Urban Digital Map Updating From Satellite High Resolution Images Using GIS Data as A Priori Knowledge

Timothée Bailloeul, Jinghui Duan, Véronique Prinet, Bruno Serra

***Abstract*—In this paper, we propose two methods aiming at updating Geographic Information System (GIS) urban maps using satellite high-resolution remote sensed images. Both methods are based on the input of *a priori* knowledge provided by GIS data, and a Digital Surface Model (DSM). This article introduces theoretical aspects of our methods; implementation is currently ongoing.**

***Index Terms*—A priori knowledge, deformable models, digital map updating, classification, Geographic Information System, urban change detection.**

I. INTRODUCTION

Computerized map updating is a challenging task that has been tackled for more than twenty years. Indeed, manual map revision is a time- and cost-consuming process that makes automatic or semi-automatic updating attractive. The need of up-to-date urban maps is worldwide, however it may be even more desirable in countries where the urban development is higher. Beijing, capital city of China is a site of special interest for urban change monitoring and map updating because of its rapid urban growth. The main reasons are the will to populate outskirts of Beijing city, the preparation of the Olympic Games in 2008 and the development of a sustainable environment. The objective of this work is to update outdated digital urban maps using recent information from high-resolution satellite imagery and prior information provided by maps. In other words, our task is to detect and analyse changes with localization (spatial detection), and identification (semantic interpretation). Two approaches are proposed for this purpose: the first makes use of a contour-based segmentation approach; the second one focuses on a region-based segmentation. Both approaches are based on the use of *a priori* knowledge derived from outdated GIS data that constrains the segmentation process. The use of existing maps enables to gain specific information from the satellite imagery. It increases the confidence in object extraction processes compared to regular "bottom-up" approaches, which fail in dense urban environment sensed with a high resolution. Both methods will be tested on images and GIS data covering the north part of Beijing city. The rest of the paper is organized as follows: section II is a brief literature review on digital map updating using a priori knowledge and data fusion. In section III we introduce our methods and conclude in section IV by limitations and open problems of the proposed schemes.

Manuscript received March 1, 2003. This work was supported in part by Alcatel Space Industry, the High-Tech program of the Chinese Ministry of Science and Technology, and the Sino-French joint Laboratory in Computer Science, Control and Applied Mathematics (LIAMA).

T. Bailloeul is with Alcatel Space Industry, 100 boulevard du Midi, BP 99, 06156 Cannes La Bocca Cedex, France and with the Sino-French joint Laboratory in Computer Science, Control and Applied Mathematics (LIAMA), Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, China (phone: +86 82614490; Fax: +86 62647458; e-mail: tbailloeul@liama.ia.ac.cn).

J. Duan and V. Prinet are with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, China (e-mail: prinet@nlpr.ia.ac.cn; jhduan@nlpr.ia.ac.cn).

B. Serra is with Alcatel Space Industry, 100 boulevard du Midi, BP 99, 06156 Cannes La Bocca Cedex, France (e-mail: Bruno.Serra@space.alcatel.fr).

I. BACKGROUND

With the improvement of the resolution of aerial and satellite remote sensed images (Ikonos, Quickbird), the semantic richness of the image increases and makes image analysis more difficult. Dense urban environment sensed by high-resolution optical sensors is even more challenging: occlusions and shadows due to buildings hide some objects of the scene, hindering exhaustive automatic or manual extraction. For more than twenty years, researches in image processing and computer vision have tackled the issue of map revision. As pointed out by the authors of [1] and [3], most of the approaches have been "bottom-up", providing good results on peculiar low-resolution scenes. Most of those methods are tested on simple scenes such as low scale images, rural or peri-urban sites where objects of the sensed scenes are quite visible, with less shade or occlusion artefacts than in inner cities. Those standard methods for object recognition fail in dense urban environment and require the input of additional information. Additional information can take on several forms:

A. Additional sensed data

Additional sensed data could be colour or near infrared spectral channels or lidar imagery. In [5], a lidar Digital Surface Model (DSM) is considered as an additional colour channel of multispectral and Near Infrared (NIR) aerial images. The results of a supervised classification are then enhanced, especially towards differentiation between grass-covered land,

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 11,2023 at 04:23:51 UTC from IEEE Xplore. Restrictions apply.

trees, streets and buildings.

