Building Extraction and Change Detection in Multitemporal Remotely Sensed Images with Multiple Birth and Death Dynamics

Csaba Benedek, Xavier Descombes and Josiane Zerubia

Ariana Project Team (INRIA/CNRS/UNSA)

2004 route des Lucioles, BP 93, 06902 SOPHIA ANTIPOLIS Cedex - FRANCE

`firstname.lastname@sophia.inria.fr`

Abstract

In this paper we introduce a new probabilistic method which integrates building extraction with change detection in remotely sensed image pairs. A global optimization process attempts to find the optimal configuration of buildings, considering the observed data, prior knowledge, and interactions between the neighboring building parts. The accuracy is ensured by a Bayesian object model verification, meanwhile the computational cost is significantly decreased by a non-uniform stochastic object birth process, which proposes relevant objects with higher probability based on low-level image features.

1. Introduction

Following the evolution of built-up regions is a key issue of high resolution aerial and satellite image analysis. Numerous previous approaches address building extraction [5, 7, 9] at a single time instance. This process can be highly facilitated by using Digital Elevation/Surface Model inputs (DEM/DSM) [3, 7] extracted from stereo image pairs as the buildings can be separated from the ground by the estimated height data. However in lack of multiview information, the building identification becomes a challenging monocular object recognition task [8].

Recent approaches on building change detection [3] assume usually that for the earlier time layer a topographic building database is already available, thus the process can be decomposed into old model verification and new building exploration phases. On the other hand, many image repositories do not contain meta data, therefore the task requires automatic building detection in each image.

Several low level change detection methods have been proposed for remote sensing [2], which search for statistically unusual differences between the images without using explicit object models. Although they are usually considered as preprocessing filters, there have been less attempts given to justify how they can support the object level inves-tigations. In contrary, our method combines object extraction with local low level similarity information between the corresponding image parts in a unified probabilistic model. It will be shown that we can benefit from evidences such as building changes can be found in the *changed* areas, while multiple object views from the different time layers may increase the detection accuracy of the *unchanged* buildings.

Another important issue is related to object modeling. The *bottom-up* techniques [5] construct the buildings from primitives, like roof blobs, edge parts or corners. Although these methods can be fast, they may fail if the primitives cannot be reliably detected. On the other hand, *inverse methods* [4] assign a fitness value to each possible object configuration and an optimization process attempts to find the configuration with the highest confidence. In this way, flexible object appearance models can be adopted, and it is also straightforward to incorporate prior shape information and object interactions. However, large computational cost is needed for the search in the high dimension population space meanwhile local maxima of the fitness function can mislead the optimization.

In the proposed model we attempt to merge the advantages of both low level and object level approaches. The applied Multiple Birth and Death technique [4] evolves the population of buildings by alternating object proposition (*birth*) and removal (*death*) steps in a simulated annealing framework. The exploration in the population space is driven by simple region descriptors, however the object verification follows the robust *inverse* modeling approach.

2. Problem formulation

The input of the proposed method consists of two co-registered aerial or satellite images which were taken from the same area with several months or years time differences. We expect the presence of registration or parallax errors, but we assume that they only cause distortions of a few pixels. We consider each building to be constructed from one or many rectangular building segments, which we aim to ex-

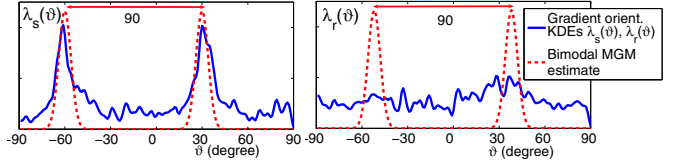Figure 1. Demonstration of the rectangle parameters

Figure 2. Kernel density estimation of the local gradient orientations over rectangles around two selected pixels: a building center s and an empty site r.

tract by the following model.

Denote by S the common pixel lattice of the input images and by $s \in S$ a single pixel. Let u be a rectangular building segment candidate. For purposes of dealing with multiple time layers we assign to u an image index flag $\xi(u) \in \{1, 2, *\}$, where '$*$' indicates unchanged object, while '1' and '2' correspond to building segments which appear *only* in the first *or* second image respectively. Let be $R_u \subset S$ the set of pixels corresponding to u. R_u is described by the five rectangle parameters: c_x and c_y center coordinates, e_L, e_l side lengths and $\theta \in [-90°, +90°]$ orientation (see Fig. 1).

