

Detecting urbanization changes using SPOT5 ☆

V. Lacroix ^{a,*}, M. Idrissa ^a, A. Hincq ^a, H. Bruynseels ^b, O. Swartenbroeckx ^b

^a *Royal Military Academy, Signal and Image Centre, Avenue de la Renaissance 30, B-1000 Bruxelles, Belgium*

^b *National Geographic Institute, Abbaye de la Cambre 13, B-1000 Bruxelles, Belgium*

Available online 28 September 2005

Abstract

An automatic system to estimate the urbanization changes on the Belgian territory, using SPOT5 images and the National Geographic Institute vectorial database is proposed. The images and the vectorial data are first co-registered. Then, the vectorial database is projected and dilated to produce a mask representing the old status of the database. On the other hand, a fusion of two classification processes on the images enables to extract the built-up area and the communication network, providing a mask representing the actual state of the urbanization in the zone. The comparison between the two masks gives a coarse information of the changes.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Cartography; Change detection; SPOT5; Built-up area detection

1. Introduction

Demand for up-to-date geographic data is increasing, due to the fast changes in many regions, and as the result of the spreading of many GIS applications in everyday life. With the arrival of high resolution sensors, space images are becoming a good source of information to gather knowledge, and to track changes.

In a near future, the National Geographic Institute of Belgium (NGI) will manage its vectorial data ranging from a conceptual scale of 1:10 000 to 1:50 000 in one single database (DB). NGI is setting up a “planning tool” to schedule the data updating process according to the changes that occurred on the field and to compute the up-to-date status of the data as information to provide to end-users. The information about the changes will come from various sources, in particular, from remote sensed data. The sensor should be such that (i) the cost of a regular territorial coverage

should be affordable; (ii) the regular territorial coverage should be technically possible; (iii) its resolution should enable the detection of changes in the built-up area and in the communication network. According to a visibility test (Lacroix et al., 2004), SPOT5 panchromatic 5 m resolution data fused with multi-spectral data seem sufficient for some photo-interpretor to detect most of sets of buildings, but not individual ones and most of the road network in open area.

In the following, we will first review some work in change detection, then expose the global strategy, and finally detail the proposed coarse change detection method.

2. State of the art

A summary of change detection methods can be found in (Li et al., 2002). Most articles deal with changes between images, and not with changes between a database and an image, often considered as a feature extraction problem. In (Vosselman and de Gunst, 1997), knowledge is used for updating road maps; the old road position is compared to the image using intensity profile. If change is observed, hypothesis of changes are made. While the incorporation of knowledge about possible lane widths and exit angles improves the interpretation results, many changes are in

☆ Colored figures are available at www.sic.rma.ac.be/Projects/ETATS_PRL/PRL.html. This study is funded by the STEREO program of the Belgian Science Policy.

* Corresponding author. Fax: +32 2 737 6472.

E-mail address: vinciane.lacroix@elec.rma.ac.be (V. Lacroix).

fact false alarms because of disturbing objects like shadows, trees and cars. In (Busch, 1998), a method is proposed to perform the revision of built-up area in a GIS using satellite imagery and GIS data. The satellite images are SPOT and IRS-1C with a ground resolution of 10 and 5.8 m, respectively. The built-up area is detected on the basis of short edge densities. A threshold obtained thanks to training on GIS data is used to separate the built-up area. The changes are then observed comparing the classified zones with the GIS data. Klang (1998) proposed an automatic detection of changes in road DB using satellite imagery. The considered satellite data are Landsat, SPOT, IRS-1C, all re-sampled to 10 m resolution. The projected road vectorial DB is matched to detected roads. Statistics over the latter are used to find a threshold which serves at extracting seed points of potential new roads where a line tracking is started. Finally, the changes with the road data base are extracted. In an invited paper, Baltasvias (2002), provides the state of the art in object extraction and revision by image analysis using existing geospatial data and knowledge. The paper mainly focuses on multi-looking aerial images or satellite images of 1 m resolution. As far as the detection of man-made objects in satellite images is concerned, some researchers have also used NDVI and edges in the form of a complexity index (Sakamoto et al., 2004).

3. Overall strategy

Changes are located by comparing a mask generated by the DB projected on the image, to the output of a classifier extracting the built-up area and the road network, called the “man-made” or MM class. We assume that the built-up area and communication network generate structures and texture in the panchromatic image. Therefore, changes inside the “old” DB extent and changes in attribute such as the road width will probably not be noticeable. On the other hand, the system should detect as change, places where the DB indicates roads or buildings while they do not exist.

