

POTENTIAL EVALUATION OF DIFFERENT TYPES OF IMAGES AND THEIR COMBINATION FOR THE CLASSIFICATION OF GIS OBJECTS CROPLAND AND GRASSLAND

J. A. Recio ^{a,*}, P. Helmholtz ^b, S. Müller ^b

^a Departamento de Ingeniería Cartográfica, Geodesia y Fotogrametría, Universidad Politécnica de Valencia, Camino de Vera s/nº, 46022 Valencia, Spain – jrecio@cgf.upv.es

^b IPI - Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover, Nienburger Str. 1, 30167 Hannover, Germany (helmholtz, mueller)@ipi.uni-hannover.de,

Commission VI, WG VI/4

KEY WORDS: Agriculture, Classification, Multitemporal, Aerial, Satellite, Orthoimage, Resolution

ABSTRACT:

In many publications the performance of different classification algorithms regarding to agricultural classes is evaluated. In contrast, this paper focuses on the potential of different imagery for the classification of the two most frequent classes: cropland and grassland. For our experiments three categories of imagery, high resolution aerial images, high resolution RapidEye satellite images and medium resolution Disaster Monitoring Constellation (DMC) satellite images are examined. An object-based image classification, as one of the most reliable methods for the automatic updating and evaluation of landuse geospatial databases, is chosen. The object boundaries are taken from a GIS database, each object is described by means of a set of image based features. Spectral, textural and structural (semivariogram derived) features are extracted from images of different dates and sensors. During classification a supervised decision trees generating algorithm is applied. To evaluate the potential of the different images, all possible combinations of the available image data are tested during classification. The results show that the best performance of landuse classification is based on RapidEye data (overall accuracy of 90%), obtaining slightly accuracy increases when this imagery is combined with additional image data (overall accuracy of 92%).

1. INTRODUCTION

The discrimination of cropland and grassland during classification of remote sensing data is still an unsolved problem, even though these classes play an important role in GIS databases in the field of agricultural policy and ecological sustainability.

The idea of the proposed approach is to use a wide range of object-based calculated features as input for classification to use as much information of the data as possible. In contrast, many other approaches rely only on some empirically determined features. To deal with the amount of features the data mining and classification tool C5.0 (Quinlan, 1993) is used.

To address the interesting question what image resolution is necessary to make the discrimination between cropland and grassland possible we use different image resolutions starting with high resolution aerial images and ending with imagery of 32m GSD.

The classification results that are achieved with the proposed approach are used for a change detection of an existing GIS database. Change detection means in this context a simple comparison of the classification outputs with the content of the

GIS database, because the object boundaries of the classified regions are identical to the GIS object boundaries.

In the following section approaches dealing with the classification of cropland and grassland are analyzed in consideration of image type and used features. In section 3 the data used for the proposed approach is presented. Section 4 describes the complete approach including feature extraction, classification and evaluation of GIS objects. An evaluation of the approach is done in section 5 using all possible combinations of input data described in section 3. At the end of the paper in section 6 a short conclusion is given.

2. RELATED WORK

In this section we briefly review approaches for the classification of cropland and grassland based on different imagery. Not only the images resolutions but also the used features have a deep impact on the classification results. For this reason, the used features are described in detail as well.

First, we start with a review of approaches which use high resolution images such as aerial images or satellite images with a resolution of one meter or better. Rengers and Prinz (2009) use the neighbourhood grey-tone difference matrix (NGTDM) to classify cropland, forest, water, grassland and urban areas in

* Corresponding author.

aerial and IKONOS images. This method is based on the differences of the grey values of two pixels and the differences of the grey values of the local neighbours, from which textural features such as coarseness, complexity and textural strength are derived. The results presented in Rengers & Prinz (2009) show that with the exception of grassland and cropland the classes mentioned above can be distinguished well. A similar conclusion is drawn by Busch et al. (2004), who apply a texture-based classification method based on Markov random fields (Gimel'farb, 1996) to aerial and IKONOS images. Their method is well-suited to classify settlement areas, industrial areas, forests, and the combined class cropland/grassland. The results reported in (Rengers & Prinz, 2009) and (Busch et al., 2004) show that using high resolution imagery but a purely textural analysis is not sufficient for separating cropland and grassland. Spectral and / or structural information is required for that purpose.