A. A priori knowledge

A priori knowledge can be roughly described as the knowledge derived from the programmer's understanding of a scene or/and the information provided by existing maps. The pre-cited "urban artefacts" (occlusion, shadows) can illustrate the first aspect of a priori knowledge, e.g. shade only occurs close to buildings. This knowledge is general, could be applied extensively to remote sensed scenes, but is subjective. Conversely, prior knowledge provided by existing digital maps is specific to a given site. It can yield useful semantic and spatial information when combined with satellite or aerial imagery, and is fully objective. Both aspects of prior knowledge may ease and increase the confidence in object extraction in urban environment; however existing maps have seldom been used, at least back to the mid-nineties ([1]-[6], [8], [9], [11]-[13]). In [3], a priori knowledge is used for mixed "top-down/bottom-up" road detection. Road models, existing maps, contextual reasoning and object-oriented features extraction are inputs of the process. In [9] outdated GIS data initiates raster 1:25,000 maps and images interpretation through semantic nets. The GIS is iteratively compared to the semantic nets representations of the raster data by a "top-down/bottom-up" method for hypothesis generation. In [12] GIS-guided supervised classification is carried out on aerial images of rural sites. Since supervised classification needs the input of an operator to digitise learning areas, fully automatic recognition is impossible. The authors propose to use outdated GIS data as training areas for later maximum likelihood classification on a recent image. Matching of areas and lines between map and classification output is achieved to localize and identify changes. In [6], GIS data is used to support automatic extraction of vegetation from IRS-1C satellite images. Spectral signatures of vegetation are derived from supervised classification carried out on colour channels using thematic maps as training areas. Resulting extracted objects boundaries are later refined by a panchromatic image.

More recently the ATOMI project [4] aims at roads and buildings map updating and takes advantage from existing knowledge to restrict search spaces, and ease object extraction. The authors also use redundant sources of information (1:25,000 scale GIS data, DSM, colour images) for mutual completion, in order to cope with the lack of information induced by high-resolution artefacts, such as shade or occlusions. Two different approaches are carried out for road extraction [13] and building reconstruction [8]. In [8], existing vectorized residential maps are used as a rough guess towards the location and the shape of buildings to be extracted from a colour orthoimage. Besides, approximation of buildings is computed by two other ways: rough shaped superstructures are extracted from a normalized DSM (blobs) and by unsupervised classification using several colour space projections. The subsequent rough data is matched by affine transformation to the objects contained in an orthoimage for planar accuracy improvement and change detection. The updating quality is rated to express the reliability of extracted objects. In [13], a priori knowledge is composed of rules towards road design, cues such as 3D straight lines or road marks, and outdated GIS. It enables building up a knowledge database for each GIS segment according to its class. Road class oriented processing later refines this database: cues are extracted and combined using image processing techniques dedicated to a given class, the results characterize the database with information specific to the image. Roads are finally extracted using constraints between cues, road areas detected by unsupervised classification, existing maps, and a normalized DSM. The authors of [1] use existing GIS data, aerial imagery and *snakes* active contours ([7]) in order to update and revise road digital maps. The map accuracy is first quantified by the input of an image acquired at the same time as the GIS data: *snakes* initialised on the GIS road objects move to the actual road track of the image. According to the *snake* motion, and for each of its node, an accuracy score is computed using fuzzy logic. This last score is the input for an additional energy term part of the total energy functional of the active contour. This energy will constrain the motion of the *snake* (initialised on the GIS data) within a recent image. The constraint is comparable to a spring connected to a given node and fixed to its initial position. The stiffness and influence area of each spring depends on the previously computed accuracy score: the higher the accuracy, the higher the stiffness, and the smaller the influence area. The final segmented road revises the map from erroneous digitisation and updates it from changes. This method tested in rural or peri-urban environment only handles road deviation but not the appearance of new objects. In [11], road extraction is achieved with 1-meter resolution Ikonos images of dense urban environment. An operator roughly draws roads on the fly in the image. This spatial and thematic prior information towards roads could also be provided by existing maps, however it has the property to be already up-to-date. This last rough road network is then converted into a graph that is used to initialise *snakes*. These active contours finally segment the two sides of the roads in the satellite image. A multi-resolution approach gets rid of urban artefacts such as cars. The authors of [2] also make use of *snakes* for road extraction in SAR images. Since setting up *snakes* parameters is a problem for automatic processing, the nature of the GIS object will help for such a task: the curvature of the initial *snake* position is computed according to the outdated road map in order to determine the rigidity coefficient. Working with *snakes* assumes that objects to be detected are close to these provided by prior outdated information. This considerably restricts the nature of changes between maps and recent images, and prevents extraction of new objects. Most of these methods are not using high-scale existing maps and thus do not totally benefit from the input of the available spatial and thematic information. Maps are usually considered as an approximation of the objects to be detected and enable to narrow their search space. Our approaches are quite different since we consider GIS data as an accurate source of spatial information that allows change detection and precise training areas when combined with satellite imagery.