The strategy is summarized in Fig. 1. NGI’s vectorial DB and SPOT5 images are the input of the system. NGI filters the DB to produce vector layers containing only the built-up area, the road network, and the hydrography. The road network and the built-up area are used to produce the “Old Mask”, representing the old extent of the

MM class. The images are registered with the vectorial DB using the data registration process. Then, on the one hand, the registered panchromatic image is analyzed by a “Texture and Structure algorithm” that separates textured from non-textured areas. On the other hand, the normalized difference vegetation index (NDVI) computed from the multi-spectral images, provides another two-classes separation: vegetation and non-vegetation areas. The fusion of both classifications from which the hydrological network is removed, is compared with the “Old Mask” to generate a “Change Map”.

4. Data registration

If a digital elevation model (DEM) is available, orthorectified images can be produced ensuring that images are well-positioned under the vectorial layer. As NGI does not have a DEM on the whole Belgian territory, another solution is sought.

A set of ground control points (GCP), evenly distributed, are used to obtain a second order polynomial georeferencing function g . If the RMS error of the vectorial DB projected with g^{-1} is larger than 5 m (a typical width of a secondary road), the image is cut in cells, and long bright lines are detected using the gradient line detector (GLD) (Lacroix and Acheroy, 1998); the latter exploits the fact that the gradient of the intensity is pointing towards/against each other at each side of a bright/dark line. Thus, in the 8-neighbourhood of each pixel, the maximum of the dot product of intensity gradient of symmetrical pixels is computed, if they both points towards the current pixel as seen in Fig. 2.

Then, a non-maximum suppression and line following is performed, as for edge extraction. Only long straight lines lying in the vicinity of the projected road network are considered as potential match. For each road, the best compatible line segments are considered, and a least square procedure (Borghys, 2001) is used to find the best local affine transform f_i fitting the detected lines of cell i under the projected vectors. The image is then re-sampled using bi-cubic interpolation.

The result of this process is shown in Fig. 3. If the RMS error in a cell containing a GCP exceeds 5 m, additional points are asked to the user, and another local function is computed. This option was not necessary over the two test areas (regions of Sint Niklaas and Brussels).

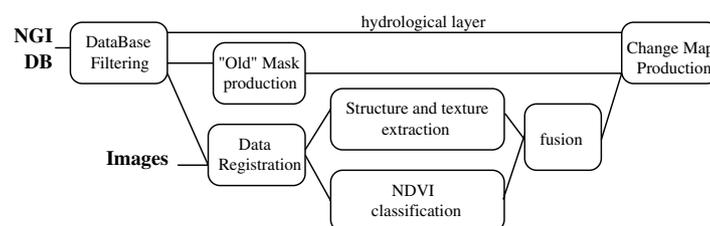


Fig. 1. The global strategy.

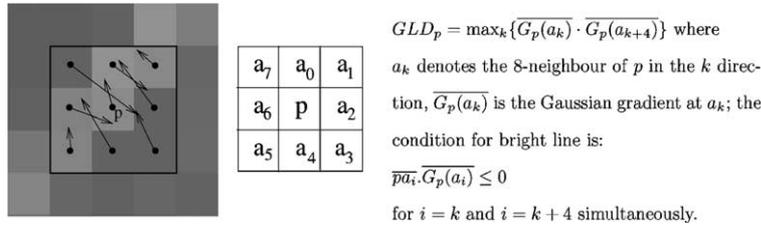


Fig. 2. Gradient field around bright line and GLD computation.

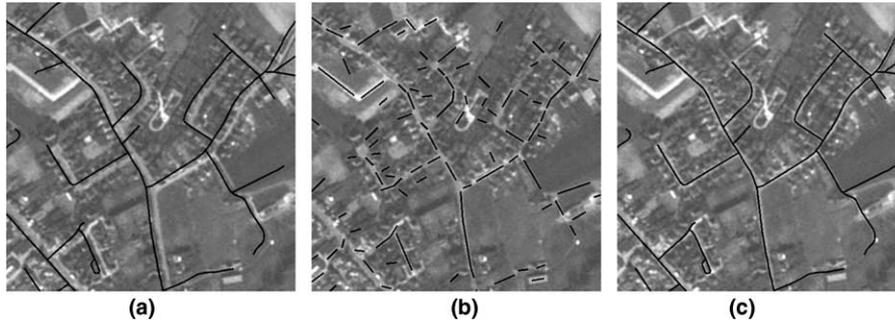


Fig. 3. (a) Projected DB; (b) bright lines in the neighbourhood; (c) registered data.