Compared to (Rengers & Prinz, 2009) and (Busch et al., 2004) Trias-Sanz (2006) uses only structural features to discriminate objects with similar spectral and textural properties, namely cropland, forest, orchards, and vineyards using aerial images with a resolution of 0.5m. Trias-Sanz (2006) distinguished between these objects only by orientation characteristics. A small window is extracted randomly inside an object to be classified, and this window (called *texton*) is used to compute a variogram of the image. A histogram of directions is derived from the Radon transform of the variogram. The maximum of this histogram corresponds to the primary direction of edges in the image, and it is used in the classification process. Using the semivariogram an overall accuracy of 95.5% could be achieved.

A similar result with an overall accuracy of 96.0% could be achieved from Balaguer et al. (2010) who also used aerial images of a resolution of 0.5m to classify different crops and grassland in the area of Spain using a decision tree classification. Compared to Trias-Sanz (2006), Balaguer et al. (2010) use not only a semivariogram but also spectral (means and standard deviation of all available bands) and textural features (Haralick features (Haralick, et al., 1973)). Similar to Balaguer et al. (2010), Helmholz et al. (2010) use a combination of spectral (mean and standard deviation of all available bands), textural features (Haralick features) and structural features using IKONOS with a SVM classification. In different to Balaguer et al. (2010) for the structural features no semivariogram is used. Helmholz et al. (2010) instead observe parallel lines visible in cropland through the cultivation to separate cropland from grassland which does not have such parallel lines. The parallel lines are detected using a histogram of directions after the calculation of a Canny edge image (Canny, 1986). Similar to Balaguer et al. (2010), Helmholz et al. (2010) achieves an overall accuracy of 96%. Both, Balaguer et al. (2010) and Helmholz et al. (2010) use an object-based approach.

Müller et al. (2010) use compared with all other approaches so far a multi-temporal high resolution dataset of images taken from a UAV (resolution of 0.17m). Similar to Balaguer et al. (2010) and Helmholz et al. (2010) spectral, textural and structural features were used. But instead to use the semivariogram or a histogram of directions for structural features, a much simpler algorithm for the detection of lines left from agriculture machines were used. An overall accuracy of 91.3% could be achieved.

In summary, with a combination of spectral, textural and structural features classification results of over 90% overall accuracy can be achieved using aerial and high resolution images of at least 1m resolution whereas the selection of the features are more important than the available of a multi-temporal dataset by comparing the results of Müller et al. (2010) and Balaguer et al. (2010) / Helmholz et al. (2010).

Next, we will briefly review approaches using medium resolution images such as IRS, Landsat TM or DMC. Haralick et al. (1973) used textural features derived from the grey level co-occurrence matrix (GLCM) such as energy, contrast, correlation and entropy along with the mean and standard deviation of the gray values of all four available channels to classify coastline, forest, grassland, urban areas and irrigated and non-irrigated cropland in California using a linear discriminate function method. By combining a textural analysis with the spectral features the classification accuracy could be improved over a purely radiometric analysis and could be achieved an overall accuracy of 83.5%.

More recently, Itzerott and Kaden (2007) tried to distinguish various types of farmland using solely the Normalised Difference Vegetation Index (NDVI) that is computed from the near infrared and the red bands of a multispectral image. Itzerott and Kaden (2007) used in their approach four different Landsat images covering a test site in Brandenburg, Germany. An overall accuracy of 65.7% using a box classifier and an overall accuracy of 72.8% using a ML-Classifier could be achieved.

Similar results could be achieved from Janssen and van Amsterdam (1991) who also worked with Landsat images, employing spectral features for the classification of different crops and grassland with a ML-classifier. Using a pixel-based approach an overall accuracy of 79.1% and using an object-based approach an overall accuracy of 76.3% could be achieved.

In a nutshell, a combination of spectral, textural and structural is necessary to achieve the best classification result. Because of the reduction of the image details using medium or low resolution images, especially structural features cannot be used, so the approaches dealing with these images is more focused on textural and spectral features. Hence, the overall accuracy decrease by using lower images resolutions as it is not possible to use as much features as using high resolution images. The review above also made clear that the used features have a deep impact on the result – not only the imagery. Hence, we want to use in this paper the same set of features for the classification of different imagery to actually evaluate the potential of different types of images and their combination for the classification of the GIS objects cropland and grassland. The best classification result could be achieved by Balaguer et al. (2010) and Helmholz et al. (2010). In this paper, we decided to use the features introduced in Balaguer et al. (2010) for the classification of different imagery.