3. Feature selection

In the proposed model, low level and object level features are distinguished. Low level descriptors are extracted around each pixel such as typical color or texture, and local similarity between the time layers. They are used by the exploration process to estimate where the buildings *can* be located, and how they *can* look like: the *birth* step generates objects in the estimated built-up regions with higher probability. On the other hand, object level features characterize a given object candidate u, exploited for the fitness calculation of the proposed oriented rectangles. Building verification is primarily based on the object level features thus their accuracy is crucial. Since apart from the similarity measure, the upcoming descriptors are calculated for the two input images separately, we often do not indicate the image index in this section.

3.1. Low level features of building identification

The first feature exploits the fact that regions of buildings should contain edges in *perpendicular* directions, which can be robustly characterized by local gradient orientation histograms [6]. Let be ∇g_s the intensity gradient vector at s with magnitude $||\nabla g_s||$ and angle ϑ_s. Let be $W_l(s)$ the rectangular $l \times l$ sized window around s, where l is chosen so that $W_l(s)$ can cover an average building narrowly. For each s we calculate the weighted ϑ_s density of $W_l(s)$:

$$\lambda_s(\vartheta) = \frac{1}{N_s} \sum_{r \in W_l(s)} \frac{1}{h} \cdot ||\nabla g_r|| \cdot k \left(\frac{\vartheta - \vartheta_r}{h} \right)$$

where $N_s = \sum_{r \in W_l(s)} ||\nabla g_r||$ and h is the kernel bandwidth parameter, we used uniform kernels for quick calcu-

lation. If $W_l(s)$ covers a building, the $\lambda_s(\vartheta)$ function has two peaks located in $90°$ distance in the ϑ-domain (Fig. 2). This property can be measured by correlating $\lambda_s(\vartheta)$ with an appropriately matched bi-modal density function:

$$\alpha(s, m) = \int \lambda_s(\vartheta) \eta_2 (\vartheta, m, d_\lambda) \, d\vartheta$$

where $\eta_2(.)$ is a mixture of two Gaussians with mean values m and $m + 90°$ respectively, and a same deviation d_λ for both components (d_λ is parameter of the process). Offset (m_s) and value (α_s) of the maximal correlation can be obtained as:

$$m_s = \text{argmax}_{m \in [-90°, 0]} \{\alpha(s, m)\} \quad \alpha_s = \alpha(s, m_s)$$

Pixels with high α_s are more likely centers of buildings, which can be coded in an α-birth map $P_b^\alpha(s) = \alpha_s / \sum_{r \in S} \alpha_r$. The nomination comes from the fact that the frequency of proposing an object in s will be proportional to the local birth factor $P_b(s)$.

On the other hand, offset m_s offers an estimate for the dominant gradient direction in $W_l(s)$. Thus for object u proposed with center s, we model its orientation as $\theta(u) = m_s + \eta_s$, where η_s is a zero-mean Gaussian random variable with a small deviation parameter σ_θ.

We have observed in various experiments that the α_s-gradient feature is usually able to roughly estimate the built-up regions. However, in several cases the detection can be refined considering other descriptors such as roof colors or shadows [9]. Some of the roof colors can be filtered using illumination invariant color representations, as the hue channel in HSV color space. Assume that we can extract in this way a $\mu_c(s) \in \{0, 1\}$ indicator mask, where $\mu_c(s) = 1$ means that pixel s has roof color. We calculate the color feature for s as $\Gamma_s = \sum_{r \in W_l(s)} \mu_c(r)$ and the color birth-map as $P_b^c(s) = \Gamma_s / \sum_{r \in S} \Gamma_r$. Note that obviously this information cannot be used for grayscale inputs, and even

(a) Object candidate (b) Gradient map (c) Masked gradient map

Figure 4. Demonstration of the gradient feature

Figure 3. Comparing the $\lambda(.)$ functions in the two image layers regarding two selected pixels. s corresponds to an unchanged point and r to a built-up change.

(a) Red roof (b) Color mask

Figure 5. Demonstration of the color roof feature

in color images the $\mu_c(s)$ filter usually finds only a part of the roofs which have typical 'red colors' ([9] and Fig 5(b)).