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 11,2023 at 04:23:51 UTC from IEEE Xplore. Restrictions apply.

I. Proposed Methodologies

A. Input Data Sets

Multi-temporal acquisitions of Quickbird-2 multispectral images over the north part of Beijing city will be used. These images cover two main spots of interest: the future Olympic Village and the new "Silicon Valley". Satellite images have a 2.8 m resolution with a 14 m root mean square error accuracy; four spectral channels are available: regular red, green, blue (RGB) and NIR bands. Radiometric and geometric corrections were applied. Geometric corrections only deal with terrain rectification; superstructures such as buildings are not ortho-rectified. The first satellite image acquisition was in March 2002.

Fig. 1. Tshinghua University area, Beijing, China: GIS road layer displayed in white (1996) registered by global affine transformation on a Quickbird satellite image (2002). The white arrow points some changes.

Besides, a DSM will be computed using aerial RGB stereo-pair images. Aerial imagery is 1:10,000 scale, with 15 cm focal length and scanned with 21 microns resulting in 21 cm footprint images. The need of a DSM is twofold: first it is required to ortho-rectify the satellite image from perspective distortion that affects its accuracy, second it is an additional source of information that may help the updating process in a complex and dense urban environment. The DSM will be computed by an edge preserving stereo correlation technique after having been projected into epipolar geometry [10]. Outdated urban digital maps are GIS data from 1996 and 2001 with a 1:10,000 scale. This vectorized data covers the Quickbird images area and contains the layers we will focus on: roads, residential areas, green spaces, rivers and stretches of water. Buildings are modelled by polygons fitting the footprint; lakes and green spaces are delineated by the same kind of polygon. Open polygonal lines fit the sides of rivers and roads. GIS data does not provide any 3D information and has a 5 m geometric accuracy. Fig. 1 shows a GIS road layer from 1996 overlaid on a 2002 Quickbird image. Many changes are noticeable, especially concerning the road network. The white arrow points out a network of narrow streets replaced by a broad avenue.

B. First Method: Use of GIS Data and One Image

The main issue of this approach is to detect changes between outdated GIS data and a recent satellite image using *a priori* knowledge and deformable models. Assuming that over a one-year period a few percentage of a urban environment has changed, we take advantage of *a priori* knowledge provided by existing GIS data to ease change detection: unchanged areas will be used for model refinement -- the models describe each "object", that is building, lake, road, etc – that will enable pattern recognition on areas of the image subject to changes. Additional knowledge such as height or spectral information derived respectively from the DSM and satellite imagery will help to disambiguate the dense urban scene complexity. Buildings of the satellite image should be ortho-rectified in order to improve the final updating accuracy of the GIS data. The method illustrated in fig. 2 can be described through the following steps:

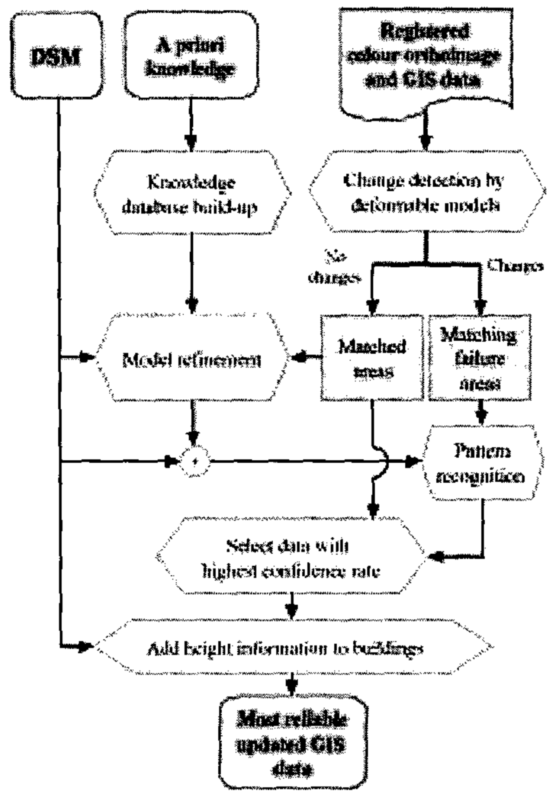

Fig. 2. Diagram synthesizing the first method for GIS data updating.

1) Global Registration: GIS data is overlaid on the satellite image by affine transformation based on an automatic method or manually detected control points (cf. Fig. 1). Unchanged areas between the existing digital map and the image should fit. First registration tries showed a good matching towards objects located on the ground (roads, lakes, rivers).

