Photogrammetrie • Fernerkundung • Geoinformation 6/2005, S. 511–522, 8 figs., 1 tab.

Digital Building Map Refinement from Knowledge-driven active Contours and very High Resolution Optical Imagery*

Timothée Bailloeul, Véronique Prinet, Bruno Serra, Philippe Marthon, Pinxiang Chen & Haitao Zhang

Keywords: photogrammetry, satellite optical imagery, Quickbird, building, accuracy, active contours, GIS, DSM, fusion.

Abstract: We propose a novel approach for digital building map refinement based on the use of knowledge-driven active contours and very high resolution panchromatic optical imagery. This methodology is designed to finely match each building symbolized in an urban Geographical Information System (GIS) database onto its counterpart representation in remote sensing data. This method is GIS map-driven: GIS data globally registered to the image allows to initialize an active contour near the target building in the image to achieve subsequent refinement. Moreover the digital map provides valuable shape information about the object in the image we aim at matching. This geometric and specific prior knowledge is embedded as a shape constraint into the active contour and enables to overcome urban artifacts issues. Besides, we propose to embed a coarse Digital Surface Model (DSM) as well as a spatio-temporal shape prior constraint within the active contour model. Experimental results carried out over Beijing city area and illustrated in this paper show how these latter contributions improve the robustness and speed of the map refinement process. Map refinement addressed in this paper is becoming an essential issue for urban planning, telecommunications, automobile navigation, crisis and pollution management, which all rely on up-to-date and precise digital maps of a city.

Zusammenfassung: *Verfeinerung digitaler Gebäudedaten aus hoch aufgelösten optischen Satellitendaten durch wissensbasiert gesteuerte aktive Konturen.* In diesem Beitrag wird ein neuer Ansatz zur Verfeinerung digitaler Gebäudedaten aus hoch aufgelösten panchromatischen Satellitendaten mithilfe wissensbasiert gesteuerter aktiver Konturen vorgeschlagen. Die Methode ermöglicht eine genaue Zuordnung eines Gebäudes einer GIS Datenbank zu der entsprechenden Repräsentation im Fernerkundungsbild. Die Methode geht von den GIS-Daten aus: global zu den Bildern zugeordnete GIS-Daten erlauben die Initialisierung einer aktiven Kontur in der Nähe des korrespondierenden Gebäudes als Startwert für die Verfeinerung. Daneben bieten die GIS-Daten wertvolle Forminformation für die zuzuordnenden Objekte. Diese Vorinformation wird als zusätzliche Bedingung für die aktive Kontur formuliert und ermöglicht es, Probleme zu lösen, die speziell in städtischen Gebieten entstehen. Daneben wird vorgeschlagen, auch Bedingungen für ein DGM und eine raumzeitliche Komponente in das Konturmodell aufzunehmen. Tests mit Bildern aus Beijing zeigen, wie die zuletzt genannten Verbesserungen die Zuverlässigkeit und die Geschwindigkeit der Gebäudedatenverfeinerung verbessern. Die in diesem Beitrag beschriebene GIS-Datenverfeinerung stellt sich zunehmend als eine wichtige Thematik für Stadtplanung, Telekommunikation, Fahrzeugnavigation, Management von Krisen und Luftverschmutzung heraus, die allesamt auf aktuelle und geometrisch genaue Daten angewiesen sind.

* Enhanced version of a paper published in the proceedings of the ISPRS Hannover Workshop 2005 "High Resolution Earth Imaging for Geospatial Information", May 17–20, 2005, Institute of Photogrammetry and GeoInformation, University of Hannover.