Another evidence for the presence of buildings can be obtained by the detection of their cast shadows [5, 9]. Exploiting that the darkness and direction of shadows are global image features, one can often extract a (noisy) binary shadow mask $\mu_{\text{sh}}(s)$, for example by filtering pixels from the *dark-blue* color domain [9]. Thereafter building candidate regions can be identified as image areas lying next to the shadow blobs in the opposite shadow direction (Fig. 6). We used a constant birth rate $P_b^{\text{sh}}(s) = p_0^{\text{sh}}$ within the obtained candidate regions and a significantly smaller constant ϵ_0^{sh} outside.

Since the main goal of the *combined birth map* is to keep focus on all building candidate areas, we derived it with the maximum operator from the feature birth maps: $P_b(s) = \max\left\{P_b^\alpha(s), P_b^c(s), P_b^{\text{sh}}(s)\right\} \forall s \in S$. For input without shadow or color information we can ignore the corresponding feature in a straightforward way. Note that we generate birth and orientation maps for both images which will be denoted by $P_b^{(i)}(s), m_s^{(i)}, i \in \{1, 2\}$.

3.2. Low level similarity feature

The gradient orientation statistics also offer a tool for low level region comparison. Matching the $\lambda_s^1(.)$ and $\lambda_s^2(.)$ functions can be considered as low level similarity checking of the areas around s in the two images, based on "building-focused" textural features (Fig 3). Moreover, these descriptors are independent of illumination and coloring effects, and robust regarding parallax and registration errors. For measuring the local textural dissimilarities, we used the Bhattacharyya distance of the distributions:
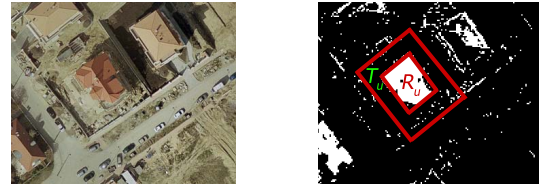
$$b(s) = -\log \int \sqrt{\lambda_s^1(\vartheta) \cdot \lambda_s^2(\vartheta)} d\vartheta$$

The binary similarity map is obtained as $B(s) = 1$ iff $b(s) < b_0$, $B(s) = 0$ otherwise.

3.3. Object-level features

In this section we introduce different object level image features. Based on them we define energy terms denoted by $\varphi^{(i)}(u)$ which evaluate the building hypothesis for u in the ith image (hereafter we ignore the i superscript). $\varphi(u)$ is interpreted as the negative building fitness value and a rectangle with $\varphi(u) < 0$ is called an *attractive* object. Since adding attractive objects may decrease the energy of the population [4], they are efficient building candidates.

We begin with gradient analysis. Below the edges of a relevant rectangle candidate R_u we expect that the magnitudes of the local gradient vectors are high and the orientations are close to the normal vector of the closest rectangle side (Fig. 4). Λ_u feature is calculated as:

$$\Lambda_u = \frac{1}{q_u} \cdot \sum_{s \in \tilde{\partial} R_u} ||\nabla g_s|| \cdot \left|\cos\left(\vartheta_s - \Theta_u^s\right)\right|$$

where $\tilde{\partial} R_u$ is the dilated edge map of rectangle R_u, $\Theta_u^s \in \{\theta(u), \theta(u) + 90°\}$ is the edge orientation of R_u around $s \in \partial R_u$, and q_u is the number of the pixels in $\tilde{\partial} R_u$. The data-energy term is calculated as: $\varphi_\Lambda(u) = \mathcal{Q}(\Lambda_u, d_\Lambda, D_\Lambda)$ where the following non-linear $\mathcal{Q}$ function is used [4]:

$$\mathcal{Q}(x, d_0, D) = \begin{cases} \left(1 - \frac{x}{d_0}\right) & \text{if } x < d_0 \\ \exp\left(-\frac{x - d_0}{D}\right) - 1 & \text{if } x \geq d_0 \end{cases}$$

