DETECTION OF BUILDINGS IN COLOUR AND COLOUR-INFRARED AERIAL PHOTOS FOR SEMI-AUTOMATED REVISION OF TOPOGRAPHIC DATABASES

Thomas Knudsen and Brian Pilemann Olsen

Kort- & Matrikelstyrelsen, 8 Rentemestervej, DK-2400 Copenhagen NV, Denmark

`mailto:{thk|bpo}@kms.dk` `http://research.kms.dk/~{thk|bpo}`

KEY WORDS: Change Detection, Urban, Remote Sensing, Automation, Updating, TOP10DK

ABSTRACT

With the growing use of digital map data, the requirements for frequent and efficient revision of digital map databases is equally growing. Revision of map databases can be divided into 3 steps: Change detection, classification of changes and registration/updating of the thematic layers of the database. Often the change detection step is carried out in an entirely manual workflow, which calls for automation since it is tedious, labour intensive, and hence very costly. In this paper, we concentrate on automated change detection for the "building" layer in a fully 3D geo-spatial database. We present a method based on the use of unsupervised classification as supervising input to a Mahalanobis classification algorithm. The method is evaluated in test cases based on building registrations from the Danish TOP10DK map database, in combination with a plain RGB colour aerial photo and a colour infra red (CIR) aerial photo covering the same area. The test case presented here is from a suburban residential area. At the present stage, plain RGB aerial photos are not enough to significantly reduce the update task. CIR photos show more promise, as the change algorithm, in general, detects all new buildings, although it still needs refinements to reduce the number of false alarms. Buildings with flat asphalt roofs are an entirely separate problem, since, in the lack of height information, they are extremely difficult to discern from roads.

1 INTRODUCTION

The map production at the National Survey and Cadastre—Denmark (Kort & Matrikelstyrelsen, KMS) is based on the TOP10DK database. TOP10DK is a fully 3D topographic map database, where geo-spatial input data was registered photogrammetrically from aerial photos, in order to meet the accuracy requirements and topological specifications (Kort & Matrikelstyrelsen, 2001).

The update process for TOP10DK is also based on aerial photos, and since the entire production cycle is digital, we may eliminate much of the tedious and costly manual change detection work by introducing digital automated change detection methods.

We present a procedure for change detection concentrating on buildings and based on a direct comparison between the map database and new aerial images (i.e. a raster-to-vector comparison). Buildings are particularly important mapping objects, as they are good indicators for urban dynamics, but also notoriously hard to detect, since they are often spectrally ill defined and highly diverse (cf. figure 1).

The aim of the procedure is to detect at least the same percentage of changes as a manual operator does. On the other hand, generation of a number of false alarms is acceptable as long as the time spent for the evaluation and 3D object registration for an area is still reduced (the 3D object registration is the final step in the update process, and is not considered further here).

Projects with similar aims are going on in Germany, for updating the ATKIS map database. ATKIS is comparable to TOP10DK but has weaker precision requirements: 3 m for ATKIS, vs. 0.5 m for TOP10DK (Arbeitsgemeinschaft der Vermessungsvervaltungen der Länder der Bundesrepublik Deutschland, 1988, Kort & Matrikelstyrelsen, 2001). The ATKIS update projects focus on registering more generic surface types (water, forest, settlement, greenland, street) using a supervised maximum likelihood classification with training areas created from the existing ATKIS registrations (Walter and Fritsch, 1998, Walter and Fritsch, 2000, Walter, 2000a, Walter, 2000b). Reduction to generic surface types results in spectrally more well defined targets which have shown to be automatically detectable with excellent precision, especially if near-infrared and/or height observations are available (Petzold and Walter, 1999, Petzold, 2000).

For the training area creation, we follow a route similar to the ATKIS update work, but since we aim at detecting individual buildings we do not follow the strategy of reduction to generic surface types. Additionally, the ATKIS update work is based on digital orthophotos (which has a direct geometric relation to the map coordinate system), where we base our work on the projection of map data directly onto the raw aerial photo using stereo modelling. This approach allows the full 3D object registration to be carried out from the stereo photo pair, overlaid by the change map.