3. DATA

In this paper we use different imagery covering a wide range of different resolutions, and the CORINE Land Cover GIS (CLC) to verify the objects cropland and grassland. The test site is situated in Halberstadt, Germany.

All used images are orthorectified. Detailed information is given in Table 1. The DMC images mentioned in this table are

images of the Disaster Monitoring Constellation operated by the company DMC International Imaging (DMCii). To describe the bands available from the different data sources the common abbreviations are used (red: R, green: G, blue: B, near infrared: NIR, red edge: RE).

Images	Bands	Resolution	Acquire date
Aerial	R, G, B	0.2m	2009-April
RapidEye	NIR, RE, R, G, B	5m	2009-08-20
DMC	NIR, R, G	32m	2009-04-24, 2009-08-24

Table 1: Overview of image data

The European CLC GIS data model was set up to be used in the scale of 1:100.000; its minimum mapping unit is 25 ha for new polygons and 5 ha for changes on existing polygons. The main landcover class in our test site is cropland. Out of 191 km² with 3455 objects, 1506 urban objects covering 19.7 km² with an average size of 1.3ha, 701 cropland objects covering 132.1 km² with an average size of 18.8ha, 323 grassland objects covering 9.9 km² with an average size of 3.1ha and 925 forest objects covering 30 km² with an average size of 3.2ha can be found.

For the evaluation we added 10% of errors randomly to the GIS reference dataset. The reference dataset was produced using field inspection and visual interpretation of the images.

4. APPROACH

In this work, each object, defined by its boundaries in a GIS database, is described by means of a set of image based features. Subsequently, it is classified with a set of decision trees and the assigned class is compared with the information contained in the database in order to detect changes in the landuse.

4.1 Features

The representation of each object in the images is described by means of a set of features extracted with the feature extraction software FETEX 2.0 (Ruiz et al, in press) specifically designed for the analysis of agricultural polygons. The employed descriptive features are grouped in three categories: spectral, textural and structural.

4.1.1 Spectral Features

Spectral features inform about the distribution of digital values of pixels inside the objects in each channel of the images, which depends on land coverage types, soil composition, state of vegetation etc. For each available band, mean, standard deviation, minimum and maximum values, range, summatory and mode are computed. Besides, when near infrared band is available; the same features are extracted from the Normalized Difference Vegetation Index (NDVI).

4.1.2 Textural Features

Texture features describe the spatial distribution of the intensity values in the image, quantifying properties such as contrast, heterogeneity, or uniformity related to each object (Ruiz et al., 2004). For each object a grey level co-occurrence matrix (GLCM) characterizing the entire object is computed. From this matrix, the following features proposed by Haralick et al. (1973) are computed: uniformity, entropy, contrast, inverse difference moment, covariance, variance and correlation.

Texture description is completed with the values of skewness and kurtosis of the histogram obtained from the intensity values of the pixels belonging to each object in the analyzed channel. For each plot, the density of edges present in a neighborhood is described by the mean and the standard deviation of the edgeness factor (Laws, 1985).

In this work, texture descriptive features are derived from the red channel of each image.

4.1.3 Structural Features

Structural features quantify the spatial arrangement of the elements contained in the objects. In this study, the structural features are calculated only for the aerial orthoimages. The limited spatial resolution of the rest of the images do not enable to detect spatial patterns in the parcels.

These features are extracted from the semivariogram graph computed for the red band of the orthoimages. The semivariogram curve quantifies the spatial associations of the values of a variable, and measures the degree of spatial correlation between different pixels in an image. This is a particularly suitable tool in the characterization of regular patterns. The expression that describes the experimental semivariogram for continuous variables is:

$$(\hat{h}) = \frac{1}{2N} \sum_{i=1}^N [z(x_i) - z(x_i + h)]^2$$

where:

$z(x_i)$ = value of the variable in position x_i .

N = number of pairs of data considered.

h = separation between elements in a given direction.

The experimental semivariogram representing each object is obtained by computing the mean of the semivariograms calculated in six directions, ranging from 0° to 150° with a step of 30°. Afterwards, each semivariogram curve is filtered using a Gaussian filter with a stencil of 3 positions, in order to smooth its shape and to eliminate experimental fluctuations. Eight structural features, whose description in detail is in Balaguer et al. (2010), are extracted for each object and are computed considering the singular points of the semivariogram, such as the first maximum, the first minimum, the second maximum, etc.