The calculation of the *roof color* feature is demonstrated in Fig. 5. Here we define the T_u object-neighborhood and calculate the $\mathcal{C}_R(u) = \frac{1}{\#R_u} \cdot \sum_{s \in R_u} \mu_c(s)$ internal and $\mathcal{C}_o(u) = \frac{1}{\#T_u} \cdot \sum_{s \in T_u} [1 - \mu_c(s)]$ external filling factors (# denotes the area in pixels). Finally the energy term is set as $\varphi_\mathcal{C}(u) = \max\left[\mathcal{Q}(\mathcal{C}_R(u), d_R^\mathcal{C}, D_R^\mathcal{C}), \mathcal{Q}(\mathcal{C}_o(u), d_o^\mathcal{C}, D_o^\mathcal{C})\right]$

The *shadow term* is derived in analogous manner, but we locate the checked neighborhood area T_u^{sh} in the shadow

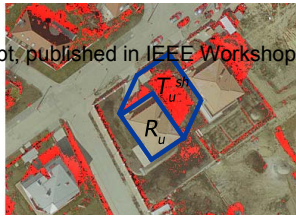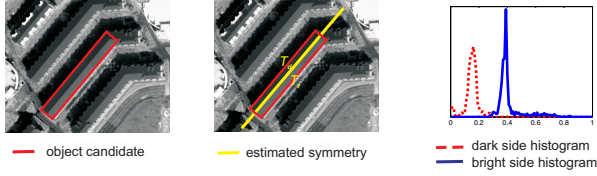Figure 6. Demonstration of the shadow feature

— object candidate   — estimated symmetry   -- dark side histogram
                                              — bright side histogram

Figure 7. Demonstration of the roof homogeneity feature

☐ Incomplete object   ☐ Floodfill propagated   ☐ u: true object   ☐ F_u: floodfill mask
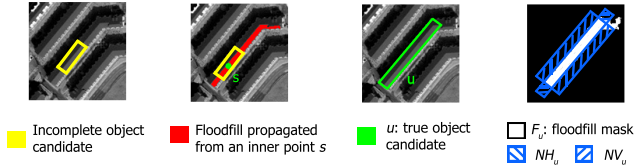   candidate   from an inner point s candidate   ◪ NH_u ◪ NV_u

Figure 8. Floodfill based feature for roof completeness

direction (Fig. 6). Thereafter we derive the internal resp. external values $\chi_R(u) = \frac{1}{\#R_u} \sum_{s \in R_u} \left[1 - \mu_{\text{sh}}(s) \right]$ and $\chi_o(u) = \frac{1}{\#T_u^{\text{sh}}} \sum_{s \in T_u^{\text{sh}}} \mu_{\text{sh}}(s)$, while the energy term $\varphi_\chi(u)$ is calculated in the same way as $\varphi_\mathcal{C}(u)$. Note that the $\varphi_\chi(u)$ term proved to be robust even if the shadow blobs had various sizes due to the diversity of building heights.

In grayscale satellite images *roof homogeneity* offers often another useful feature. Fig. 7 shows an example of how to describe two-side homogenous roofs. After extracting the symmetry axis of the object candidate u, we can characterize the "peakedness" of the dark and bright side histograms by calculating their kurtosis $\kappa_d(u)$ and $\kappa_b(u)$ respectively. However, as shown in Fig. 8 the homogeneity feature may have false maxima for incomplete roofs, therefore roof completeness should be measured at the same time. Thus we derive the F_u floodfill mask of u, which contains the pixels reached by floodfill propagations from the internal points of R_u. If the homogenous roof is complete, F_u must have low intersection with the NH_u resp. NV_u 'horizontal' and 'vertical' neighborhood regions of R_u (Fig. 8). Finally, the $\varphi_\kappa(u)$ energy term can be constructed from the kurtosis and completeness descriptors in a similar manner to the previous attributes.

The proposed framework enables flexible *feature integration* depending on the image properties. For each building prototype we can prescribe the fulfillment of one or many feature constraints whose φ-subterms are connected with the max operator in the prototype's joint energy term (logical AND in the negative fitness domain). In a given

image pair several building prototypes can be detected simultaneously, we combine the different prototypes with the min (logical OR) operator. For example, in the Budapest pair (Fig. 11) we use two prototypes: the first prescribes the edge and shadow constraints, the second one the roof color alone, thus the joint energy is calculated as:

$$\varphi(u) = \min \left\{ \max \left\{ \varphi_\Lambda(u), \varphi_\chi(u) \right\}, \varphi_c(u) \right\}.$$

4. Marked Point Process model

Let be $\mathcal{H}$ the space of u objects. Using a bounded Borel set $H \in \mathcal{H}$, the Ω configuration space is defined as [4]:

$$\Omega = \bigcup_{n=0}^{\infty} \Omega_n, \quad \Omega_n = \left\{ \{u_1, \ldots, u_n\} \in H^n \right\}$$

Denote by ω an arbitrary object configuration $\{u_1, \ldots, u_n\}$ in Ω. We define a $\sim$ neighborhood relation in $\mathcal{H}$: $u \sim v$ if their rectangles R_u and R_v intersect.