Another related project aims at updating the Dutch TOP-10Vector digital map database (Hoffmann et al., 2000, van Asperen, 1996). This project is based on fully digital multi spectral, multi stereo data from the *High Resolution Stereo Camera—Airborne,* HRSC-A (Neukum, 1999). The HRSC-A dataset combines the high-resolution (15 cm) spectral data with the possibility of automatic generation of a high-resolution surface height model. Data at this level of resolution makes it possible to carry out a multi-scale analysis and representation of the data. This leads to a better definition of the map objects, due to the higher stability of the spectral characteristics when aggregating multiple pixels (Hoffmann et al., 2000).

In Switzerland, the ATOMI project (Eidenbenz et al., 2000,

Figure 1: Buildings are often spectrally ill-defined, and highly diverse

Niederöst, 2000, Niederöst, 2001) aims at detecting changes, and enhancing planimetric accuracy from the 5 m level to the 1 m level, for the 2D VECTOR25 database, based on colour aerial photos, a high resolution digital elevation model (DEM) and a photogrammetrically generated digital surface model (DSM). The ATOMI update work uses the DSM as the primary data source, and uses the image information primarily in order to discern man made objects from natural objects.

2 DATA

The development and update of TOP10DK is currently based on colour (RGB) aerial photos. In the present work additional experimental colour infrared (CIR) photos are also used.

All aerial photos used are taken from an altitude of approximately 3800 m, resulting in a coverage of about 6 km $\times$ 6 km on ground, and a nominal scale of 1:25000. The digital versions of these photos are created by scanning at a resolution of 21 μm (48 lines/mm) for a nominal ground resolution of 0.525 m, and a size of 11000$\times$11000 pixels (the case study area used below consist of small fractions of that).

Contrary to the Swiss ATOMI project, we do not have access to DEM and DSM data. We do, however, have access to camera calibration reports and image orientation data, and as TOP10DK vectors are fully 3D they can be projected directly onto the photos with very high precision.

3 METHOD

The change detection method presented here is based on a modular workflow consisting of 7 self-contained steps, labelled *fusion, preprocessing, clustering, classification, iteration, create change map* and *post processing*.

The steps are shown schematically in figure 2, where they are also grouped conceptually into 3 higher level procedures, labelled *preparation*, *classification*, and *detection*.

The preparation procedure (fusion and preprocessing steps) unifies the reference systems for the vector data and the image data and extracts the training set used in the classification procedure.

The classification procedure (clustering and classification steps, running a number of times under the control of the iteration step) does the actual classification.

The detection procedure (change map and post processing steps) detects all potential changes and eliminates spurious elements and changes too small to require an update.

The selfcontainment of the individual steps simplifies experimentation, since each step can be reimplemented using a different algorithm without affecting the rest of the workflow. Each step is described below, with the currently used algorithm as example. In some cases examples of alternative algorithms are given. These examples are given merely for illustrative purposes—not in an attempt to give complete listings.

A main point in the method is that the use of photogrammetric tools is confined to step 1. The remaining steps may be carried out using any plain image processing tool that fits. Plain image processing tools are more abundant (and usually *much* cheaper) than the more specialized photogrammetric tools.

The heart of the method is based on clustering of the spectrally non-uniform building pixels into a number of more uniform classes, using an unsupervised clustering algorithm. These classes are used for controlling a following supervised classification algorithm, which does the actual discrimination between buildings and background.

The two main assumptions behind the use of this procedure for change detection is that (1) the number of changes is much smaller than the total number of buildings. This holds for most urban areas, and that (2) the spectral characteristics of new buildings are similar to those of the existing buildings. This may or may not hold and in cases where it does not hold, the method may fail.

Data fusion: In order to use the objects in the digital map database as training areas for the determination of the class characteristics, the raster image data and the vector map database must be co-registered and combined.

As mentioned previously, we use a projection of the map database onto the raw image for this data fusion process, rather than the more common way of going through an orthorectification process. This implies that we avoid the slight, but undesirable, change in spectral characteristics and degradation of edge representations which are unavoidable when resampling an image. It also retains the possibility of digitizing full 3D polygons off the image.