2) Knowledge database build-up: a priori knowledge provided by existing outdated GIS data (or layer-maps) is used to define "object" models. The object classes are given by the layers of interest of the GIS data (roads, buildings...). Models are defined by the spatial and geometric information gained

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 11,2023 at 04:23:51 UTC from IEEE Xplore. Restrictions apply.

from the GIS data. General rules towards the objects composing the urban scene are also input to the knowledge database (e.g. buildings and roads do not overlap). At this step, the modelling is quite coarse. It will be later refined using some information specific to the image. This approach is similar to the work achieved in [13] towards road extraction.

3) Change detection: deformable models initialised on the GIS objects are used to fit the corresponding objects of the image. If the matching succeeds for a given object, we can assume that no change occurred at that location. On the contrary, mismatching indicates change. Geometric constraints of the initialisation enable to set up automatically the parameters of the models. Deformable models will be also constrained by previously defined rules and additional data (e.g. a deformable model initialised on a road side can not fit an object above the ground). Quantification of change detection reliability will be needed. At the opposite of the works achieved with *snakes* active contour ([1], [2], [11]), we do not expect deformable models to extract objects from the scene since this task is too complex and cannot handle new objects detection. Deformable models are constrained to change detection only.

4) Model refinement: unchanged areas will be used to improve object modelling by gaining radiometric, texture and height information from the image and the DSM.

5) New objects recognition: regions previously detected as "changed areas" are subjected to classification using the refined models. The output of this task is a semantic interpretation and localization of the changes: new objects are identified and spatially localized. Newly detected objects are vectorized to have a consistent representation with the GIS data. The confidence in new detected objects will also have to be evaluated.

6) Map updating: newly detected and unchanged GIS objects with the highest confidence rate are merged to make the final updated GIS data. Buildings height derived from DSM is added.

C. *Second Method: Use of GIS Data and Two Images*

The input data of the second method should fulfil the following requirements: two satellite images represent a same scene acquired at a different time. The content of one satellite image is consistent with the GIS data that means no changes occurred between this image acquisition and the GIS data. Satellite imagery has to be ortho-rectified to enable accurate later registration with the GIS data.

Since the content of one image is the same as the GIS data, we expect to discover changes between digital maps and the recent image through the changes between the two images. Prior spatial and thematic knowledge provided by the existing digital maps will provide training areas for subsequent supervised learning ([6], [12]). Besides, image-to-image change detection may divide the recent image into two regions: changed and unchanged areas. Then, changed areas of this image can be subject to classification with spectral, radiometric, or other statistical information gained from the former image. The output of that process is the extraction of new objects of the urban scene. At last, GIS data is updated according to classification results. The method can be summarized through the four following steps and fig. 3.

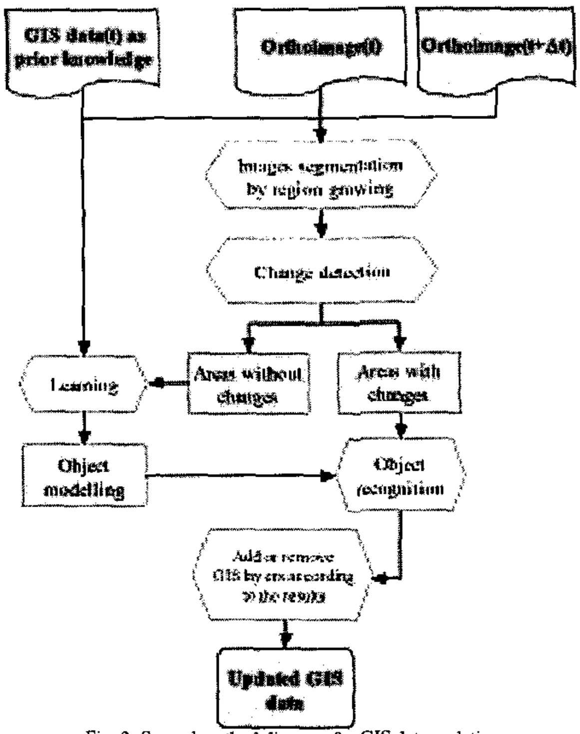

Fig. 3. Second method diagram for GIS data updating

1) Registration: GIS data and the former ortho-rectified image are overlaid by an affine transformation.