In summary, for each object in the database, 40 features (21 spectral, 11 textural and 8 structural) are extracted from the orthoimages, 39 features (28 spectral and 11 textural) from the DMC images and 53 features (42 spectral + 11 textural) from the RapidEye image.

4.2 Classification

After the feature extraction process, the described objects are classified by means of decision trees created with the C5.0 algorithm (Figure 1). The C5.0 algorithm creates a set of decision trees from a set of training data using the concept of information entropy (Quinlan, 1993) and employing only those features that are relevant, avoiding the use of redundant information. In each classification, ten decision trees, created with C5.0 and the multi-classifier algorithm boosting, are employed. In order to avoid overfitting of the trees to the training samples, a minimum number of ten elements for the

resulting classified objects sets is imposed to the algorithm. The training sample set is composed by 400 objects, with 100 objects describing each class.

In total 15 classifications are done with different descriptions of the objects obtained as the possible combinations of the features derived from the available images. Figure 1 shows the two first decision trees used in the classification of the objects combining the aerial and the RapidEye images. In the first decision tree, three spectral features (ST_DEV_R_Aerial, MEAN_RE_RapidEye and MEAN_G_RapidEye) are combined with two texture features derived from the aerial image (KURTOSIS_Aerial and ST_DEV_EDGENESS_Aerial) whereas in the second tree, five spectral features (MINIMUM_NDVI_RapidEye, ST_DEV_R_RapidEye, SUMMATORY_G_RapidEye, MINIMUM_NIR_RapidEye and ST_DEV_NDVI_RapidEye) are used with two textural features (MEAN_EDGENESS_RapidEye and MEAN_EDGENESS_Aerial).

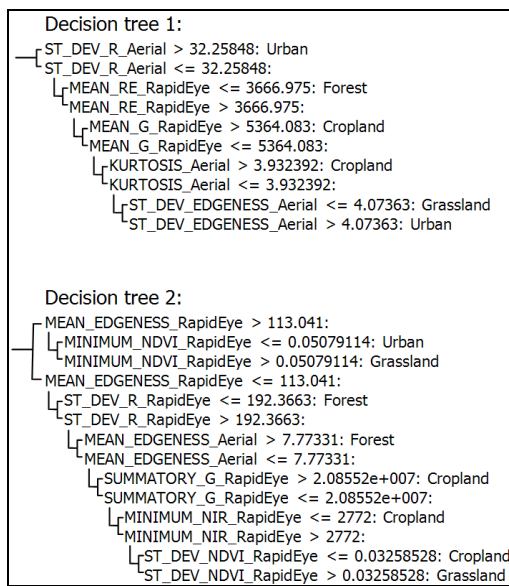


Figure 1. Two first decision trees used in the classification of the objects combining the aerial and the RapidEye images.

4.3 Change Detection

Once the objects are classified, the comparison of the new assigned landuse with the old database enables for the detection of mismatches, that is, those parcels in which the previous landuse is different from what is classified. In this way, parcels with coincident landuse in both sources are accepted, whereas parcels with non coincident landuse are rejected.

The results of the change detection process are expressed using a confusion matrix (Table 2). The efficiency of the approach is estimated by the addition of the percentage of unchanged objects accepted by the system and the percentage of changed objects rejected by the system. There will be undetected errors if objects which have been accepted by the system correspond to changed objects. The percentage of undetected errors has to be as small as possible. The false alarms are the unchanged objects rejected by the system.

		System	
		Accepted	Rejected
Reference	Unchanged	Success	False alarms
	Changed	Undetected errors	Success

Table 2: Confusion matrix

5. EVALUATION

The evaluation sample set is made up by the parcels in the database not used for training purposes. The performance of the methodology is analyzed in two aspects: accuracy of classifications and results of change detection process.

5.1 Classification

The relevance of the features in the classification process is analyzed for the four classifications based on the employment of only one image. In each case, the percentage of parcels which are classified according to a rule derived from a feature is computed. As each classification is made with the combination of ten decision trees, the average usage of each feature in the ten decision trees of the classification is calculated.

Figure 2 shows the percentage of parcels classified considering the most relevant features in the case of the classification of the aerial image. In this case, the most relevant feature is the mean of the semivariogram values up the first maximum (MFM), being evaluated for the classification of 52.6% the parcels. Following, textural and spectral features are combined, dominating the spectral features extracted from the red band.