1432-8364/05/2005/0511 $ 3.00

1 Introduction

1.1 *Context and focus*

The era of sub-meter resolution satellite imagery presents new opportunities for users of spatial data. Indeed high resolution satellite imagery is becoming an affordable solution to add large-scale and high level of geographic knowledge and detail to geospatial databases. The more regular revisit capabilities of satellites also enable a higher frequency of map revision and monitoring. However the maintenance of such Geographical Information System (GIS) data is time and cost consuming when achieved manually. Efforts have been undertaken for more than thirty years by the Computer Vision and Image Processing communities to assist and automate the photogrammetric processing chain in order to shorten revision cycles and therefore improve the currency of information. Image interpretation from very high resolution images raises difficulties and challenges that do not appear with low and mid-resolution data: profusion of details makes automatic analysis of images arduous, and causes traditional bottom-up approaches to fail. This is particularly critical for urban environments where shadows, occlusions and apparent perspective distortion of high buildings are common *artifacts* to cope with. In order to ensure a reliable automatic image understanding of dense urban environments, a recent trend is to use multiple sources of information, which complementarity may disambiguate the analysis. Multiple sources of information can embody collateral imagery data of the same scene or prior knowledge towards the target object to be extracted from the data, the input data to be used as well as the processing methods to be applied (Baltsavias 2002). Map revision comprehends three main aspects. The first one deals with the detection of new objects to be incorporated into the map from more recent imagery data. The second applies to the issue of removing from the map any object that is no longer present in imagery data. The last one focuses on improving the spatial quality of the map from imagery as well as enhancing its level of information (such as adding 3D information to a 2D map). In this paper we address the latter aspect of map revision by proposing a novel method based on the use of active contours and very high resolution panchromatic optical remote sensing imagery in order to improve the spatial location of cartographic buildings objects included in a 2D digital map. This methodology is designed to automatically improve the accuracy of urban GIS databases while overcoming the difficulties of analyzing urban scenes sensed at a high resolution. We take advantage of the geometric prior knowledge derived from the map and adapt region-based and shape constrained active contours models exposed in (Paragios & Deriche 2002, Chan & Zhu) in order to accurately match each building symbolized in the map to its counterpart representation in the satellite image. Besides, we propose two approaches to increase the robustness of the active contours matching. The first deals with adding an exogenous source of information in the active contour model. In our application, additional data is a coarse orthoscopic Digital Surface Model (DSM) encoding the altitude of the same scene as the satellite image. The second approach consists in allowing a spatio-temporal change of the prior shape constraint during the active contours convergence, which may robustly accelerate the refinement process. In the next subsection we briefly outline former works using active contours for roads or buildings extraction from remote sensing data. In section 2, we review the prerequisite background towards knowledge-based active contours and how we adapt them to building map refinement. We detail the contributions of our scheme as well as its domain of application. In section 3, we present some results achieved with 1:10,000 scale cartographic data and Quickbird imagery over Beijing city area. We finally conclude in section 4 with a discussion about future improvements of the proposed scheme.

1.2 *Related works*

Recently active contours or deformable models have raised the interest of the Photogrammetric and Image Processing communities for the purpose of object extraction from remote sensing imagery (Agouris et al. 2001, Vinson et al. 2001, Guo & Yasuoka 2002, Péteri & Ranchin 2003, Rochery et al. 2003, Oriot 2003). Active contours are attractive since they are flexible, can be easily interfaced with the user in a semi-automatic fashion, and can readily embed high level information, which may be useful to ease the extraction process and to make it more reliable. High level information and *a priori* knowledge embrace multiple aspects which are comprehensively reviewed in (Baltsavias 2002). In this paper we focus on the incorporation of geometric prior knowledge within active contours based frameworks. Geometric prior knowledge has two aspects: it can be *generic* or *specific*. Generic information is derived from common sense knowledge and empirical learning. Statements like "buildings roof outlines often have ninety degrees angles" or "roads have parallel borders" are examples of generic knowledge and are already extensively used to enable object extraction. In (Péteri & Ranchin 2003) double snakes are used to extract both sides of roads from high resolution images of a dense urban environment. The snakes (Kass et al. 1987) are initialized from an existing road network graph which may be derived from a map or manually. The snakes evolve according to parallelism inner constraints as well as gradient-based external image forces in order to drive the active contour close to the road borders. In (Rochery et al. 2003) active contours derived from a variational approach are used for road extraction from mid-resolution images. The authors propose a novel quadratic energy to model non local interactions between contour points. This enables to incorporate generic knowledge towards road parallelism and the minimum width of the roads to be extracted. Unlike the previously cited method, this scheme is not sensitive to initialization and it naturally embeds roads geometrical properties and intrinsically incorporates the concept of network. The authors of (Agouris et al. 2001) use existing GIS data, aerial imagery and snakes active contours 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 initialized 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 of an additional energy which is part of the total energy functional of the active contour. This energy will constrain the motion of the snake in a more recent image of the same scene. The final segmented road revises the map from erroneous digitization and updates it from changes. In (Guo & Yasuoka 2002) an Ikonos image and a laser scanning DSM are jointly used for snake-based building extraction. The snake is initialized from a multiple height bins thresholding of the DSM and evolves according to edge information derived from the image and the DSM. In (Oriot 2003) some statistical snakes are used for building extraction from aerial images. They embed a correlation cost function from stereoscopic images, which favours the inclusion within the active contour of higher disparity measures than the background. Building extraction is simultaneously refined by edge information derived from the images and by generic shape constraint favouring ninety degrees corners. Initialization is achieved by human interaction while the optimization process is based on insertion/updating/deletion of vertices. Good results are demonstrated even if mistakes arise with vegetation closely surrounding buildings. In (Vinson et al. 2001) deformable templates are used to finely extract rectangular buildings from the output of an above-ground structures detection. Optimal rectangular model parameters are later found from the edge information derived from an orthoimage.