5. Structures and texture detection

A filter providing a high value for the MM class is sought. At the scale under study, roads, small and large buildings, will form bright ridges, spots and rectangles, respectively. Edge detectors have been used in (Busch, 1998) and in (Sakamoto et al., 2004) to detect built-up areas. However, they will respond to the borders of fields and forests, introducing false alarms. In that respect, ridge detectors should be better, but will miss large buildings. Gabor filters on the other hand could detect both.

5.1. Gabor filters

Gabor filters perform a local Fourier analysis thanks to sine and cosine functions modulated by a Gaussian window. In the complex space these filters are defined as

$$G(x, y, k_x, k_y) = e^{-\frac{(x-X)^2 + (y-Y)^2}{2\sigma^2}} \cdot e^{j(k_x x + k_y y)}, \quad (1)$$

where x, y represent the spatial coordinates while k_x, k_y represent the frequency coordinates. X and Y are the spatial localization of the Gaussian window.

In this paper, two simplifications are proposed. The first one makes use of short time Fourier transform (STFT). The second introduces binomial window as approximation of the Gaussian. The basis functions of this decomposition are

$$S_{k,l}(n, m) = W(n, m) \times \sin 2\pi \left(\frac{kn}{N+1} + \frac{lm}{M+1} \right) \quad \text{and} \quad (2)$$

$$C_{k,l}(n, m) = W(n, m) \times \cos 2\pi \left(\frac{kn}{N+1} + \frac{lm}{M+1} \right),$$

where $W^2(n, m) = \frac{1}{2^{(N+M)}} C_N^{\left(\frac{N}{2}+n\right)} C_M^{\left(\frac{M}{2}+m\right)}$ is the $(N+1) \times (M+1)$ binomial window. The coefficients are given by

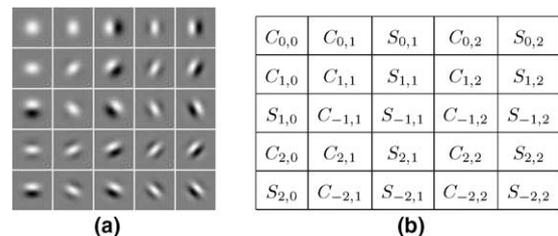
$$C_j^i = \begin{cases} \frac{j!}{i!(j-i)!} & 0 \leq i \leq j \\ 0 & \text{elsewhere} \end{cases} \quad (3)$$

The indexes $n = -\frac{N}{2}, \dots, \frac{N}{2}$ and $m = -\frac{M}{2}, \dots, \frac{M}{2}$ are the window spatial coordinates with N and M even integers. The filters selectivity in frequency (expressed in number of cycle per pixel) and orientation (in radian) are derived from the following equation:

$$f = 2\pi \sqrt{\left(\frac{k}{N+1} \right)^2 + \left(\frac{l}{M+1} \right)^2} \quad \text{and} \quad (4)$$

$$\theta = \arctan \frac{k}{N+1} \frac{M+1}{l}$$

with $k = -\frac{K}{2}, \dots, \frac{K}{2}$, $l = -\frac{L}{2}, \dots, \frac{L}{2}$ and $K \leq N$, $L \leq M$. Fig. 4(a) shows a subset of filter functions for a 31×31 binomial window with $K=4$ and $L=4$, while Fig. 4(b) shows their corresponding sine and cosine terms.

Fig. 4. (a) Filter bank for $N = M = 30$ (grey = 0, white = positive values, black = negative values), (b) filters legend: $S_{k,l}$ and $C_{k,l}$ are respectively the sine and cosine terms.

A set of feature images is obtained by convolving the image I with each filter providing a local “Energy” (Kruizinga and Petkov, 1999) given by

$$E_{k,l}(n, m) = [(C_{k,l} \otimes I)(n, m)]^2 + [(S_{k,l} \otimes I)(n, m)]^2 \quad (5)$$

Here, \otimes denotes the convolution and k, l the frequency–orientation coordinates.