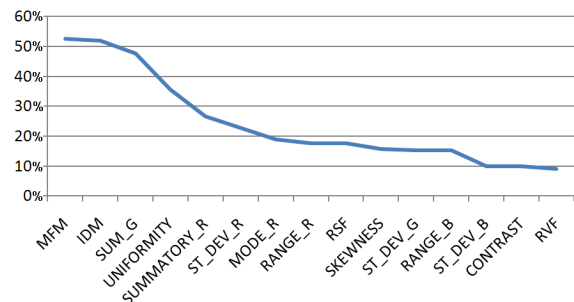


Figure 2: Relevance of features in classification of aerial image.

In the classification of the RapidEye image, the most relevant feature is the mean of edgeness factor (Figure 3). The great majority of following features are spectral and none of the GLCM texture features are employed.

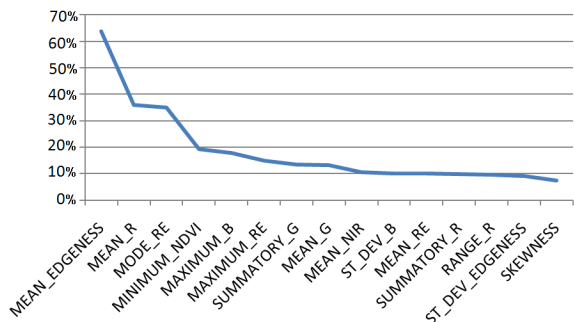


Figure 3: Relevance of features in classification of RapidEye image.

In the case of the first DMC image, the most employed features describe the spectral information, using only one texture descriptor (Figure 4).

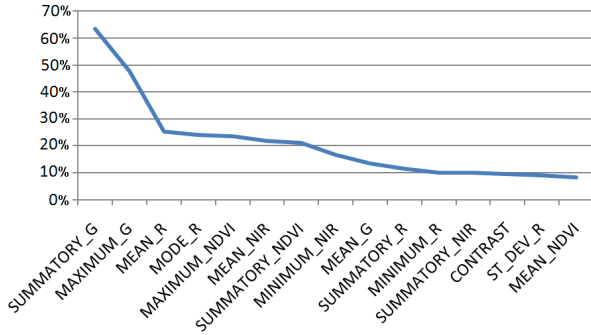


Figure 4: Relevance of features in classification of the first DMC image.

For the classification of the second DMC image, Figure 5 confirms that the most relevant features are the spectral based and only one texture descriptor is included in the classification with a low relevance.

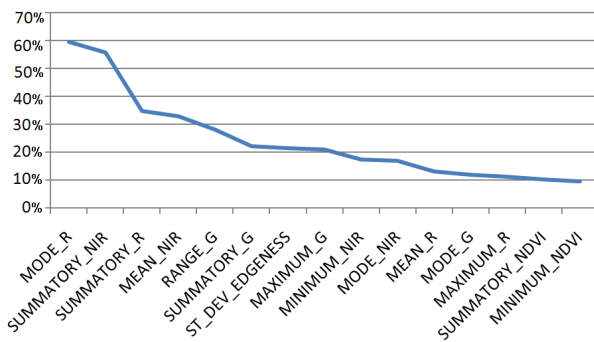


Figure 5: Relevance of features in classification of the second DMC image.

The analysis of the relevance of features in every classification reflects that the importance of spectral features increases when spatial resolution of the images decreases. Texture and structural descriptors are mainly used in the classification of images with high spatial resolution.

Figure 6 shows the overall accuracies obtained for the classifications based on the entire possible combinations of the input data. When only one image is employed, the best results are obtained with the RapidEye image. Its spectral resolution enables to describe the spectral properties of the objects more accurately than the rest of the images. Being the features derived from the red edge channel frequently used in the classification of the objects. Besides, its spatial resolution is suitable for the representation of the objects in this database. Comparison of the two classifications based on DMC images reveals that the different acquisition dates of these images (spring and summer) have not influenced significantly in the results.

When two images are combined, the most accurate classifications are obtained in combinations of RapidEye image with the other images. The most accurate classification is obtained with the combination of RapidEye image and the aerial orthoimage. In this case, the overall accuracy of the classification with both images is 91.58%, two points higher than the obtained with only the RapidEye Image.