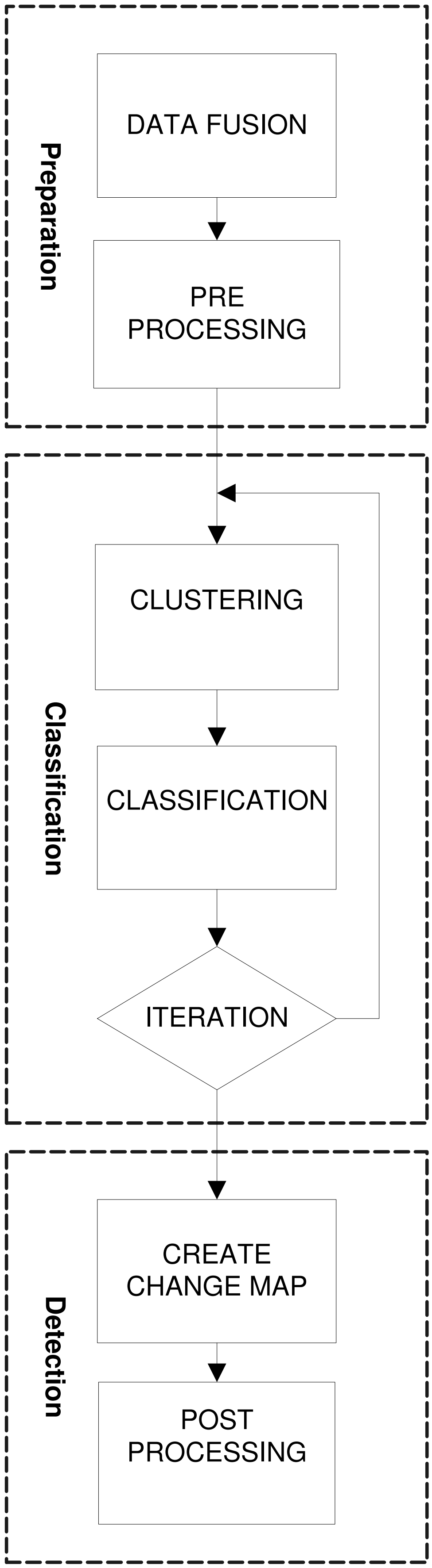

Figure 2: The change detection work flow—cf. section 3 for description

Orientation parameters for the image data and full 3D registrations for all map database points are required in order to project the map data onto the stereo image pair. As these requirements are fulfilled, the image coordinates for all database points can be calculated using the basic photogrammetric equations (Kraus, 1993, e.g.). The polygons of the transformed map data are then rasterized onto an image with the same geometry and resolution as the aerial photo. This new image is used as a mask defining the training areas for the supervised classification.

At this stage the geodetic/photogrammetric part of the work has been totally decoupled from the image analysis part: from here, the work continues on the two input images (the mask and the aerial photo) using ordinary image processing tools and concepts.

The method could also work using orthophotos. The disadvantage of using orthophotos is that to reach sufficient precision in the orthorectification, a high resolution and very precise DSM (including buildings) is required.

Preprocessing: Currently, the preprocessing step only consists of one operation, namely the decorrelation of RGB images, by transforming the R,G,B channels to R-G,G,B-G. Other possible preprocessing operations include the transformation into principal components or into the HSI colour space.

Clustering: As indicated in figure 1, buildings are usually spectrally ill defined and highly diverse. To handle this, all pixels registered as buildings are split up in smaller groups using a simple migrating means clustering process (called PRECLUST) based on the ISODATA algorithm (Ball and Hall, 1965). The initial cluster centers are selected at random positions in phase space, and the initial number of clusters is arbitrarily set to 15. The algorithm is iterative, and during the iterations the cluster centers are moved around algorithmically and the number of clusters is adjusted up and/or down (i.e. clusters are splitted or merged) in the attempt of finding an optimal fit to the actual input data.

The clustering process results in a number of class centers, representing classes which are spectrally more uniform than the initial single building class.

Other clustering methods like the recently published SYNERACT algorithm (Huang, 2002), or neural network based methods (Knudsen et al., 2002) could be used, and may improve the separation process.

Classification: The clusters generated from the pixels representing buildings, are now used in a Mahalanobis classification of the entire image; each image pixel is assigned to the class having the smallest Mahalanobis distance from the class cluster to the pixel value (Richards and Jia, 1999, e.g.). Pixels far from all clusters are not assigned to any class. Here, the meaning of “far” is defined by a somewhat arbitrarily selected threshold value; this implies that we depend on having the building pixels separated from the non-building pixels by a large gap in minimum Mahalanobis distance. Otherwise, the classification will be very sensitive to the actual value of this parameter.