We introduce a non-stationary data-dependent Gibbs distribution on the configuration space as: $P_\mathcal{D}(\omega) = 1/Z \cdot \exp\left[-\Phi_\mathcal{D}(\omega) \right]$, where Z is a normalizing constant, and

$$\Phi_\mathcal{D}(\omega) = \sum_{u \in \omega} A_\mathcal{D}(u) + \gamma \cdot \sum_{\substack{u,v \in \omega \\ u \sim v}} I(u, v) \tag{1}$$

Here $A_\mathcal{D}(u)$ and $I(u,v)$ are the data dependent unary and the prior interaction potentials, respectively and γ is a weighting factor between the two energy terms. Thus the maximum likelihood configuration estimate according to $P_\mathcal{D}(\omega)$ can be obtained as $\omega_{\text{ML}} = \arg\min_{\omega \in \Omega} \left[\Phi_\mathcal{D}(\omega) \right]$.

Unary potentials characterize a given building segment candidate $u = \{c_x, c_y, e_L, e_l, \theta, \xi\}$ as a function of the local image data in both images, but independently of other object of the population:

$$A_\mathcal{D}(u) = \mathbb{I}_{[\xi(u) \in \{1,*\}]} \cdot \varphi^{(1)}(u) + \mathbb{I}_{[\xi(u) \in \{2,*\}]} \cdot \varphi^{(2)}(u) +$$

$$+ \frac{\gamma_\xi}{\#R_u} \left\{ \mathbb{I}_{[\xi(u)=*]} \sum_{s \in R_u} \left(1 - B(s) \right) + \mathbb{I}_{[\xi(u) \in \{1,2\}]} \sum_{s \in R_u} B(s) \right\}$$

where $\mathbb{I}_{[E]} \in \{0,1\}$ is the indicator function of event E, and as defined earlier $\varphi^{(1)}(u)$ and $\varphi^{(2)}(u)$ are the building energies in the 1st resp. 2nd image (Sec. 3.3), while $B(.)$ is the low level similarity mask between the two time layers (Sec. 3.2). The last term penalizes unchanged objects ($\xi(u) = *$) in the regions of textural differences, and new/demolished buildings ($\xi(u) \in \{1,2\}$) in *changeless* areas.

On the other hand *interaction* potentials enforce prior geometrical constraints: they penalize intersection between different object rectangles sharing the time layer (Fig. 4):

$$I(u, v) = \mathbb{I}_{[\xi(u) \simeq \xi(v)]} \cdot \frac{\#(R_u \cap R_v)}{\#(R_u \cup R_v)}$$

where $\xi(u) \simeq \xi(v)$ relation holds iff $\xi(u) = \xi(v)$, or $\xi(u) = *$, or $\xi(v) = *$.
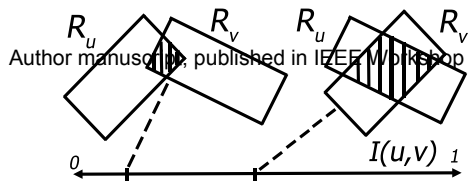
Figure 9. Intersection feature

Table 1. Main properties of the test data sets.