The combination of three images produces slightly accuracy increases when the RapidEye and the aerial images are combined with the DMC1 image. This increase is so short that addition of the DMC image is not justified. In the rest of combinations the overall accuracies obtained are smaller than the obtained with the RapidEye and the aerial images.

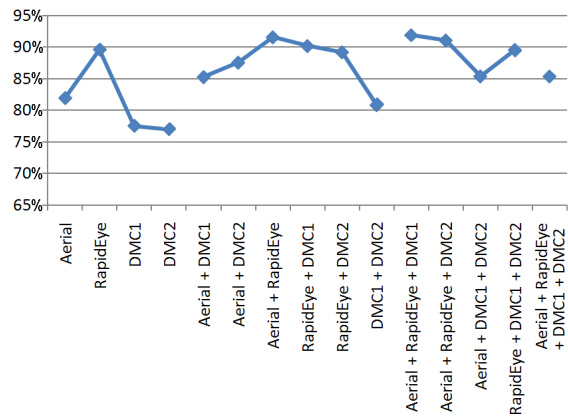


Figure 6: Global accuracies obtained for the classifications based on the entire possible combinations of input data.

To join the features derived from the four available images does not increase the overall accuracies obtained in the previous tests.

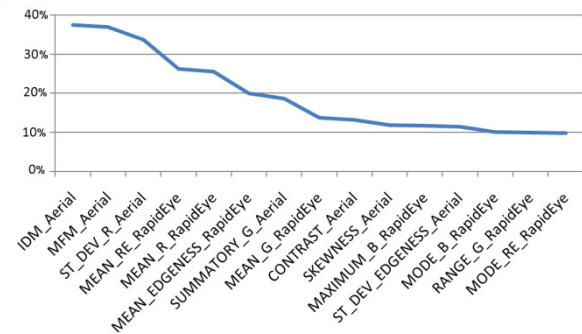


Figure 7: Relevance of features in classification of the aerial image combined with the RapidEye image.

The relevance of features in the most accurate classification (Figure 7) shows that aerial image based features are combined alternately with RapidEye derived features. Features derived from aerial image describe mainly the texture of the objects, whereas the RapidEye based features provide the spectral description.

Table 3 shows the confusion matrix for the most accurate classification obtained in combination of RapidEye with the aerial image. In this classification, urban and forest objects are classified accurately, obtaining the biggest confusion ratios between the cropland and grassland objects.

CLASSIF.	REFERENCE				
	Urban	Cropland	Grassland	Forest	TOTAL
Urban	1467	12	13	7	1499
Cropland	4	565	45	2	616
Grassland	14	115	250	34	416
Forest	18	9	15	882	924
TOTAL	1506	701	323	925	3455

Table 3: Confusion matrix for the classification of the aerial image combined with the RapidEye image.

5.2 Change detection

In this work, the amount and type of occurred changes with respect to the database is known, therefore it is possible to analyse the efficiency of the change detection methodology. The highest efficiency in the change detection process based on classifications of only one image is obtained with the RapidEye image (Figure 8) reaching a value of 90%, obtained as the addition of correctly detected categories. Slightly increases in efficiency are obtained when RapidEye image is combined with the aerial image or with the aerial image and the first DMC image, achieving 92.1%.

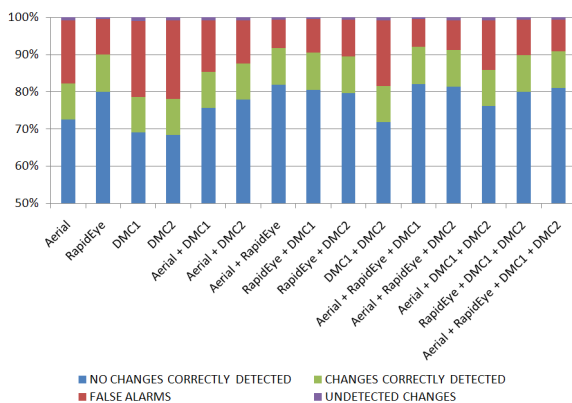


Figure 8. Performance of the change detection method in the whole dataset.

In the same way, the lowest percentage of false alarms (7.6%) and the lowest percentage of undetected changes (0.64%) are obtained with the combination of RapidEye and the aerial image. In this case, the percentage of parcels rejected from the system is 17.5%, detecting the 9.9% of changed parcels and remaining a 0.64% of undetected changes.