The minimum distance measure could substitute the Mahalanobis distance, for a faster, but probably less stable classification. Fuzzy logic and neural network based methods may also be considered.

Iteration: Since the the clustering process is not necessarily convergent, a stop criterion is necessary. This makes the final class centers dependent on the initial configuration of the cluster centers. Therefore the clustering/classification steps are repeated a small number of times (usually 5). Each run generates a new set of initial cluster centers and results in a new classification. At the end of this *Monte Carlo* process, only pixels identified as buildings in all the (usually 5) steps, are accepted as buildings. This typically results in a significant reduction of misclassifications.

Computing the change map: The actual change detection is carried out by comparing the building mask, generated under step 1, with the set of classification results generated in the iterations of steps 2 and 3. This splits the pixels into a set of 3 groups, defined as follows:
NO CHANGE pixels are *either* pixels registered as buildings and consistently detected as buildings *or* pixels NOT registered as buildings and consistently UNdetected as buildings,
POTENTIAL CHANGE pixels are *either* pixels registered as buildings and sometimes detected as buildings *or* pixels NOT registered as buildings and sometimes detected as buildings,
EVIDENT CHANGE pixels are *either* pixels registered as buildings and consistently UNdetected as buildings *or* pixels NOT registered as buildings and consistently detected as buildings.

Post processing: The change map computed includes all potential changes in the building layer and therefore also contains some blunders and spurious pixels, which must be removed.

This can be done in many ways and the different steps in the post processing may be done in a different order. The first step in the post processing used here is the removal of *no change* and *potential change* pixels leaving only evident changes.

The post processing continues with an elimination of spurious pixels by the use of plain mathematical morphology. A closing (with a 3×3 circular kernel) is performed in order to fill out small gaps and is followed by an opening (using the same kernel) in order to remove single pixels and small pixel groups.

Remaining pixel clusters smaller than the minimum detection requirement (25 m^2 = 90 pixels in our case) are removed from the dataset. Finally shadow covered pixels are removed using an intensity threshold. These pixels are removed because object detection/registration is difficult and in many cases impossible in shaded areas.

Additional tests could be applied on a pixel or areal basis: check against other object classes in the existing database, evaluation of size and shape, comparison between solar angle and nearby shadows, comparison with edges detected in the original image, etc. etc.

The result of the post processing is an image having a value of 1 for all pixels where change has been registered, and 0 everywhere else. This is the final change detection product.

4 EXAMPLE

Figures 3 and 4 shows an example of the entire processing chain from aerial photo to change detection map in CIR and RGB data, respectively. The CIR image (figure 3) is 536 rows by 530 columns and includes 73 buildings. The RGB image (figure 4) is 555 rows by 555 columns and includes 61 buildings.

The test area is a traditional Danish suburban residential area. All buildings are single family houses with gabled roofs, and their sizes and roof colours vary widely. Small gardens with trees, grass etc. surround the buildings.

The two figures 3 and 4 have identical structures, which is as follows: The *upper left panel* shows the aerial photo. The *upper center panel* shows the result of the data fusion step. The buildings masked in white are registered in the map data base while the five (RGB) or six (CIR) buildings masked in red are deliberately removed from the map data base to test the change detection algorithm. The *upper right panel* is the result of the comparison between the building mask and the initial classification result. White pixels mean no change, grey pixels mean potential change, and black pixels mean evident change. The *lower left panel* shows the evident change areas (no change and potential change pixels are removed). The *lower center panel* shows the effect of spurious pixel removal, and the filling in of gaps while the *lower right panel* shows the remaining areas of change after removal of all clusters smaller than the change detection limit (25 m^2), and shadow covered pixels.

Results: The six change detection targets are all captured in the CIR case, while only three (and a tiny bit of the fourth) are captured in the RGB case.

The RGB scene has large amounts of misclassified vegetation all over the area. As expected, the CIR scene does not show any signs of misclassified vegetation.

In this case, the RGB scene has a large number of road pixels misclassified as buildings. The CIR scene eliminates almost all road pixels. Experience from other areas, however, indicate that in presence of asphalt roofs, the discrimination between roads and buildings become almost impossible using RGB and/or CIR data alone.

The remaining false alarms in the CIR image are almost all due to shadow areas which have slipped through the shadow elimination process.