Generic geometric knowledge includes social and cultural aspects which augment its variability across geographical locations and therefore decrease the robustness of this information (roads widths and buildings shapes may vary at a regional/national level and even more at a worldwide scale). Unlike the aforementioned works we propose to make use of *specific* geometric information, which is derived from symbolized buildings contained in a digital 2D map. Specific shape information is highly discriminative, object and scene dependent and may enable better recognition and matching performances. This specific geometric prior information derived from the map will be embedded as a shape constraint within an active contours framework. Shape constrained active contours have been extensively studied since the early nineties, especially by the Medical Imaging community which has to deal with data corrupted by noise, occlusions or low contrast. Their use has been recently extended to natural scenes or manufactured objects images and object tracking from video sequences (Chen et al. 2001, Rousson 2002, Chan & Zhu 2003, Cremers et al. 2004). However, the already proposed shape constrained models are not robust enough to deal with complex dense urban scenes. The next section describes how prior shape knowledge has been incorporated within region-based active contours as well as our contributions to increase their robustness.

2 Methodology

2.1 *Active contour model*

We propose to adapt knowledge-driven active contours to digital building map refinement from very high resolution optical satellite imagery. Our goal is to finely match each building symbolized in a map to its counterpart representation in a panchromatic high resolution image. This image is assumed to be the ground truth and has higher geocoding accuracy than the map. Cartographic objects are initially and coarsely registered to the satellite image (this could be the result of rough registration process). Their accuracy is later improved by our proposed fine matching technique. The information provided by the map is used to initialize active contours: initial location is provided by the global map-to-image registration, and the initial shape is similar to the considered cartographic object. Since region based active contours are known to be less sensitive to initialization than their gradient-based counterparts, we make use of the region-based formulation of the Bayesian MAP (Maximum a Posteriori) deformable model formerly proposed in (Paragios & Deriche 2002) in order to drive the active contour to the target building in the image. This approach best befits segmentation of piecewise smooth components of an image. Since we deal with buildings that exibits shape singularities (such as corners) we choose to implicitly represent active contours by their level set functions which naturally model sharp corners (Osher & Sethian 1988). Derived from a variational approach, such active contours minimize the following energy functional:

$$
\begin{aligned}
J^s(\phi) = & \int \left\{ \frac{\left(I^s(\mathbf{x}) - \overline{I}^s_{in}(t)\right)^2}{2\,\sigma^s_{in}(t)^2} + \ln \sqrt{2\,\pi\,\sigma^s_{in}(t)^2} \right\} \\
& \cdot H\big(\phi(\mathbf{x})\big)\, d\mathbf{x} \\
& + \int \left\{ \frac{\left(I^s(\mathbf{x}) - \overline{I}^s_{out}(t)\right)^2}{2\,\sigma^s_{out}(t)^2} + \ln \sqrt{2\,\pi\,\sigma^s_{out}(t)^2} \right\} \\
& \cdot \big(1 - H\big(\phi(\mathbf{x})\big)\big)\, d\mathbf{x}
\end{aligned}
\qquad (1)
$$

where $\overline{I}^s$ and σ^{s^2} respectively denote the image mean and variance grey level. Subscripts *in* and *out* refer to the computation of these statistical quantities inside and outside the evolving active contour. The active contour is embedded in a level set function ϕ which is assumed to be positive inside the contour. The superscript s refers to the satellite image to be analyzed. H represents the Heaviside function. Edge-based and contour regularization terms have been deliberately omitted in equation (1) since we only investigate region-based active contours. Besides, the shape constraint introduced in the next step will act as a contour regularizer. Shape knowledge directly derived from the map is incorporated as a shape constraint in the active contour to make it akin the considered cartographic reference template. The gain of shape constraint is twofold: i) it enables to match the right building in the image according to shape information. ii) it overcomes common urban artifacts such as occlusions or low contrast of the target building. We propose to use the shape constraint energy proposed in (CHAN & ZHU 2003), which compares the area within the active contour and the reference template:

$$J_{shape}(\phi, \psi) = \int \left\{ H(\phi(\mathbf{x})) - H(\psi(\mathbf{x})) \right\}^2 d\mathbf{x} \tag{2}$$

ψ is the level set function embedding the prior shape. This term is made invariant from any similarity transformation: $\psi(\mathbf{x}) = \psi_0(T_{sim}(\mathbf{x}))$ where ψ_0 is the level set function embedding the static prior shape derived from the map. In addition to the active contour evolution process, invariance from similarity transformation requires an additional optimization scheme to estimate the best parameters (rotation, translation, scale), which minimize (2). Shape prior incorporated into region based active contours yields the functional J_{SC}:

$$J_{SC}(\phi, \psi) = J^s(\phi) + \lambda J_{shape}(\phi, \psi) \tag{3}$$

The constant weight λ balances the influence of the shape prior regarding to image information. To perform the gradient descent method and retrieve ϕ that minimizes equation (3), we need to calculate the derivative of J_{SC} with respect to ϕ, which yields the following evolution equation of the active contour:

$$\phi_t(\mathbf{x}, t) = \left\{ -\frac{\left(I^s(\mathbf{x}) - \overline{I}^s_{in}(t)\right)^2}{2\,\sigma^s_{in}(t)^2} + \frac{\left(I^s(\mathbf{x}) - \overline{I}^s_{out}(t)\right)^2}{2\,\sigma^s_{out}(t)^2} + \ln\left(\frac{\sigma^s_{out}(t)^2}{\sigma^s_{in}(t)^2}\right) + 2\,\lambda\left[H_a(\phi(\mathbf{x})) - H_a(\psi(\mathbf{x}))\right] \right\} \cdot \delta_a(\phi(\mathbf{x})) \tag{4}$$

where H_a and δ_a are regularized approximations of the Heaviside and Dirac functions. The matching algorithm for map refinement is as follows for each building:

1. Build the shape template level set ψ_0 from the map and initialize the active contour level set function: $\phi(t = 0, \mathbf{x}) = \psi_0(\mathbf{x})$.
2. Compute the mean and variance $\overline{I}^s_{in}(t)$, $\overline{I}^s_{out}(t)$, $\sigma^s_{in}(t)^2$, $\sigma^s_{out}(t)^2$.
3. Evolve the constrained active contour according to (4).
4. Optimize the parameters of T_{sim}.
5. Loop steps 2 to 4 until convergence. The convergence is reached when evaluations of the overall energy functional at two consecutive iterations is below a given threshold ε. In other words, the active contour minimizing the energy functional J_{SC} has converged at time $t_c = n_c \Delta t$ when the following condition is fulfilled:

$$|J_{SC}(\phi(\mathbf{x}, n_c \Delta t)) - J_{SC}(\phi(\mathbf{x}, (n_c - 1)\Delta t))| < \varepsilon$$

where $n_c \in \mathbb{N}^+$ and Δt is the time step discretizing the continuous temporal variable t.

From experiments, we found that a threshold ε ranging from 10^{-6} to 10^{-3} is a good compromise between computational time and matching accuracy. It is however important to notice that a too high threshold would make the active contour sensitive to local minima of the functional. We propose two contributions to address the latter issue in order to improve the robustness of the active contours matching:

Exogenous DSM fusion: First we propose to support building map refinement with the input of an exogenous DSM. The DSM data encoding the altitude of the scene components is not redundant with the satellite image. Therefore we could make them cooperatively drive the active contour to the building target to achieve fine matching. Unlike the satellite image, the DSM enables a good contrast of buildings from the rest of the scene, which is a desirable property for the piecewise smooth segmentation model that we use. The joint use of DSM and optical imagery data has already been achieved in (GUO & YASUOKA 2002). Unlike the former method, we do not need a high quality DSM since we do not use gradient-based active contours, which are sensitive to noise and artifacts. Moreover the embedded shape constraint and the implicit active contours representation of our scheme enable to overcome occlusions and manage complex buildings shapes, which is not the case in (Guo & Yasuoka 2002).

Flexible shape prior incorporation: The tuning of the shape constraint weight that balances the influence of the shape prior with respect to the image information is not trivial. Indeed, a too low weight prevents from accurate matching and from overcoming image corruption. Conversely a too high weight will weaken the intrinsic property of flexibility of active contours. Besides, a remote initialization of an active contour embodying a strong shape constraint may be sensitive to local minima of the functional: far from the target building, the image-based information, which drives the active contour, may be penalized by the predominant shape constraint. The active contour may converge to an undesired solution, preventing from carrying out map-to-image matching. We address this issue in turning the constant weight λ into a monotonically increasing function of the iteration time t. The weight is low at the beginning of the iterative process, allowing more shape freedom to the active contour. As a result, the active contour will converge more surely to the desired target in the image. As time goes by, the shape prior is enforced to recover contour regularization and to overcome image alterations. Additionally, λ is also a function of the shape prior ψ to confine the active contour freedom within a restricted space: λ is lower close to the reference template, and asymptotically tends to a higher constant far away from the reference template. This freedom space is gradually reduced as the amplitude of λ increases. Such spatio-temporal shape constraint weight conveys more freedom to the active contour in a restricted space in order to overcome the problem of local minima leading to unsuccessful map-to-image matching.

In summary, the two proposed contributions are formalized as additional terms in the original energy functional:

$$J_{flux,flex}(\phi,\psi) = J^s + \lambda_{DSM} J^d + \lambda_{flex}(\psi,t)\, J_{shape} \tag{5}$$

Superscript d refers to the DSM data, which influence is balanced by the constant weight λ_{DSM}. Details of the formulation of λ_{flex} and the optimization scheme to retrieve T_{sim} can be found in Bailloeul et al. 2005.

2.2 *Application scope*

Implicit assumptions have been made for the design of the newly proposed methodology. We intend to detail them in this section in order to define the application scope of our scheme. First the image-based terms of equation (5) partition image data into piecewise smooth components, which may limit our study to buildings with a quite smooth and homogeneous roof reflectance. Second, exogenous data to be merged must fulfill two consistency criteria:

a. Data must be superimposable. This raises the issue of data geometry and registration accuracy. Ideally both satellite image and DSM might be projected into the same geometry and may have a high registration precision in order to ensure that a given pixel in both data represents the same part of the considered building.

b. Data must depict the same scene. This raises the issue of data acquisition time. Since the DSM is made from different acquisition means then the image ones, it may be possible that some changes (building removal) happened between the acquisition times of the satellite image and the DSM. This will constrain DSM data fusion to be solely applicable to unchanged areas.

