure 3. Percentage of compressed ratios segmentation objects with respect to the uncompressed segmentation objects, in level 1.

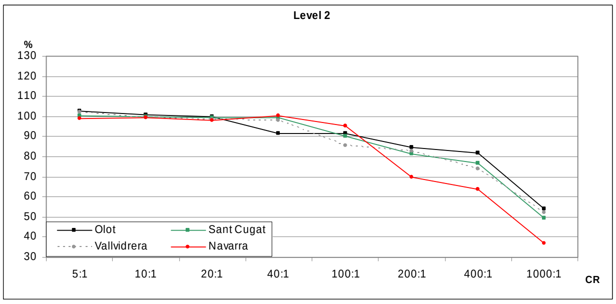

Figure 4. Percentage of compressed ratios segmentation objects with respect to the uncompressed segmentation objects, in level 2.

3.2 Classification

One of the aims of this study was to create an automatic fuzzy user-defined classification process that could be applied to all orthophotos using the same object features for discriminating the different categories. After different tries, as explained in section 2.4, it is proved that it can be achieved, but it is important to note that it is necessary to adapt the values of this object features to each RGB orthophoto, and that, in some cases, another new object feature is needed as well.
Due to the high number of classifications carried out, it is not possible to show all of them here. Figure 5 shows a detail of recovered images (on the left) and the respective obtained classifications (on the right), for three compression levels and over the same area.

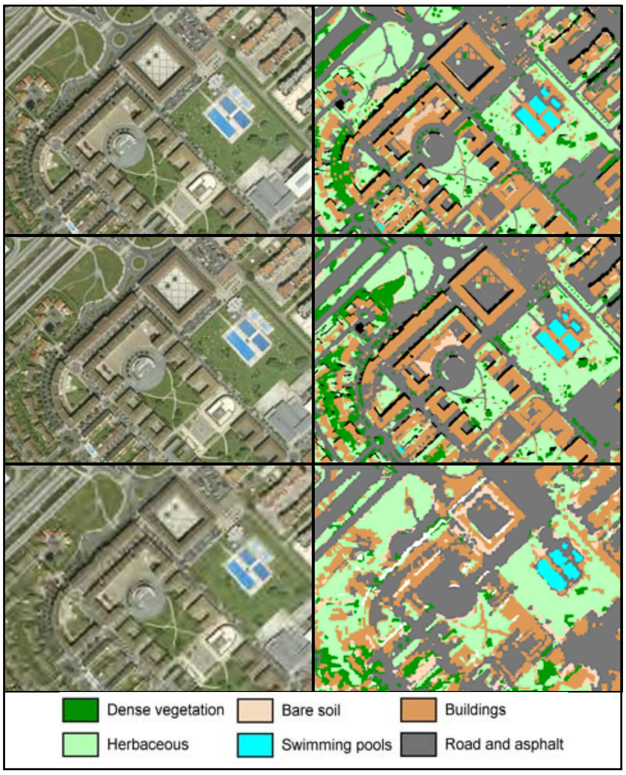

Figure 5. Detail of the images (on the left) and of the classifications (on the right) of original, 10:1 and 200:1 compression ratios.

Figure 6 shows the overall accuracy for each classification and each compression ratio. As in the case of segmentation, the four areas follow the same trend: better overall accuracy was obtained in the three first compressed levels, but if more compression is applied, this value descends considerably. This descend is less prominent for the Olot area, because this orthophoto has an important area of dense vegetation, and this area presents less problems to be classified at any compression level. In the case of the Zizur Mayor area, this descends is more significant because the most important cover in this area is bare soil and dense vegetation area is minimal. Bare soil turned out difficult to be well classified due to the significant confusion with bright and white covers as roofs.
If we focused on the different overall accuracies among areas, Vallvidrera and Sant Cugat have the lowest accuracies. This is because these areas are characterized by single semi-detached and detached houses with private swimming pools surrounded by dense vegetation. This means that they are fragmented areas and high levels of compression trend to homogenise, blurring small objects with the objects situated around them.
The Olot area it is the most homogeneous zone with well differentiated classes, implying that high compression does not produce significant effects in the classification.
Although Zizur Mayor area is an homogeneous zone as well, the important confusion of bare soil as asphalt, white roofs or herbaceous vegetation imply that the classification with high compression levels obtain low accuracy.