Fig. 5(b) illustrates the effect of the application of the selected filter bank on the area displayed in Fig. 5(a). Notice the grey background generated by cosine filters, due to their response at low center frequencies. As the detection should not involved the DC component of the signal, the DC component introduced by these filters should be subtracted. The considered energy is therefore

$$E(n, m) = \max_{k,l=-\frac{N}{2}, \dots, \frac{N}{2}} [((C_{k,l} - DC_{k,l}) \otimes I)(n, m)]^2 + [(S_{k,l} \otimes I)(n, m)]^2. \quad (6)$$

The energy is averaged using a binomial window of the filter size. For symmetry reason, N is set equal to M . In order to have the largest set of frequencies and orientations, K and L are set to N .

5.2. Finding the filter parameters

A “learning” region over which the DB is up-to-date, and containing most of the various structures of buildings and roads one could probably meet in Belgium, is used to find the best filters.

According to Eq. (6), N is the only parameter to set to compute the energy; N being set, the threshold separating the textured from non-textured area is another parameter to determine. Each set of $C_{k,l}$ and $S_{k,l}$ filters corresponding to a given window size are first evaluated, then, the best threshold separating the two classes is sought.

Sensitivity and specificity are classification indices often used to evaluate a two-classes classifier (Provost and Fawcett, 2001). In this case, the two classes are respectively 1 for the MM class and 0 for the rest. The indices are based on the confusion matrix made of the ground truth and of the classification results, as shown in Fig. 6. For example, “TP” in this table is the number of pixels that are considered as belonging to the MM class according to ground truth and detected as such by a classifier. Sensitivity s and specificity p are defined as follows:

$$s = TP / (TP + FN) \quad \text{and} \quad p = TN / (FP + TN) \quad (7)$$

where TP, FP, TN and FN denote True Positive, False Positive, True Negative and False Negative, respectively.

For each filter set, there is a compromise to solve: a low threshold of the energy texture measure will generate a lot of false alarms (low specificity) while providing a good sensitivity, and vice versa. The receiver operating characteristic (ROC) analysis consists in analyzing the plot of the “TP rate” (s) in function of the “FP rate” ($1 - p$) (Provost and Fawcett, 2001), summarizing the evolution of this compromise when the threshold is modified. In this representation, the more a curve tends to the step function at 0, the better it is. Having a serie of ROC curves, the ideal combined method would consist in choosing for each threshold, that is, for each FP rate, the method providing the highest TP rate (Provost and Fawcett, 1997). Actually, the point closest to the (0,1) point provides the best compromise.

ground truth \rightarrow	1	0
detected classes \downarrow	1	FP
	0	FN
		TN

Fig. 6. Confusion matrix.

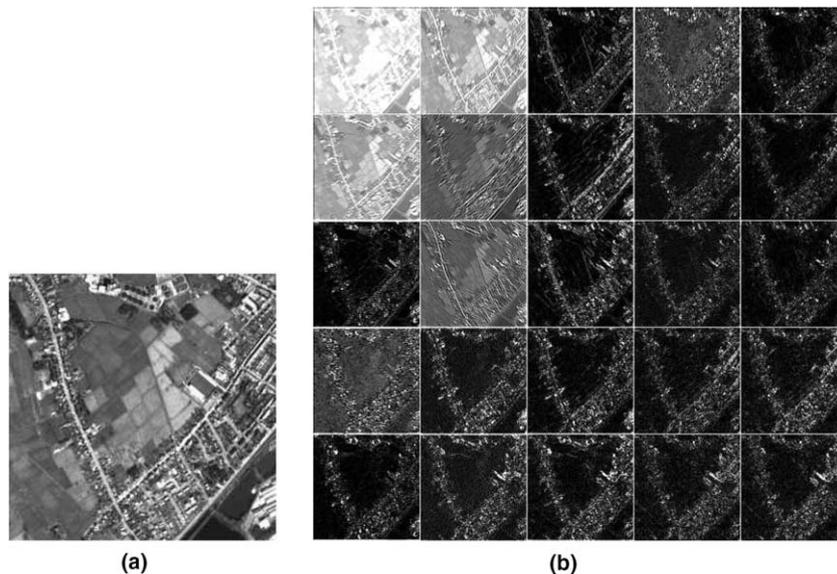


Fig. 5. (a) Part of a SPOT5 panchromatic image (b) feature images.