Cartographic data projection is orthoscopic. As a result, the initial active contour derived from the map will be closer to the target building footprint than its roof that we aim at matching. Matching the building roof is the most tractable solution since it is the most visible part of a building, which is moreover the part represented in the map. A too high footprint-to-roof discrepancy may be problematic since active contours techniques are intrinsically local and may not be able to match a too remote target building roof. This effect is non-existent in case we deal with orthoscopic remote sensing data. Otherwise it is significant for high buildings which exhibit sharp perspective distortion but negligible for low buildings. As a consequence, our scheme is applicable to nearly orthoscopic data or low buildings areas.

Last but not least, we assume that the map is free from shape errors and from generalization effect. A mistaken prior shape derived from the map may bias the matching process since the shape is not consistent with its representation in the image data (Fig. 8). Generalization effect embodies two aspects. The first one deals with the simplification in the map of a single building outline. This may have the same side effect as a mistaken cartographic object. The second is the inclusion of a group of buildings within the same cartographic object. In that case the entity to be matched in the image might not be homogeneous, which violates our first assumption toward piecewise smooth buildings roofs.

3 Results

3.1 *Data and preprocessing*

In our application, additional exogenous data is an orthoscopic DSM encoding the altitude of the same scene as a panchromatic Quickbird satellite image of Beijing city (0.6 m/pixel). Both data depict a dense urban area. This DSM was computed by edge preserving correlation of digitized stereoscopic aerial images couples (Paparoditis et al. 1998). The subsequent DSM was next orthorectified and reached 1 m planimetric and altimetric geocoding accuracy. The satellite image was rectified from terrain variations by the Beijing Institute of Surveying and Mapping (BISM) to reach 0.4 m geocoding accuracy. Since both data are geocoded in the same cartographic system, overlay is straight-forward, satisfying the first exogenous consistency requirement. However we may stress that the satellite image is not in orthoscopic geometry, which obliges us to carry out experiments over low buildings areas. The second data fusion requisite dictates data to represent the same object, and therefore raises the issue of data acquisition time. Indeed, the DSM we used in experiments was generated from 1999 aerial images whereas the satellite image is from year 2002. This anachronism constrained us to carry out experiments on areas where no change is noticeable between these dates. GIS data were manually generated from aerial imagery of the same area by the BISM. Buildings represented in the 2D map are vectorized polygons. The satellite image is pre-processed by anisotropic diffusion to enhance piecewise homogeneity.

3.2 *Experimental results*

Shape and location information from the GIS map enables a good global initial superimposition between the initial active contour and its counterpart representation in the image. However, we intentionally corrupted it in order to examine how well our method could manage inaccurate initial overlay of the data and to illustrate fine map-to-image matching. To do so, any original cartographic vectorized polygon is subject to a global similarity transformation before being overlaid onto the image to initialize the refinement process. This transformation is to be retrieved by the estimation of T_{sim} during the optimization step 4 exposed in the pseudo-code of section 2.1, which will in the end enable a successful matching. The first experiment illustrated in Fig. 2 shows

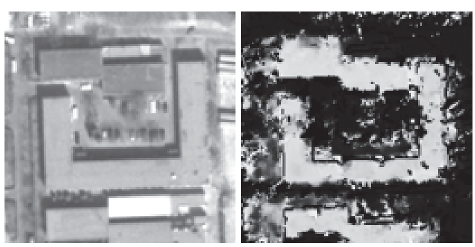

Fig. 1: U-shaped building represented in a satellite image (left) and a DSM which contrast has been enhanced by nonlinear clipping (right).

Fig. 2: Matching without shape prior. Left: initial state close to the desired solution. Right: result without shape prior, ''leak'' of the active contour.

the need for prior shape knowledge in the matching process. We deliberately initialize the active contour very close to the matching solution of a U-shaped building. The matching task performs poorly without shape prior derived from the map, even though the initialization is close to the desired solution: lack of contrast at the building borders makes the active contour ''leak'' all over the image, segmenting areas having similar statistical features. This problem is solved in incorporating shape prior (Fig. 3). A more remote initialization than Fig. 2 is possible since shape constraint incorporation is invariant from similarity transformation in order to achieve building fine matching. Figs. 4 to 6 illustrate experiments results carried out with two kinds of buildings and a remote initialization. The figures compare the constrained active contours model of equation (3) with our improved scheme (5). We notice that our scheme outperforms the model without exogenous data fusion or flexible shape prior constraint, enabling in both cases a satisfying matching (Figs. 5–6).

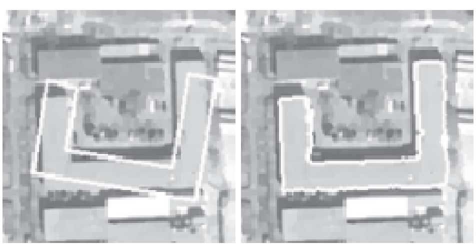

Fig. 3: Matching with shape prior and remote initialization ($\lambda = 10$). Left: initial state. Right: successful matching result even with remote initialisation.