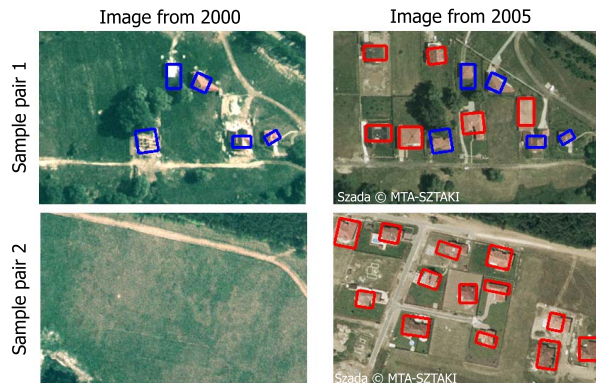| Data Set	Type	Color	Shadow	Gradient	Kurtosis
Budapest	Optical	Yes	Yes	Good	Partial
BEIJING	QBird	No	Yes	Weak	Partial
SZADA	Optical	Yes	No	Weak	No
ABIDJAN	Ikonos	No	No	Sharp	Yes

Figure 10. Results on two samples from the SZADA images (source: MTA-SZTAKI©). Blue rectangles denote the detected unchanged objects, red rectangles the changed (new, demolished or modified) ones.

5. Optimization

We estimate the optimal object configuration by the Multiple Birth and Death Algorithm [4] as follows:

Initialization: calculate the $P_b^{(i)}(s)$ and $m_s^{(i)}$ ($i \in \{1, 2\}$) birth maps, and start with an empty population $\omega = \emptyset$.

Main program: initialize the inverse temperature parameter $\beta = \beta_0$ and the discretization step $\delta = \delta_0$, and alternate birth and death steps.

1. *Birth step*: for each pixel $s \in S$, if there is no object with center s in the current configuration ω, pick up $\xi \in \{1, 2, *\}$ randomly, let be $\widehat{P}_b = P_b^{(\xi)}(s)$ if $\xi \in \{1, 2\}$, $\widehat{P}_b = \max\{P_b^{(1)}(s), P_b^{(2)}(s)\}$ if $\xi = *$; and choose birth in s with probability $\delta \widehat{P}_b$.

   If birth is chosen in s: generate a new object u with center s, image index ξ, set the $e_L(u)$, $e_l(u)$ parameter randomly between prescribed maximal and minimal side lengths, and orientation $\theta(u)$ following the $\eta(., m_s^{(\xi)}, \sigma_\theta)$ Gaussian distribution as shown in Sec. 3.1. Finally, add u to the current configuration ω.

2. *Death step*: Consider the configuration of objects $\omega = \{u_1, \ldots, u_n\}$ and sort it from the highest to the lowest value of $A_\mathcal{D}(u)$. For each object u taken in this order, compute $\Delta\Phi_\omega(u) = \Phi_\mathcal{D}(\omega/\{u\}) - \Phi_\mathcal{D}(\omega)$, derive the *death rate* as follows:

$$d_\omega(u) = \frac{\delta a_\omega(u)}{1 + \delta a_\omega(u)}, \quad \text{with} \quad a_\omega(u) = e^{-\beta \cdot \Delta\Phi_\omega(u)}$$

and remove u from ω with probability $d_\omega(u)$.

Convergence test: if the process has not converged yet, increase the inverse temperature β and decrease the discretization step δ with a geometric scheme, and go back to the birth step. The convergence is obtained when all the objects added during the birth step, and only these ones, have been killed during the death step.

6. Experiments

We evaluated our method on four significantly different data sets[1], whose main properties are summarized in Table 1. Qualitative results are shown in Fig. 10–12.

---

[1] The authors would like to thank the test data providers: András Görög, Budapest; French Defense Agency (DGA); Liama Laboratory of CAS, China; and MTA-SZTAKI, Hungary.

To justify the fact that we addressed both object extraction and change detection in the same probabilistic framework, we compared the proposed method (hereafter joint detection - JD) to the conventional approach where the buildings are separately extracted in the two image layers, and the change information is posteriorly estimated through comparing the location and geometry of the detected objects (separate detection - SD). As Fig. 12 shows, the SD method causes false change alarms as low contrasted objects may be erroneously missed from one of the image layers, and due to noise, false objects can appear frequently in case of the less robust one-view information.

Relevance of the applied multiple feature based building appearance models is compared to the Edge Verification (EV) method. In EV similarly to [9], shadow and roof color information is only used to coarsely detect the built-in areas, while the object verification is purely based on matching the edges of the building candidates to the Canny edge map extracted over the estimated built-in regions.

In the quantitative evaluation we measured the number of missing and falsely detected objects (MO and FO), missing and false change alarms (MC, FC), and the pixel-level accuracy of the detection (DA). For the DA-rate we compared the resulting building footprint masks to the ground truth mask, and calculated the F-rate of the detection (harmonic mean of precision and recall). Results in Table 2 confirm the generality of the proposed model and the superiority of the joint detection (JD) framework over the SD and EV approaches (lower object-level errors, and higher DA rates). Further details of evaluation can be found in [1].