The area under the curve (AUC) is another cue for selecting a classifier that will perform better in average. However, the system will mostly be used in the upper left region of the curve, that is, below 20% of false alarms and above 80% of good detection. Therefore, it is rather the AUC in that interval that matters. Three sets of filter parameters ($N + 1 = 5, 7$ and 9) were tested on a SPOT5 5 m resolution panchromatic in the region of Sint-Niklaas. The ground truth is taken by projecting the DB vectors on the image according to the registration process already described, and by dilating them. The global AUC values are all 0.92, while the AUC in the interval of interest are 0.45, 0.52 and 0.54, respectively. $N + 1 = 7$ has been chosen for a best performance-complexity compromise.

As a basis of comparison, we have drawn in Fig. 7 the ROC curves of Laws filters (Laws, 1980) and of the DCT considered as efficient in texture classification (Randen, 1999), as well as the ridge detector (Lacroix and Acheroy, 1998) already used for the data registration in Section 4, the complexity index (Sakamoto et al., 2004) and the NDVI as defined in Eq. (8). At low false alarm rate, the ridge detector is the best choice, while at an alarm rate around 12%, the Gabor energy becomes better.

5.3. Finding the appropriate threshold

If the ground truth is available on part of the image on a region representative enough of the built-up structures of the whole image, the threshold corresponding to the closest point to $(0, 1)$ on the ROC curve should provide the best compromise. In the context of NGI, however, this informa-

tion will not be available. A threshold based on a learning region extracted from another image cannot be used neither, as the energy will depend on the image.

The distribution of texture values can be modeled by a mixture of two γ distributions. As an iterative algorithm to obtain the parameters of these distributions is time-consuming, an initial guess assuming a Gaussian mixture is used. Starting from the a priori information given by the old mask, an initial mean and variance for the MM and non-MM classes are used as an initial guess. Once the parameters of the two γ distributions are obtained, a Maximum A Posteriori rule is used to separate the two classes.

Fig. 8 shows the histogram of the texture values in the Sint-Niklaas region. The initial Gaussian distributions are displayed, and the final mixture of γ distributions resulting from the iterative algorithm shows a good accordance with the actual histogram.

5.4. Fusing texture and NDVI

If the texture measure is considered alone, some industrial buildings characterized by a low reflectance and a relatively homogeneous roof may be missed. The introduction of the NDVI enables to extract these areas. The NDVI computed on SPOT5 image is defined as

$$\text{NDVI} = \frac{XS_3 - XS_2}{XS_3 + XS_2}, \quad (8)$$

where XS_3 and XS_2 denote the infrared band, and the red band, respectively. In order to obtain two classes (vegeta-