The active contour driven by the model of equation (3) fails in matching the target building in the image (Fig. 4). Lack of discrimination of the building in the image as well as a predominant shape constraint are the two main reasons which may trap the energy functional minimization in a local minimum. The input of a soften shape constraint (Fig. 5) conveys more flexibility to the active contour while globally preserving the reference template shape in order to reach the target building. As we can see on the convergence sequence, the active contour topology may change at the beginning of the iterative process due to a low shape constraint. At the end of the matching action, extra blobs are naturally erased by a stronger shape constraint. Fig. 1 shows the DSM integrated into our model. This DSM looks rough, which is inherent to the method that generated it as well as the complexity of urban scenes. We may notice that some parts of the buildings are not well reconstructed, especially at the boundaries. However this representation of the building is not corrupted by shadows or peripheral objects located on the ground and allows a better discrimination of the building from the background. On the other hand, the representation from the satellite image yields quite clear building boundaries but with lack of discrimination and presence of artifacts. Fig. 6 shows how the complementarity of both satellite and DSM representations overcomes far initialization to carry out a successful matching. Since the DSM has a lower geocoding accuracy than the satellite image and exhibits reconstruction artifacts, we choose to drastically lower its contribution at the end of the convergence process ($\lambda_{DSM} \ll 1$) to favour the satellite image where the building outline is better defined. The use of our full model incorporating both flexible shape constraint and DSM data fusion would yield the same results as Figs. 5–6. The only difference arises in the convergence time (this issue will be investigated in the next sub-section).

The result of Fig. 3 showed there is no restriction towards the shape convexity of the target building to be matched in imagery data. The Fig. 7 illustrates this aspect further with a more complex topology of the building which contains an inner courtyard. As a consequence the cartographic object includes a hole, which is naturally handled by the flexible topology of the active contour implicitly represented by its level set function.

3.3 *Convergence time*

The Tab. 1 displays the computational convergence time ratio of the method in equation (3) with respect to our scheme performance (5). We investigate computational time distinction with the U-shaped and rectangular buildings and two different initializations. Time comparisons are shown with the sole fusion of DSM data, with the sole flexible shape constraint and finally our full model. The results demonstrate that DSM data fusion enables a faster convergence than the model of equation (3). The DSM represents more obviously the target building and drives more surely the active contour, which may explain the convergence time gain. The input of a flexible shape constraint increases even more the convergence efficiency. A lower shape constraint allows increasing the active contour evolution speed while globally keeping the prior shape information. This enables to quickly reach a rough and close solution to the target before the prior shape is gradually enforced. The flexible shape constraint always performs faster results, however while coupled with the DSM data it makes the active contour more sensitive to the DSM reconstruction artifacts and drives it away from the final solution before the DSM weight is relaxed and the prior shape enforced. As a result we obtain a slower convergence with the full model than the sole input of a flexible shape constraint. It is important to stress that the convergence time gain is image and initialization dependent and can not be stated theoretically from algorithmic considerations. However the results of Tab. 1 empirically show an obvious trend of computational cost decrease compared to the model of equation (3). Statistics over a high number of cases would be needed to indubitably confirm this inclination.

Tab. 1: Convergence time comparison.

Time ratio	DSM data fusion	Flexible shape constraint	DSM + flexible shape
U-shaped,1	1.4	1.8	1.6
Rect,1	1.3	3.1	1.8
U-shaped,2	2.5	4.3	2.1
Rect,2	1.3	2.0	2.0

3.4 *Discussion*

From experiments illustrated in Figs. 2 to 7 we showed how the input of an exogenous DSM and a flexible shape constraint could alleviate the sensitivity of the proposed refinement method with respect to incorrect GIS data position. It is worth noticing that a too low registration accuracy would not yield successful matching results as active contour techniques strongly rely on a good initialization, close to the desired solution. Since we address the issue of map refinement, we assume that the initial overlapping area between the map object and its counterpart representation in the image is quite

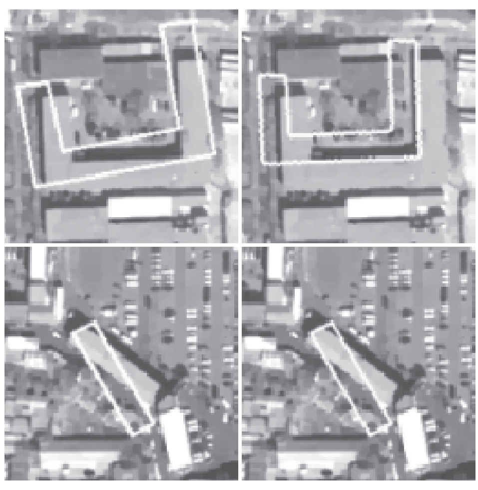

Fig. 4: U-shaped and rectangular buildings: failed matching with the model of equation (3) and $\lambda = 10$. Left: initial state. Right: result.