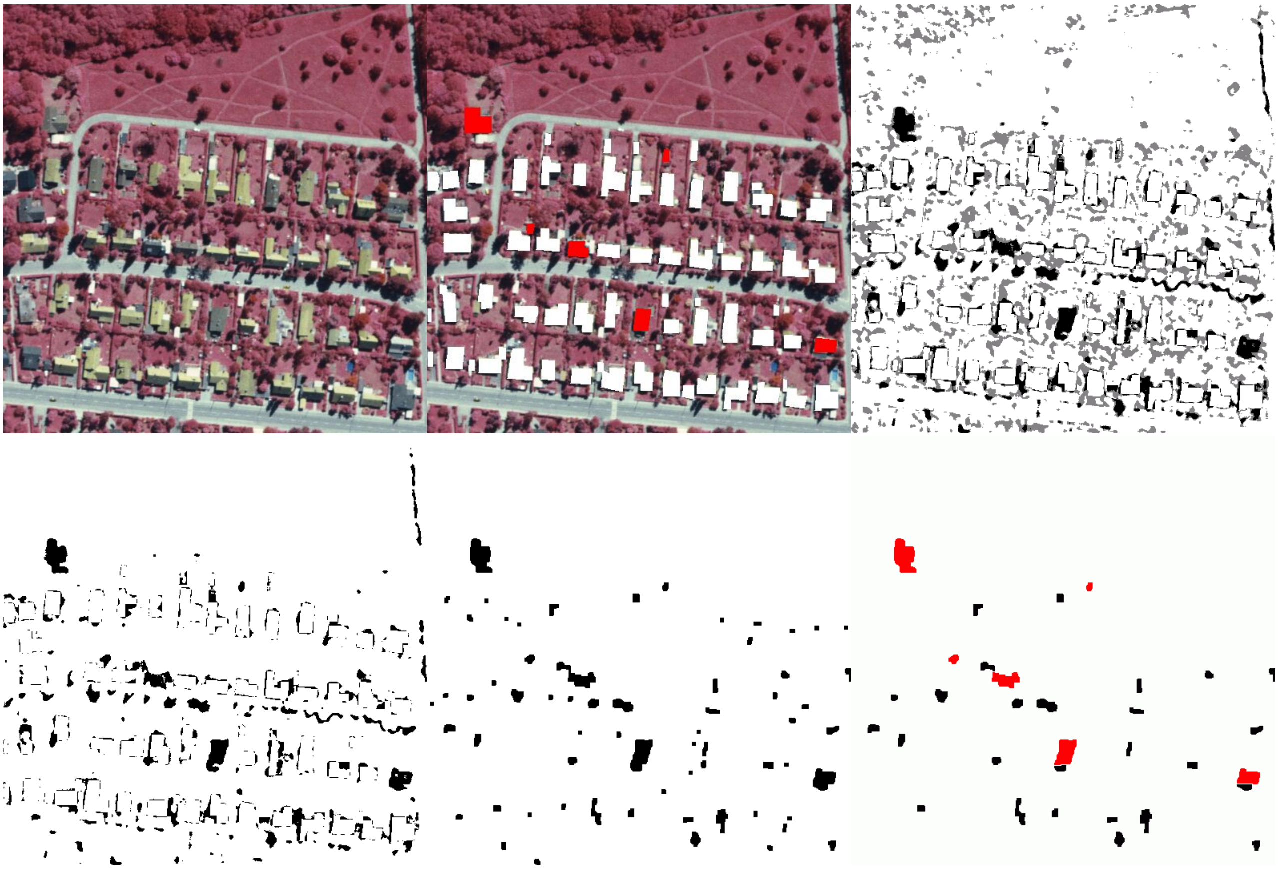

Figure 3: Steps of the change detection algorithm in CIR—cf. section 4 for description