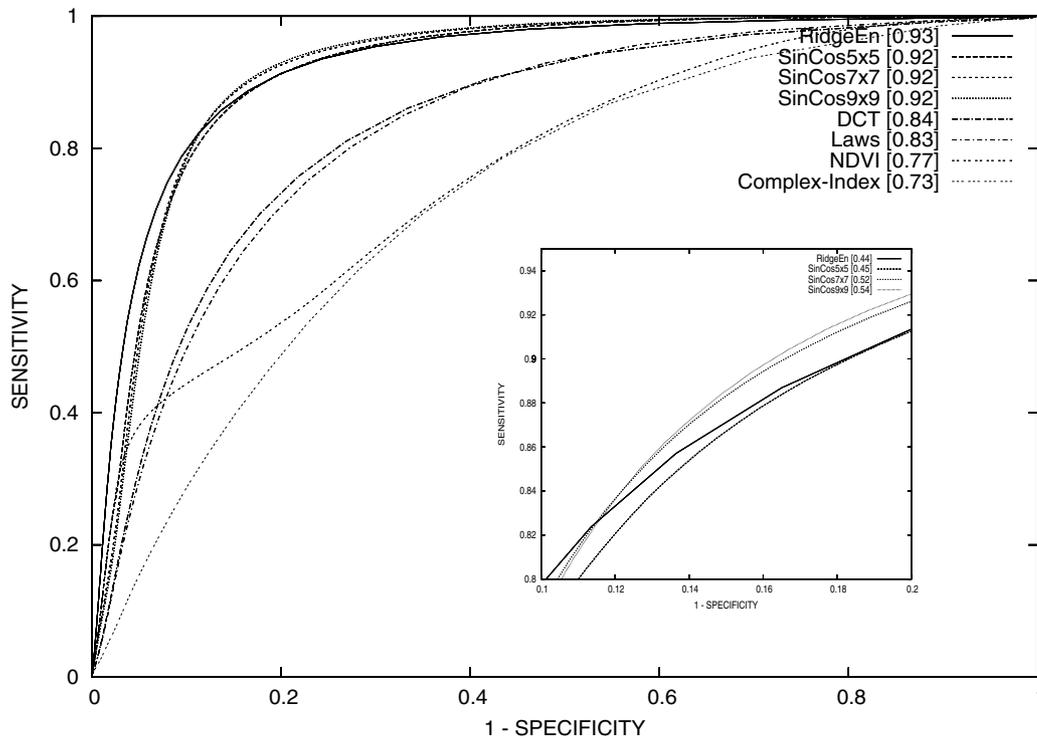


Fig. 7. ROC curve.

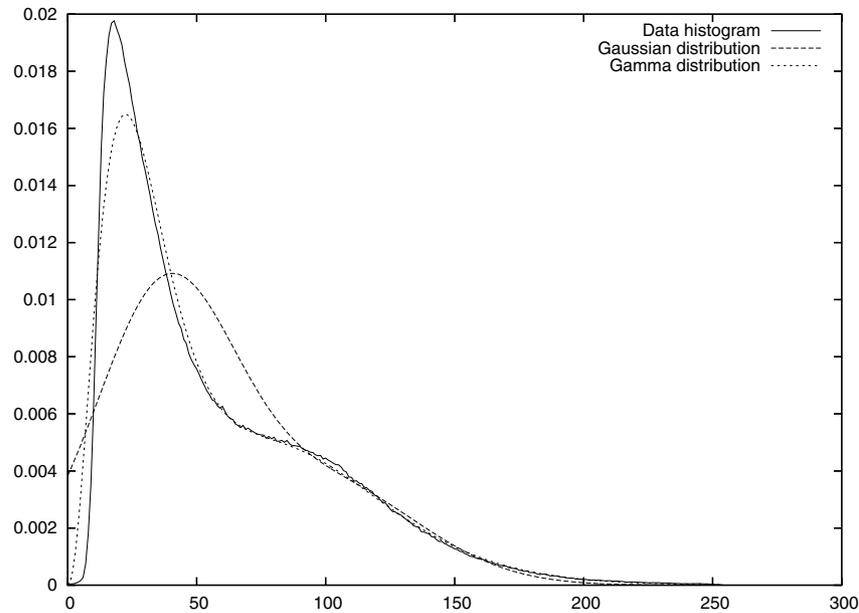


Fig. 8. Texture histogram over Sint-Niklaas.

tion and non-vegetation), a fuzzy clustering (Gath and Gev, 1989) algorithm is used.

As the hydrological network may produce texture or should be classified as non-vegetation thanks to the NDVI index, the fusion rule is: (texture “OR” non-vegetation) “AND” (“NOT” in the hydrology layer). That is, if the pixel belongs to the built-up zone according to the texture algorithm or to the NDVI classification then, and is not in the hydrological layer, it is considered as belonging to the MM class.

6. Results

A Change Map over the Sint-Niklaas region has been produced. Fig. 9 shows the considered area. Fig. 10

shows the overlapping of the mask and of the final classifier: the detected DB elements (TP), corresponding to no change, are displayed in dark grey (0.40), or in an average grey (0.55), if they were detected by the NDVI only; the missing DB elements (FN) are in light grey (0.70), and zones of potential changes (FP) in black (0), or in deep grey (0.25) if they were detected by the NDVI only.

An analysis of these results shows that for the missing DB elements (FN): (i) one area corresponds to a “ghost area”: buildings and roads were planned in this area but were never built; (ii) some secondary roads are missed; (iii) several areas are situated in forest, thus occluding buildings and roads; (iv) the information brought by the NDVI only is quite small.



Fig. 9. © SPOT5 5 m resolution over Sint-Niklaas region.