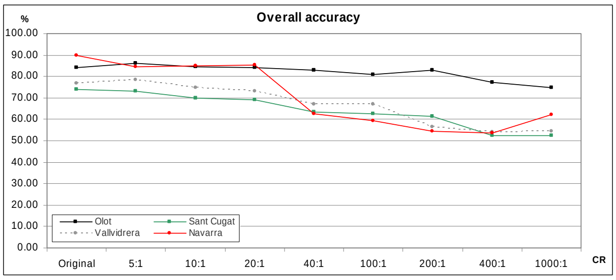

Figure 6. Overall accuracy of each classification areas for different compression ratios.

User's accuracy is now analysed for each class.

Figure 7 shows user's accuracy for dense vegetation. Except in one case, relatively high accuracy is maintained for the areas at all compression levels as explained above. Only Zizur Mayor area has a significantly descend of user's accuracy due to the fact that area of dense vegetation is minimal. Some isolated trees or group of two or three trees forms this cover, especially surrounded by herbaceous vegetation, which at high compression levels is homogenized with surrounding covers.

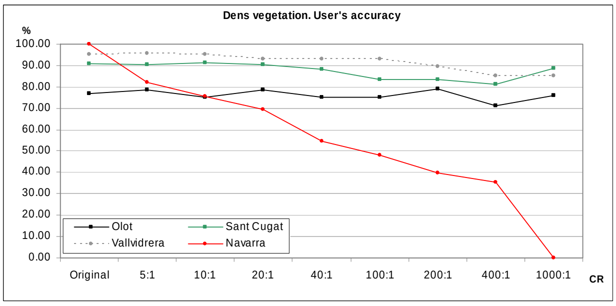

Figure 7. Dense vegetation user's accuracy.

Figure 8 shows user's accuracy for the herbaceous class. This cover presents important commission errors from dense vegetation. When using multispectral classification methods, there are no problems to discriminate vegetation from other covers; however, discriminating between dense and herbaceous vegetation in RGB orthophotos sometimes is not feasible. This confusion implies different accuracy values between different compression levels.

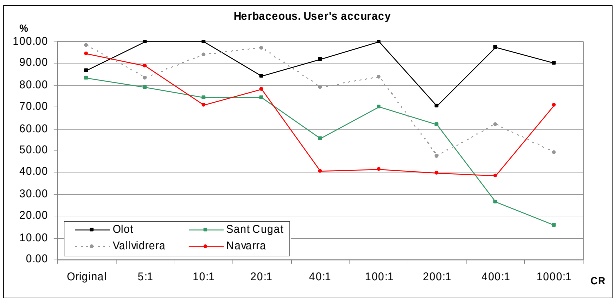

Figure 8. Herbaceous user's accuracy

Buildings cover is heavily depending on the color and type of roof. Brown roofs are confused with fallow lands and white roofs are confused with asphalt areas. Better user's accuracies, shown in figure 9, were obtained in the areas where roofs are more similar among them, as Olot or Sant Cugat. Worst results were obtained in areas were roofs are very different among them, specially between single detached houses with gable roofs and apartments building with flat roofs, as in Vallvidrera and Zizur Mayor.