Fig. 5: Convergence sequence (from top to bottom and left to right) illustrating the matching process with a flexible shape constraint ($5 < \lambda_{flex} < 30$, $\lambda_{DSM} = 0$).

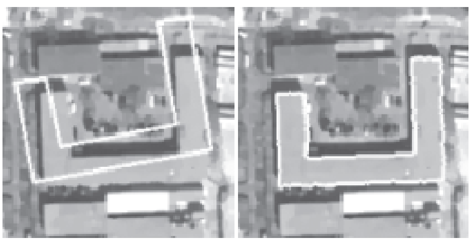

Fig. 6: Matching with DSM fusion ($\lambda_{flex} = cst$, $\lambda_{DSM} = 0.75$). Left: initial state. Right: successful matching with DSM fusion.

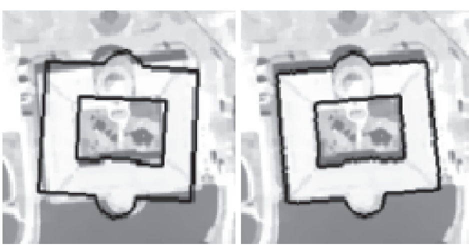

Fig. 7: Matching with flexible shape prior ($5 < \lambda_{flex} < 30$, $\lambda_{DSM} = 0$) and complex topology. Left: initial state. Right: successful matching result with complex topology.

high, which would guarantee a satisfying subsequent refinement. The quantification of such initial overlapping rate allowing a

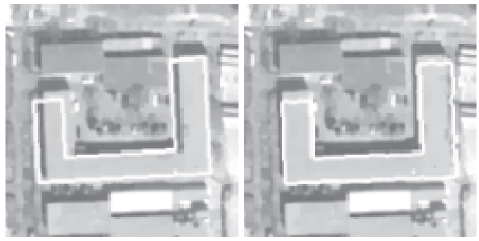

Fig. 8: Matching with flexible shape prior and DSM fusion and a locally mistaken cartographic object ($5 < \lambda_{flex} < 30$, $\lambda_{DSM} = 0.75$). Left: initial state derived from the map. Right: unsuccessful complete matching result.

successful matching is not an easy task as it may be image and object dependent. However, we can qualitatively state that the more the buildings are discriminated from the image background, the more successful the matching with a remote initialization.

Experiments were carried out while assuming that map objects are free from shape errors or generalization effects and that the pose inaccuracy can be modelled by a simple global transformation (similarity transform). The invariance of the algorithm from more complex global transformations (affine or perspective projections) is however possible to handle an extended class of form inaccuracies which may arise in the GIS data. Nevertheless, shape incorrectness from the map is usually local and cannot be modelled by a global transform. We illustrate this aspect with the experiment carried out with the U-shaped building which cartographic representation is locally mistaken at the top of the U branches (Fig. 8). The matching result is then a compromise between the image and the erroneous shape information contained in the map. The similarity invariance of the scheme enables a global refinement of the cartographic data, but local discrepancies between the target building and the reference template are too large at the top of the U branches to be recovered. This is inherent to the fact that no transformation input in J_{shape} is able to model such kind of local discrepancies and because the active contour shape is compelled to be akin to the mistaken shape information. Such hard incorporation of the shape constraint prevents from a completely successful matching. An intuitive solution consisting in relaxing locally the shape constraint to correct local discrepancies would face two difficulties. First it is not trivial to decide where the constraint has to be weakened. Second, the active contour may be subject to image corruptions and urban artifacts at locations where the prior shape information is relaxed. An alternative to handle the problem would be to reformulate the shape energy functional in order to allow a restricted class of displacement or discrepancies of the active contour from the prior shape template. Such restricted motion would allow the active contour to deviate from the cartographic shape prior while preserving its geometric properties at a local level. The formulation of such shape energy and its implementation is still an ongoing work.

4 Conclusions

We have introduced a novel scheme based on the use of active contours to refine digital building map using a high resolution satellite image. Our method is supported by data fusion: active contours are initialized and constrained by a GIS map and make use of an exogenous DSM to achieve successful matching. We demonstrated how the input of the DSM and a flexible spatio-temporal shape constraint could outperform traditional knowledge-based active contours while decreasing the computational cost. Future works will attempt to get rid of the limitations of the presented approach. Extension of our methodology to non-homogeneous or clustered buildings will be tackled with the incorporation in the active contours functional of edge information derived from the remote sensing image. Presence of mistakes or simplifications in the map is still an open and challenging question that we may consider by incorporating a new class of flexible shape constraint allowing larger discrepancies from the reference template. Finally, the proposed method is a preliminary step to building map-to-image change detection as it increases the consistency and comparability of both representations. The map refinement would alleviate the problem of exogenous discrepancies between the map and the image (inaccurate registration, mistakes and generalization effects in the map), enabling more robust subsequent change detection to state what cartographic buildings are no longer present in a more recent remote sensing image.