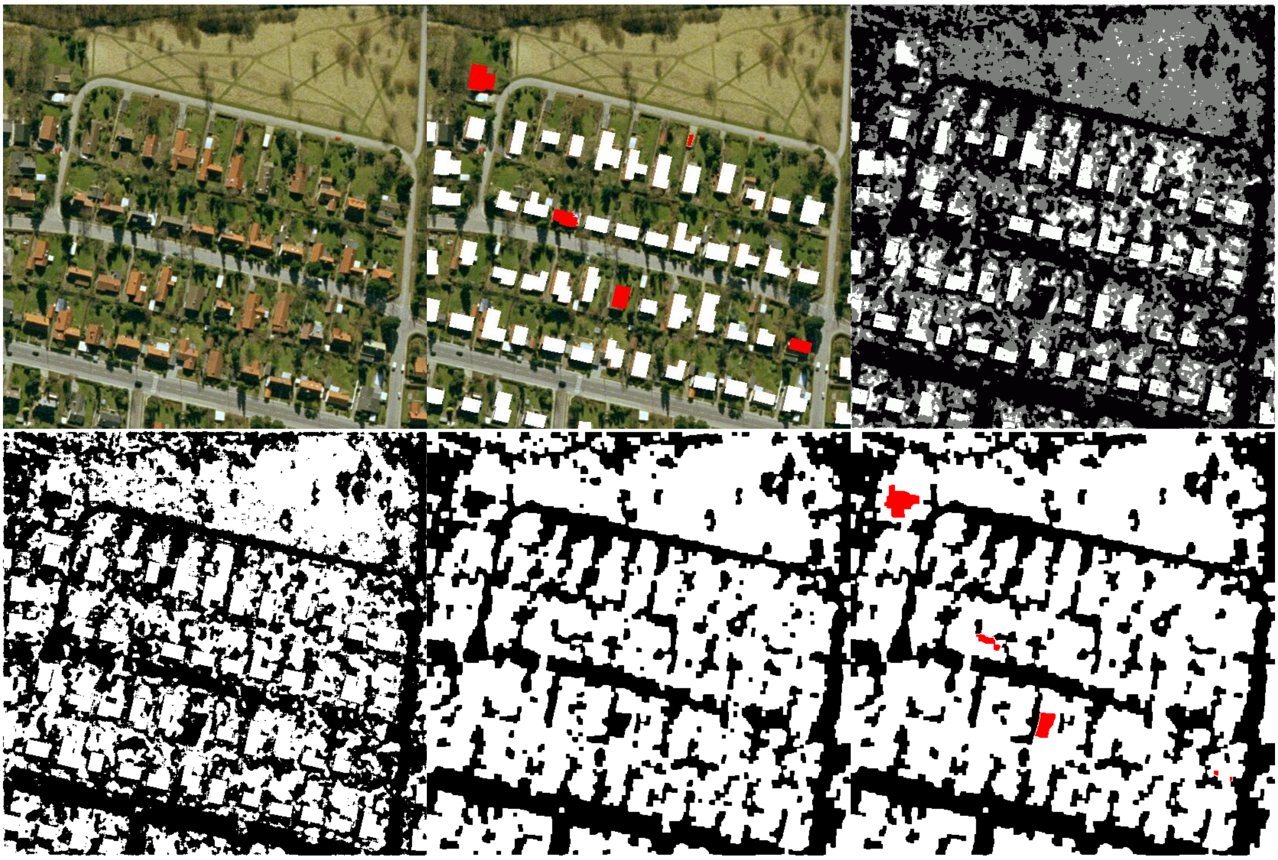

Figure 4: Steps of the change detection algorithm in RGB—cf. section 4 for description

5 DISCUSSION AND CONCLUSION

We have presented a method for change detection in urban areas, primarily aimed at the production of change detection maps for the building theme in digital map databases.

In the test case presented, the method works well in the CIR case, but produces an unacceptable number of false alarms in the RGB case. Experience from other test areas have shown that in some cases even the CIR case produces an unacceptable number of false alarms (although it is much smaller than for the RGB case presented here). Especially in areas with flat asphalt roof buildings, the CIR case have big difficulties in discerning buildings from roads (which is also true for the RGB case). The performance may be improved by using a classification introducing roads as a specific class in the classification step, or by introducing additional data, such as DSM, DEM, texture measures, object size considerations, shadow size metrics, and/or linear features.

CIR observations results in fewer false alarms than true colour observations. The two information sources do, however, complement each other so a full four channel sensor system (as provided by the latest generation of high-resolution multispectral satellites), or systems with an even higher number of bands would be very useful for this kind of work. But the spatial resolution of satellite data is still not quite high enough for the final 3D registration, and in our areas of interest, their cost dwarfs that of aerial photos by more than an order of magnitude.

The shadow detection step still leaves room for improvement; either through use of more advanced spectral methods, or by physical modelling of solar angle in combination with the existing 3D building registrations. In general, the postprocessing step is open ended, and allows for the attachment of any number of additional checks and tests. In our case, it would be obvious to use the abundance of ancillary data registered in the TOP10DK database.