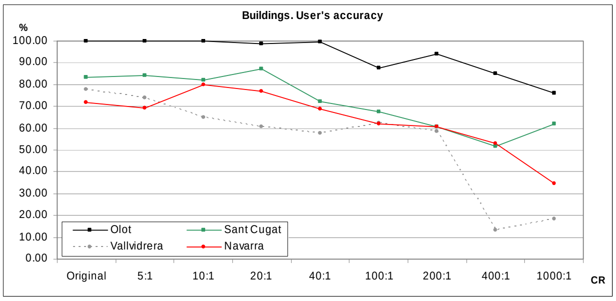

Figure 9. Buildings user's accuracy.

Figure 10 shows roads and asphalt area user's accuracy. As we explained in other covers, better results were obtained in semi urban areas as Olot and Zizur Mayor, where important roads exist. Worst results were obtained in Vallvidrera and Sant Cugat because there are few important roads and these are situated in low density residential zones. In addition, many of narrow roads are covered by canopy trees.

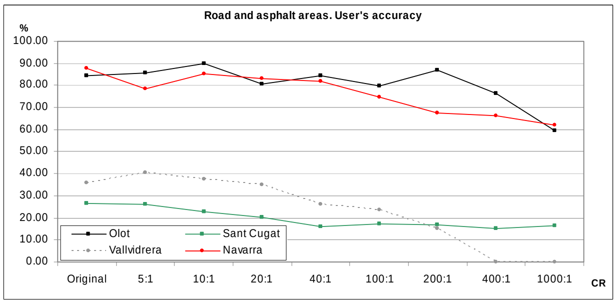

Figure 10. Road and asphalt user's accuracy.

Swimmimg pools user's accuracy is shown in figure 11. Clearly is the best classified cover at almost all compression levels, because this cover is very different of the other ones. Except in the case of Zizur Mayor area, user accuracy is significantly reduced at high compression levels. It occurs due to the small size of many private swimming pools, because at high compression levels these are blurred with other covers.

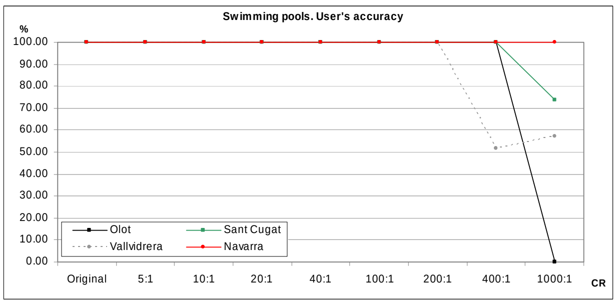

Figure 11. Swimming pools user's accuracy.

This fifth cover is in all four areas, and bare soil and river are in some of them. The size of the bare soil areas plays an important

role to discriminate these from other covers, even though presents confusions with other covers especially at high compression levels because of the different types of bare soil (future building, fallow land, etc.)
The last class, river, was well classified because a topographic map was introduced in the classifier.

4. CONCLUSIONS

The main conclusion of this study is that at high compression levels, poor accuracies are obtained. Regarding, fragmentation landscape, it occurs in all cases, but more significantly in more fragmented areas.
For each type of cover, the accuracy depends on the size of this cover in each orthophoto, the objects that enclose this cover and the heterogeneity with respect to the other covers. Bare soil, buildings and herbaceous categories are more difficult to classify. Dense vegetation, asphalt and swimming pools present user's accuracies above 50% at all compression levels, except in the case of Zizur Mayor.
Similar overall accuracies have been obtained for uncompressed and the three first compression levels, especially first and second, for all areas.
It is important to note that classifying using 3 visible bands and using object-based methods, acceptable results have been obtained; however, adding an infrared band could reduce confusion errors among categories and improve discrimination, especially in categories such as bare soil, vegetation and water bodies. Nevertheless it is also true that up to date, most aerial information available lacks of these infrared components, especially in older dates.

5. EMERGENCY MANAGEMENT APPLICABILITY

The last aim of this study was to assess if segmentation processes applied to lossy compressed images may be usefully applied to emergency management scenarios. In such situations, a quick response implies a quick analysis too. The research carried out shows that the segmentation and classification processes are not faster in the compressed images than in the original images. Furthermore, with significant compression ratios is more difficult to define the categories covering the studied area. Thus, the only advantage of compressed images would be their faster transference, but not a quicker or better analysis of them. In an emergency environment, then, the best option would be using lossless or low-compression-ratio lossy compressed images.

6. ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish Government, by FEDER, and by the Catalan Government, under Grants TSI2006-14005-C02-02, TIN2009-14426-C02-02 and SGR2009-1511, and also by the European Commission through the FP7-242390-GEO-PICTURES project (FP7-SPACE-2009-1)

7. BIBLIOGRAPHY

References from Journals:

Blaschke, T. 2010. Object based image analysis for remote sensing. *ISPRS Journal of Photogrametry & Remote Sensing*, 65, pp. 2-16.

Conchedda, G. *et al*, 2008. An object–based method for mapping and change analysis in mangrove ecosystems. *ISPRS Journal of Photogrametry & Remote Sensing*, 63, pp. 578-589.

Kressler, F. P., M. Franzen, and K. Steinnocher, 2005. Segmentation based classification of aerial images and its potential to support the update of existing land use data bases, *Proceeding ISPRS Hannover Workshop 2005, High Resolution Earth Imaging for Geospatial Information,* Heipke C., Jacobsen K., Gerke M. (Eds.), Hannover, Germany.

Kiema, J.B.K. 2000. Wavelet compression and the automatic classification of Landsat imagery. *Photogrammetric Record*, 16 (96), pp 997-1006.

West, B., Zhang, Y. 2009. Region based segmentation of QuickBird multispectral imagery trough band ratios and fuzzy comparison, *ISPRS Journal of Photogrametry & Remote Sensing*, 64, pp. 55-64

Xiaoxia, S *et al*. 2005 A comparison of object-oriented and pixel-based classification approaches using quickbird imagery. *ISPRS STM*, Beijing, pp.281-284.

References from Books:

Aulí-Llinàs, F., J.R. Paton, J. Bartrina-Rapesta, J.L. Monteagudo-Pereira, J. Serra-Sagristà. 2005. "J2K: introducing a novel JPEG 2000 coder", In. *Proc. SPIE*, vol. 5960, 2005, pp. 1763-1773.

Baatz, M., Shäpe, A. 2000. Multiresolution segmentation: an optimization for high quality multi-scale image segmentation. Strobl, J., Blaschke, T., Griesebner, G (Eds.). 2006. *Angewandte Geograsphische Information-Verarbeitung* XII. Wichmann Verlag, Karlsruhe, pp. 12-23.

Campbell, J.B. 2002. *Introduction to Remote Sensing*, 3rd ed.(Book Review) -. Guilford Press, 2002. ISBN 1-57290-640-8.

Definiens (2005). User's guide, Munich.

ISO/IEC, 2004, 15444-1:2004, Information technology – JPEG 2000 image coding system: Core coding system. Geneva, Switzerland, 2004, 194 pp.

Zabala, A., Pons, X., Díaz-Delgado, R., García, F., Aulí-Llinàs, F. and Serra-Sagristà, J. 2006. Effects of JPEG and JPEG 2000 Lossy Compression on Remote Sensing Image Classification for Mapping Crops and Forest areas. In Proceedings of the 2006 *IEEE International Geoscience & Remote Sensing Symposium & 27th Canadian Symposium on Remote Sensing (IGARSS2006)*, organised by the Geoscience and Remote Sensing Society. Denver, Colorado, EE.UU. ISBN: 0-7803-9510-7. pp790-793. 10.1109/IGARSS.2006.203 PDF Document

Zabala, A.., Pons, X., Aulí-Llinàs, F., Serra-Sagristà, J. 2007. Implications of JPEG 2000 lossy compression on multiple regression modeling. *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology VII*. Ehlers M, Michel U (ed.). ISBN: 9780819469076 (DOI: 10.1117/12.738028)