Acknowledgements

This project is supported by the Chinese Ministry of Science and Technology 863 programme and Alcatel Alenia Space. We thank the editors of the ISPRS Hannover Workshop 2005 and PFG for their invitation to publish the present article in a special issue of the journal.

References

AGOURIS, P., STEFANIDIS, A. & GYFTAKIS, S., 2001: Differential snakes for change detection in road segments. – Photogrammetric Engineering and Remote Sensing **67**(12): 1391–1399.

BAILLOEUL, T., PRINET, V., SERRA, B. & MARTHON, P., 2005: Spatio-temporal shape constraint for level set segmentation. – Proc. EMMCVPR, LNCS 3757, Springer-Verlag, Berlin–Heidelberg, pp. 503–519.

BALTSAVIAS, E., 2002. Object extraction and revision by image analysis using existing geospatial data and knowledge: State-of-the-art and steps towards operational systems. – Commission VI Mid-Term Symposium, Xian, China, International Archives of Photogrammetry and Remote Sensing, XXXIV (2): 13–22.

CHAN, T. & ZHU, W., 2003: Level set based shape prior segmentation. – Computational and Applied Mathematics technical report no 03–66 of the University of California, Los Angeles.

CHEN, Y., THIRUVENKADAM, S., TAGARE, H., HUANG, F., WILSON, D. & GEISER, E. A., 2001: On the incorporation of shape priors into geometric active contours. – Proc. of the IEEE 1st Workshop on Vatiational Framework and Level Sets methods, pp. 145–152.

CREMERS, D., OSHER, S. & SOATTO, S., 2004: Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: teaching level sets to walk. – Pattern Recognition, Proc. of DAGM 2004, Springer Verlag LNCS Vol. 3157, pp. 36–44.

Guo, T. & Yasuoka, Y., 2002: Snake-based approach for building extraction from high-resolution satellite images and height data in urban areas. – Proc. of the 23rd Asian Conference on Remote Sensing.

Kass, M., Witkin, A. & Terzopoulos, D., 1987: Snakes: active contour models. – 1st International Journal of Computer Vision **1**(4): 321–331.

Oriot, H., 2003: Statistical snakes for building extraction from stereoscopic aerial images. – International archives of photogrammetry, remote sensing and spatial information sciences, vol XXXIV-3/W8.

Osher, S. & Sethian, J., 1988: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. – Journal of Computational Physics **79:** 12–49.

Paparoditis, N., Cord, M., Jordan, M. & Cocquerez, J.-P., 1998: Building detection and reconstruction from mid- and high-resolution aerial imagery. – Computer Vision and Image Understanding **72**(2): 122–142.

Paragios, N. & Deriche, R., 2002: Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision. – Journal of Visual Communication and Image Representation **13**(1/2): 249–268.

Péteri, R. & Ranchin, T., 2003: Multiresolution Snakes for urban road extraction from IKONOS and Quickbird images. – In: Benes, T. (ed.): 23nd EARSeL Annual Symposium "Remote Sensing in Transition". – Mill-press, Rotterdam, Netherlands, Ghent, Belgium.

Rochery, M., Jermyn, I. & Zerubia, J., 2003: Higher order active contours and their application to the detection of line networks in satellite imagery. – IEEE Workshop on Variational, Geometrical, and Level Set Methods in Computer Vision, VLSM'03 at ICCV, Nice, France.

Vinson, S., Cohen, L. & Perlant, F., 2001: Extraction of rectangular buildings using DEM and orthoimage. – Proc. of Scandinavian Conference on Image Analysis (SCIA'01).

Addresses of the authors:

Timothée Bailloeul, PhD student
Dr. Véronique Prinet
LIAMA, Institute of Automation, Chinese Academy of Sciences, P. O Box 2728, Beijing 100080, China
Tel.: (+86 10) 82 61 44 62
Fax: (+86 10) 62 64 74 58
e-mail: tbailloeul@liama.ia.ac.cn
prinet@nlpr.ia.ac.cn

Dr.-Ing. Bruno Serra
Alcatel Alenia Space, 100 bd du Midi
06156 Cannes La Bocca, France
Tel.: (+33) 4 92 92 67 26
Fax: (+33) 4 92 92 76 60
e-mail: bruno.serra@alcatelaleniaspace.com

Prof. Dr. Philippe Marthon
LIMA (IRIT), 2 rue Camichel, 31071 Toulouse
France – Tel.:/ Fax: (+33) 5 61 58 83 53
e-mail: Marthon@enseeiht.fr

Pinxiang Chen, Haitao Zhang
Beijing Institute of Surveying and Mapping
15 Yangfangdianlu, Haidian District
Beijing 100038, China
Tel.:(+86 10) 63951326, Fax:(+86 10) 63963144
e-mail: cpx@bism.cn
Tel.:(+86 10) 63953948, Fax:(+86 10) 63971019
e-mail: zhanght@bism.cn

Manuskript eingereicht: Juni 2005
Angenommen: Juli 2005