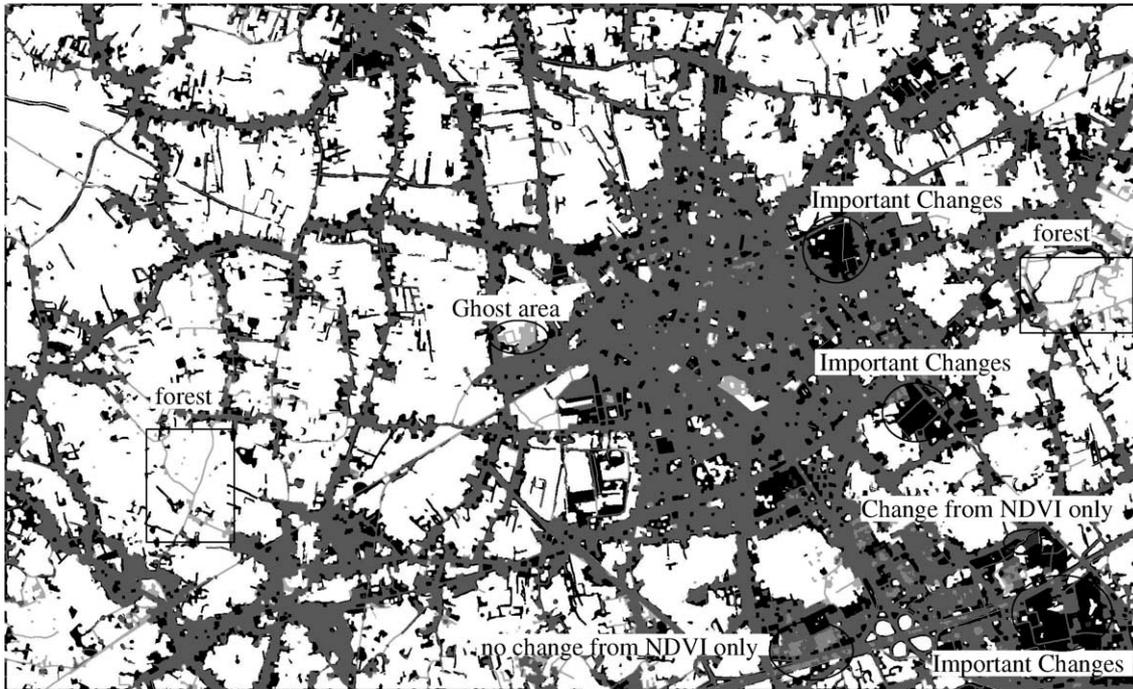


Fig. 10. Change map; dark average grey [0.40]: no change (MM → MM); average grey [0.55]: no change detected by NDVI only; light grey [0.70]: missed objects or DB errors (MM → non-MM); black [0]: changes or false alarms (non-MM → MM); deep grey [0.25]: false alarms or changes detected by NDVI only; white [1]: no change (non-MM → non-MM).

The system has then been tested and evaluated on 10 zones selected on SPOT5 5 and 2.5 m resolution images in order to include sub-urban and rural areas. The results were compared to the results of visibility tests performed by an experimented photo-interpreter. Fig. 11 shows that the number of elements missed by the system are of the same

order (10%) in half of the experiments, independently of the landscape type or image resolution, while it is of approximately 20% in the other cases. In all cases, the system is presenting many more false alarms, which could probably be reduced by a post-processing. Finally, it is not sure whether or not the cost of multi-spectral images is justified.