Despite the number of false alarms, the overall conclusion is that the method works in our CIR case, as it tends to detect all deliberately introduced "missing" registrations wholly or partially, while still keeping the number of false alarms at a reasonable level.

REFERENCES

Arbeitsgemeinschaft der Vermessungsvervaltungen der Länder der Bundesrepublik Deutschland, 1988. Amtliches topographisch-kartographisches informationssystem (ATKIS). http://www.adv-online.de/english/products/atkis.htm.

Ball, G. H. and Hall, D. J., 1965. A Novel Method of Data Analysis and Pattern Classification. Stanford Research Institute, Menlo Park, CA.

Eidenbenz, C., Kaeser, C. and Baltsavias, E., 2000. ATOMI—automated reconstruction of topographic objects from aerial images using vectorized map information. International Archives of Photogrammetry and Remote Sensing XXIII, pp. 8.

Hoffmann, A., van der Vegt, J. W. and Lehmann, F., 2000. Towards automated map updating: is it feasible with new digital dataaquisition and processing techniques? International Archives of Photogrammetry and Remote Sensing XXXIII(B2), pp. 295–302.

Huang, K.-Y., 2002. A synergistic automatic clustering technique (SYNERACT) for multispectral image analysis. Photogrammetric Engineering & Remote Sensing 68(1), pp. 33–40.

Knudsen, T. (ed.), 2000. Proceedings of the seminar on remote sensing and image analysis techniques for revision of topographic databases. National Survey and Cadastre—Denmark, Technical Report no. 13, National Survey and Cadastre—Denmark, Copenhagen, Denmark.

Knudsen, T., Muhammed, H. H. and Olsen, B. P., 2002. A comparison of neuro-fuzzy and traditional image segmentation methods for automated detection of buildings in aerial photos. This Volume.

Kort & Matrikelstyrelsen, 2001. TOP10DK geometrisk registrering, specifikation udgave 3.2.0. Copenhagen, Denmark: Kort & Matrikelstyrelsen.

Kraus, K., 1993. Photogrammetry. Vol. 1, Ferd Dümmlers Verlag, Bonn.

Neukum, G., 1999. The airborne HRSC-A: Performance results and application potential. In: Photogrammetrische Woche, 1999, pp. 83–88.

Niederöst, M., 2000. Reliable reconstruction of buildings for digital map revision. International Archives of Photogrammetry and Remote Sensing XXIII, pp. 8.

Niederöst, M., 2001. Automated update of building information in maps using medium-scale imagery (1:15,000). In: E. P. Baltsavias, A. Grün and L. V. Gool (eds), Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), A. A. Balkema, Rotterdam.

Petzold, B., 2000. Revision of topographic databases by satellite images—experiences and expectations. In: (Knudsen, 2000), pp. 15–23.

Petzold, B. and Walter, V., 1999. Revision of topographic databases by satellite images. In: M. Schroeder, K. Jacobsen, G. Konecny and C. Heipke (eds), Sensors and mapping from space 1999, ISPRS, Hanover, Germany, p. 9.

Richards, J. A. and Jia, X., 1999. Remote Sensing Digital Image Analysis: an introduction. 3 edn, Springer, Berlin.

van Asperen, P., 1996. Digital updates at the Dutch Topographic Service. International Archives of Photogrammetry and Remote Sensing 31(B4), pp. 891–900.

Walter, V., 2000a. Automatic change detection in GIS databases based on classification of multispectral data. International Archives of Photogrammetry and Remote Sensing XXXIII(B4), pp. 1138–1145.

Walter, V., 2000b. The design and implementation of an automated system for revision of topographic databases. In: (Knudsen, 2000), pp. 7–14.

Walter, V. and Fritsch, D., 1998. Automatic verification of GIS data using high resolution multispectral data. International Archives of Photogrammetry and Remote Sensing XXXII(3), pp. 74–79.

Walter, V. and Fritsch, D., 2000. Automated revision of GIS databases. In: K.-J. Li, K. Makki, N. Pissinou and S. Ravada (eds), ACM-GIS 2000, Proceedings of the Eighth ACM Symposium on Advances in Geographic Information Systems, November 10-11, 2000, Washington D.C., USA, ACM, pp. 129–134.