The confusion matrices of the change detection with the combination of the aerial and the RapidEye images for the classes croplands and grasslands are shown in the tables 4 and 5. In both classes, efficiency is around 80%, the percentages of false alarms are relatively high and these are parcels to be reviewed unnecessarily, whereas the percentages of undetected changes are quite low ensuring a high degree of accuracy of the change detection process.

Reference		System	
		Accepted	Rejected
Reference	Unchanged	67.33%	16.83%
	Changed	2.14%	13.70%

Table 4: Confusion matrix for the cropland objects.

Reference		System	
		Accepted	Rejected
Reference	Unchanged	74.30%	20.74
	Changed	0.93%	4.03%

Table 5: Confusion matrix for the grassland objects

6. CONCLUSION

We describe an object-based methodology for the automatic discrimination of the farmland types cropland and grassland combining multitemporal images with different spatial resolutions. The object-based approach enables to combine effectively different data sources for the description of the objects contained in a geospatial database.

The most accurate classifications are obtained with the features derived from the RapidEye imagery. The combination of RapidEye image derived spectral features with aerial image texture features produces a slightly accuracy increase.

When an accurate classification is available, the change detection method proposed enables to reduce significantly the amount of objects rejected from the system ensuring a low degree of undetected changes.

ACKNOWLEDGEMENTS

This work was supported by the German Federal Agency for Cartography and Geodesy (BKG) and by the Polytechnic University of Valencia in the frame of the Program PAID-00-10.

REFERENCES

- Balaguer, A., Ruiz, L. A., Hermosilla, T., Recio, J.A., 2010: Definition of a comprehensive set of texture semivariogram features and their evaluation for object-oriented image classification. In *Computers & Geoscience*, vol. 36, 2010, S. 231- 240.
- Busch, A., Gerke, M., Grünreich, D., Heipke, C., Liedtke, C. E., Müller, S., 2004. Automated verification of a topographic reference dataset: System design and practical results. *IntArchPhRS XXXV-B2*, pp. 735-740.
- Canny, J.F., 1986. A computational approach to edge detection. *IEEE TPAMI* 8(6), pp. 679-698.
- Gimel'farb, G.L., 1996. Texture Modelling by Multiple Pairwise Pixel Interactions. *IEEE TPAMI* (18), pp.1110-1114.
- Haralick, R.M., Shanmugam, K., Dinstein, 1973. Texture features for image classification. *IEEE Trans. systems, man. & cybernetics*, SMC-3, pp. 610-622.
- Helmholz, P., Rottensteiner, F., Heipke, C., 2010. Automatic quality control of cropland and grassland GIS objects using IKONOS Satellite Imagery. *IntArchPhRS* (38), Part 7/B, Vienna, pp. 275-280.
- Itzerott, S. and Kaden, K., 2007. Klassifizierung landwirtschaftlicher Fruchtarten. *PFG* 2(2007), pp. 109- 120.
- Janssen, L. L. F., van Amsterdam, J.D., 1991. An Object Based Approach to the Classification of Remotely Sensed Images. In *IGARSS*, 1991, S. 2191- 2195.

Laws, K.I., 1985. Goal-directed texture image segmentation. *Applications of Artificial Intelligence II*, SPIE (548), pp.19-26.

Müller, S.; Heipke, C.; Pakzad, K., 2010. Classification of farmland using multitemporal aerial images. *IntArchPhRS* vol. XXXVIII-4-8-2/W9, Haifa, 2010, p. 70-74 (on CD-ROM).

Quinlan, J. R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers.

Rengers, N. and Prinz, T., 2009. JAVA-basierte Texturanalyse mittels Neighborhood Gray-Tone Differenz Matrix (NGTDM) zur Optimierung von Landnutzungsklassifikation in hoch auflösenden Fernerkundungsdaten. *PFG* 5(2009):455- 467.

Ruiz, L.A., Fdez-Sarria, A., Recio, J.A., 2004. Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study. *IntArchPhRS*, vol. XXXV, pp. 1109- 1115.

Ruiz, L.A., Recio, J.A., Fernández-Sarria, A., Hermosilla, T., 2011. A feature extraction software tool for agricultural object-based image analysis. *Computers and Electronics in Agriculture*. (Article in Press, online available)

Trias-Sanz, R., 2006. Texture orientation and period estimation for discriminating between forests, orchards, vineyards, and tilled fields. *IEEE TGARS* vol. 44(10), pp. 2755- 2760.