2) Change detection: Segmentation by region growing is performed on both images for change detection. Each GIS object allows to determine a seed that initiates a region growing process at the same location of the sensed scene in the two images. A comparison of the two segmented regions in both images enables to state whether a change occurred or not. Moreover, the output of this process provides a segmented region in the former image for each GIS object. This may overcome erroneous registration and map digitisation that prevent from perfect matching between the former image and the GIS data. Change detection based on region growing should yield: (i) areas in both images where no changes occurred . The corresponding areas in the former image enhance the spatial localization of registered GIS data. (ii) areas in the recent image where changes occurred.

3) GIS-guided supervised learning and object recognition: regions without changes of the former image define areas for supervised learning. Therefore, layers of interest of the GIS data can be characterized by some radiometric, statistical, or geometric properties of the image. Objects can be recognized in areas of change belonging to the recent image according to

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 11,2023 at 04:23:51 UTC from IEEE Xplore. Restrictions apply.

the knowledge gained by this learning. New objects are then extracted by classification.

4) *GIS data updating:* objects extracted from changed areas are vectorized and inserted to the GIS database.

II. Conclusion

We have introduced two methods aiming at GIS map updating in an urban environment. The two approaches are based on image-GIS data fusion to make the urban map updating process more reliable. Compared to other methods using existing maps as prior information, our methods do not use GIS data as a rough approximation of objects that may be extracted in the image. We make use of high-resolution digital maps that can outline objects of the image. As a result, we take advantage of high quantity and quality of learning areas for later classification. A theoretical limitation of the first methodology is its intrinsic contour-based change detection. Therefore an object in the image that is semantically different from its GIS representation but which contours did not change, has no chance to be detected as different. This problem may only arise for object located on the ground since the DSM can disambiguate this critical situation. A limitation of both proposed schemes is the sensitivity towards GIS to image registration. Indeed, a mistaken registration may cause wrong change detection and inaccurate supervised classification for both approaches. Further works will be dedicated to reduce the dependence towards the registration quality. Confidence quantification of change detection and object extraction will be carried out in the future implementation of the methods.

Acknowledgment

The authors acknowledge the Beijing Institute of Survey and Mapping (BISM) for providing GIS data over Beijing city and intensive discussions on output results requirements.

References

[1] P. Agouris, A. Stefanidis, S. Gyftakis: "Differential snakes for change detection in road segments", *Photogrammetric Engineering and Remote Sensing*, Vol. 67, No. 12, pp. 1391-1399, 2001.

[2] L. Bentabet, S. Jodouin, D. Ziou and J. Vaillancourt, "Automated updating of road databases from SAR imagery; integration of road databases and SAR imagery information", *International Conference on Information Fusion*, Montréal, 2001.

[3] M. De Gunst, "Knowledge based interpretation of aerial images for updating of road maps", *PhD thesis, Technical University of Delft*, 1996.

[4] C. Eidenbenz, C. Kaeser, E. Baltsavias, "ATOMI - Automated reconstruction of topographic objects from aerial images using vectorized map information", *IAPRS*, Vol.XXIII, Amsterdam 2000.

[5] N. Haala, V. Walter, "Automatic classification of urban environments for database revision using lidar and color aerial imagery", *IAPRS*, Vol. 32, Band 7-4-3W6, Valladolid, pp. 76–82, 1999.

[6] C. Heipke, B-M. Straub, "Automatic extraction of GIS objects from digital imagery", *IAPRS*, Vol. 32, Part 3-2W5, pp 167-174, München, 1999.

[7] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: active contour models", *First International Conference on Computer Vision*, pp. 259-268, 1987.

[8] M. Niederoest, "Reliable reconstruction of buildings for digital map revision", *IAPRS*, Vol. XXIII, Amsterdam, 2000.

[9] K. Pakzad, H. Koch, R. Tönjes, "Knowledge based interpretation of aerial images and maps using digital landscape model as partial interpretation", *Workshop on "Semantic Modeling of Topographic Information from Images and Maps" (SMATI'97)*, Bonn, Germany, 1997.

[10] N. Paparoditis, M. Cord, M. Jordan and J-P. Cocquerez, "Building detection and reconstruction from mid- and high-resolution aerial imagery", *Computer Vision and Image Understanding*, Vol. 72, Issue 2, pp. 122-142, 1998.

[11] R. Péteri, T. Ranchin, "Extraction and update of street networks in urban areas from high-resolution satellite images", *ISPRS Proceedings*, Ottawa, Canada, 2002.

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

[13] C. Zhang, E. Baltsavias, A. Gruen, "Knowledge-based image analysis for 3D road reconstruction", *Asian Journal of Geoinformatics*, Vol. 1, No. 4, 2001.

Authorized licensed use limited to: Institut Teknologi Bandung. Downloaded on October 11,2023 at 04:23:51 UTC from IEEE Xplore. Restrictions apply.