Figure 11. Results of the proposed model (JD) on two image pairs. Top: BUDAPEST data (only an image part is shown - source: András Görög©). Bottom: BEIJING (Liama Laboratory CAS© China). Unchanged (blue) and changed (red) objects are distinguished.

Table 2. Quantitative evaluation results. #CH and #UCH denote the total number of changed resp. unchanged buildings in the set. JD refers to the proposed model; reference methods EV & SD and evaluation rates MO, FO, MC, FC & DA are defined in Sec. 6.

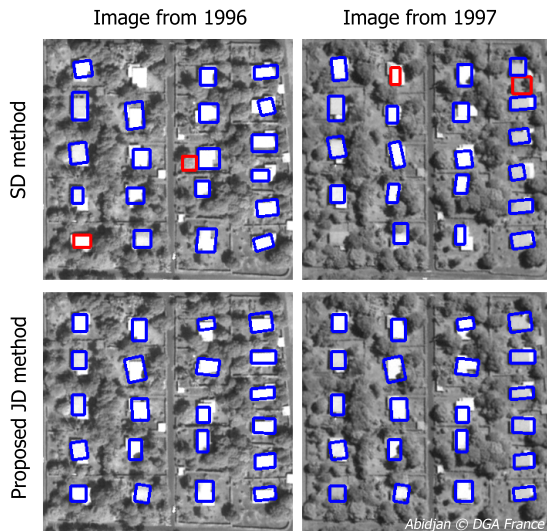| Data Set	#CH	#UCH	MO			FO			MC			FC			DA		
			EV	SD	**JD**	EV	SD	**JD**	EV	SD	**JD**	EV	SD	**JD**	EV	SD	**JD**
BUDAPEST	20	21	3	3	1	8	8	2	3	1	1	5	11	1	0.73	0.70	0.78
BEIJING	13	4	0	1	0	5	2	1	0	0	0	2	3	0	0.48	0.77	0.85
SZADA	31	6	4	3	1	1	0	1	3	3	2	2	3	0	0.78	0.74	0.83
ABIDJAN	0	21	1	2	0	0	2	0	0	0	0	0	4	0	0.84	0.78	0.91

Figure 12. Results on ABIDJAN images (DGA© France). Top: separate detection (SD) method, where all the indicated changes are false alarms. Bottom: proposed joint model (JD).

References

[1] C. Benedek, X. Descombes, and J. Zerubia. Building extraction and change detection in multitemporal aerial and satellite images in a joint stochastic approach. Research report, INRIA Sophia Antipolis, October 2009.

[2] F. Bovolo. A multilevel parcel-based approach to change detection in very high resolution multitemporal images. *IEEE GRS Letters*, 6(1):33–37, 2009.

[3] N. Champion, L. Matikainen, X. Liang, J. Hyyppa, and F. Rottensteiner. A test of 2D building change detection methods: Comparison, evaluation and perspectives. In *ISPRS Congress*, pages 297–304, Beijing, China, 2008.

[4] X. Descombes, R. Minlos, and E. Zhizhina. Object extraction using a stochastic birth-and-death dynamics in continuum. *J. of Math. Imaging and Vision*, 33(3):347–359, 2009.

[5] A. Katartzis and H. Sahli. A stochastic framework for the identification of building rooftops using a single remote sensing image. *IEEE Trans. GRS*, 46(1):259–271, 2008.

[6] S. Kumar and M. Hebert. Man-made structure detection in natural images using a causal multiscale random field. In *CVPR*, volume 1, pages 119–126, 2003.

[7] F. Lafarge, X. Descombes, J. Zerubia, and M. Pierrot-Deseilligny. Structural approach for building reconstruction from a single DSM. *IEEE Trans. PAMI*, 2009. in press.

[8] J. Shufelt. Performance evaluation and analysis of monocular building extraction from aerial imagery. *IEEE Trans. PAMI*, 21(4):311–326, 1999.

[9] B. Sirmacek and C. Unsalan. Building detection from aerial imagery using invariant color features and shadow information. In *IEEE ISCIS*, Istanbul, Turkey, 2008.