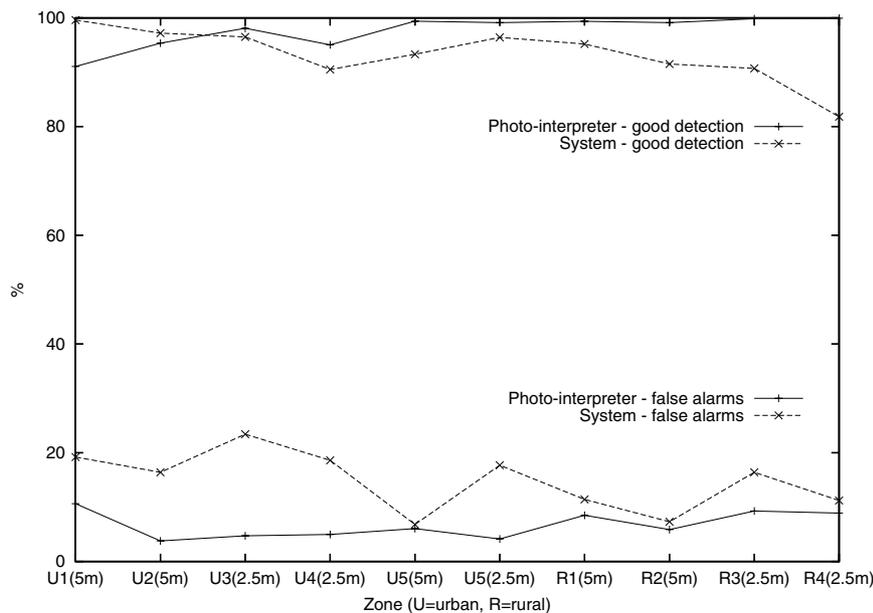


Fig. 11. Good detection and false alarm rate of 10 zones spread over two images (5 and 2.5 m). On X axis, U_i and R_i denote urban and rural zones, respectively. The image resolution is given in parenthesis. Zone U5 is common to both images.

7. Discussion and further work

ROC analysis has been used to select Gabor Filters to detect built-up areas and road network in SPOT5 images. The results show more than 90% of good detection in half of the experiments, more than 80% in the others, and approximately 20% of false alarms. Of course, zones under the forest cannot be analyzed. Therefore, an automatic forest detection algorithm would be useful to highlight zones over which the system cannot judge the urbanization status. While large change areas are detected by the system, many false alarms remain. A filtering of these regions based on some attributes such as the area, the shape, or some radiometric measurements might reduce them. A quantitative evaluation of the ability of the system to detect the changes in open area should be performed after this filtering step.

References

- Baltsavias, E., 2002. Invited paper, object extraction and revision by image analysis using existing geospatial data and knowledge: state-of-the-art and steps towards operational systems, In: Commission VI Mid-Term Symposium, vol. XXXIV, Part 2, Xi'an, China, Int. Arch. Photogramm. Remote Sensing (IAPRS), pp. 13–22.
- Borghys, D., 2001. Interpretation and Registration of High-Resolution Polarimetric SAR Images. PhD thesis, ENST, Dpt. TSI, Paris.
- Busch, A., 1998. Revision of built-up areas in a GIS using satellite imagery and GIS data. In: International Archives of Photogrammetry and Remote Sensing, ISPRS Commission IV—GIS Between Visions and Applications WG IV/III.2, vol. 32.
- Gath, I., Gev, A.B., 1989. Unsupervised optimal fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. 11 (July), 773–780.
- Klang, D., 1998. Automatic detection of changes in road data bases using satellite imagery. In: Proc. Internat. Archives of Photogrammetry and Remote Sensing, vol. 32, pp. 293–298.
- Kruizinga, P., Petkov, N., 1999. Nonlinear operator for oriented texture. IEEE Trans. Image Process. 8 (October), 1395–1407.
- Lacroix, V., Acheroy, M., 1998. Feature-extraction using the constrained gradient. ISPRS J. Photogramm. Remote Sensing 53 (April), 85–94.
- Lacroix, V., Hincq, A., Mahamadou, I., Bruynseels, H., Swartenbroeckx, O., 2004. A visibility test on spot5 images, In: ISPRS 2004, vol. Commission 4, Istanbul, Turkey.
- Laws, K. 1980. Rapid texture identification. In: Proc. SPIE Conf. Image Processing for Missile Guidance, vol. 28, pp. 376–380.
- Li, D., Sui, H., Xiao, P., 2002. Automatic change detection of geo-spatial data from imagery. In: Integrated System for Spatial Data Production, Custodian and Decision Support, Xi'an, China, Int. Arch. Photogramm. Remote Sensing (IAPRS).
- Provost, F.J., Fawcett, T., 1997. Analysis and visualization of classifier performance: comparison under imprecise class and cost distributions. Knowledge Discovery and Data Mining, pp. 43–48.
- Provost, F., Fawcett, T., 2001. Robust classification for imprecise environments. Mach. Learn. 42, 203–231.
- Randen, J.H.T., 1999. Filtering for texture classification: a comparative study. IEEE Trans. Pattern Anal. Mach. Intell. 21 (April), 281–310.
- Sakamoto, M., Takasago, Y., Uto, K., Doihara, T., Kakumoto, S., Kosugi, Y., 2004. Automatic detection of damaged area of iran earthquake by high-resolution imagery. In: Proc. IGARSS2004, Alaska.
- Vosselman, G., de Gunst, M., 1997. Updating road maps by contextual reasoning. In: Proc. 2nd Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona.