

Épületek és változásaik detekciója sztochasztikus módszerekkel^{*}

Benedek Csaba^{1,2}, Xavier Descombes², Josiane Zerubia²

¹ Elosztott Események Elemzése Kutatócsoport, MTA SZTAKI,
1111 Budapest, Kende utca 13-17
bcsaba@sztaki.hu

² Ariana Project Team (INRIA/CNRS/UNSA), 2004 route des Lucioles, BP 93, 06902
SOPHIA ANTIPOLIS Cedex - FRANCE
firstname.lastname@sophia.inria.fr

Absztrakt. Cikkünkben légi- és műholdképek feldolgozását végző valószínűségi módszert mutatunk be. A modell együttesen célozza meg épületek kinyerését és változásaik detektálását többidejű, jelentős időkülönbséggel készült képi bemeneteket feltételezve. Globális optimalizációs folyamat találja meg az épületek megfelelő konfigurációját, figyelembe véve a megfigyelt képi adatot, prior információt és a szomszédos épületszegmensek közötti interakciót. A módszer pontosságát Bayesi verifikációs folyamat biztosítja, míg a számítási idő jelentősen lerövidül egy új sztochasztikus objektum születési folyamat segítségével, mely alacsony képi jellemzők alapján nagy valószínűséggel javasol releváns objektumokat.

1. Bevezetés

Beépített területek változásainak monitorozása központi feladat nagy felbontású légi- és műholdképek vizsgálatakor. Számos módszer végez épületdetekciót különálló képeken [5, 7, 9]. Ez a folyamat jelentősen megkönnyíthető DEM/DSM adat-bemenetek felhasználásával [3, 7], melyek sztereo képpárokából vagy speciális szenzoradatokból nyerhetők ki: ekkor a házak elkülöníthetők a földfelszíntől a becsült magassági adatok alapján. Több nézeti információ hiányában azonban az épületek azonosítása komoly kihívássá válik [8].

Az épületváltozások észlelését végző számos korábbi módszer [3] feltételezi, hogy a korábbi időpontról topografikus épület adatbázis áll rendelkezésre, így a folyamat szétválasztható két részre: a korábbi épületek ellenőrzésére és az új épületek felderítésére. Ugyanakkor sok képi adatbázis nem tartalmaz kiegészítő adatokat, ezért szükségessé válik mindkét időrétegen az automatikus épületkinyerés.

A távérzékelésben korábban számos alacsony szintű változásdetekciós módszert ismertettek [2], melyek statisztikailag kiugró különbségeket keresnek a képeken konkrét

^{*} A cikk eredményei az alábbi publikációban jelentek meg: C. Benedek, X. Descombes, and J. Zerubia. Building extraction and change detection in multitemporal remotely sensed images with multiple birth and death dynamics. In *IEEE Workshop on Applications of Computer Vision (WACV)*, pages 100–105, 2009.

objektum modellek használata nélkül. Bár ezeket a módszereket általában csak előfeldolgozó lépéseknek tekintik a szakirodalomban, viszonylag kevés olyan kutatási eredményt tettek közzé, amely felhasználhatóságukat bizonyítaná objektum szintű vizsgálatokban. Ezt a hiányosságot pótolva, módszerünk kombinálja az objektumfelismerést lokális alacsony szintű hasonlóság vizsgálatával, melyek egy egyesített valószínűségi modellben működnek. Bemutatjuk, hogy különböző alapvető tényezők együttesen használhatók ki, úgy mint *épületváltozások* a textúra szinten is megváltozott területeken található, míg a *változatlan* objektumokról rendelkezésünkre álló több nézetű információ pontosítja a felismerést.

A módszerek másik fontos csoportosítási szempontja az objektumok modellezésén alapul. A direkt módszerek [5] az épületeket primitív elemekből építik fel; a leggyakrabban használt primitívek a tető blokkok, él részletek vagy sarokpontok. Bár ezek az eljárások általában gyorsak, hibáznak, ha a képeken a primitíveket nem sikerül megbízhatóan kinyerni. Egy másik jelentős módszertani csoportba tartoznak az *inverz* modellek [4], melyek fitness értéket rendelnek a különböző lehetséges objektumkonfigurációkhoz, és egy optimalizációs folyamattal keresik a legvalószínűbb populációt. Így rugalmas objektummodellek építhetők, és könnyen vehetünk figyelembe prior alakú információkat és objektum interakciókat. Ugyanakkor, nagy számítási igény szükséges a magas dimenziójú populációterben történő kereséshez, mialatt a fitness függvény lokális maximumai félrevezethetik az optimalizációt.

Módszerünkben egyaránt kihasználjuk az alacsony és az objektumszintű megközelítések előnyeit. Az felhasznált többszörös születés és halál (Multiple Birth and Death, MBD) optimalizáció [4] az épületek populációját objektumgenerálás (születés) és eldobás (halál) lépések iterációival fejleszti lépésről lépésre, szimulált lehűtés keretrendszerbe ágyazva. A felderítés fázist alacsony szintű képi leírók vezérlik, míg az objektum verifikáció a robotsztus inverz megközelítést követi.

2. Probléma definíció

A bemutatásra kerülő módszer bemenete két egymáshoz regisztrált légi- vagy műholdkép, melyeket ugyanarról a földrajzi területről készítettek több hónapos vagy éves időkülönbséggel. Bár figyelembe vesszük, hogy az eltérő nézetekből adódóan képregisztrációs hibák lehetnek jelen, feltételezzük, hogy ezek csak néhány pixelnyi torzítást okoznak. Az épületeket felülnézetből téglalap alakú szegmensek együttesének tekintjük, így a következőkben csak ezekre a szegmensekre koncentrálunk.

Jelöljük S -sel a képek közös pixelrácsát, és $s \in S$ -sel egy adott pixelt. Legyen u egy téglalap alakú épület szegmens jelölt. Azért hogy együtt kezelhessünk két időréteget, u -hoz egy képindex attribútumot is rendelünk $\xi(u) \in \{1, 2, *\}$, ahol ‘*’ változatlan épületet jelöl, míg ‘1’ illetve ‘2’ olyan szegmensekhez tartozik, melyek csak az első illetve második képen szerepelnek. Legyen $R_u \subset S$ az u -hoz tartozó a pixelek halmaza. R_u -t öt téglalap-paraméter írja le: c_x és c_y középponti koordináták, e_L , e_l oldal hosszak és $\theta \in [-90^\circ, +90^\circ]$ orientációs szög (1. ábra).

3. Jellemzőválasztás

Módszerünkben alacsony szintű és objektumszintű jellemzőket használunk együttesen. Az alacsony szintű jellemzőket a pixelek környezetéből nyerjük ki, ilyenek lehetnek tipikus színek, textúrák és lokális hasonlóság a rétegek között. E jellemzőket a felderítés folyamat használja fel annak becslésére, hogy az épületek hol helyezkednek el nagy valószínűséggel, illetve milyen *lehet* a geometriai megjelenésük: a születés lépés az objektumokat a becsült beépített területeken nagyobb valószínűséggel generálja. Másfelől, az objektum szintű jellemzők egy adott u épületjelöltet írnak le, és felhasználhatók a generált téglalapok fitnessz értékének meghatározásához. Mivel az épületek verifikációja ez utóbbi jellemzőkön alapul, a pontosságuk kritikus. Mivel a hasonlóság leírón kívül a következő jellemzők számítása a két képen függetlenül történik, gyakran nem hangsúlyozzuk ki az aktuális képindexet ebben a fejezetben.



1. ábra: A téglalap geometriai paraméterei

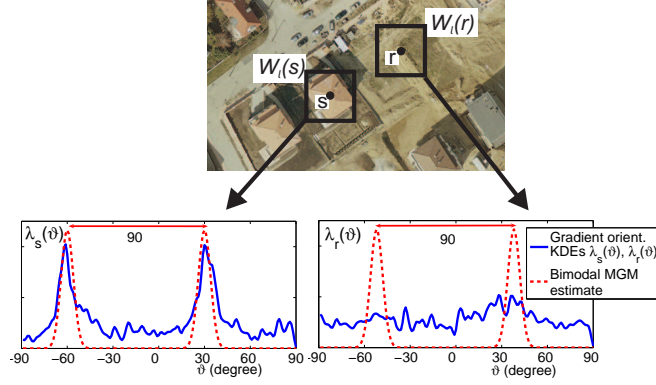
3.1. Alacsony szintű jellemzők a beépítések felismeréséhez

Az első jellemző azt a tényt használja ki, hogy az épületeket tartalmazó képrészletek hangsúlyos és egymásra merőleges éleket tartalmaznak. Ezt a tulajdonságot robusztusan jellemezhetjük lokális gradiens orientációs hisztogramokkal [6]. Jelölje ∇g_s az s pixelhez tartozó intenzitás gradiens vektort, $\|\nabla g_s\|$ abszolút értékkel és ϑ_s szöggel. Legyen $W_l(s)$ a négyzet alakú $l \times l$ méretű ablak s körül, ahol l -t úgy választjuk, hogy $W_l(s)$ szűken lefedjen egy átlagos épületet. Minden s pixel $W_l(s)$ környezetében kiszámítjuk a súlyozott ϑ_s sűrűséget:

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

ahol $N_s = \sum_{r \in W_l(s)} \|\nabla g_r\|$ és h a kernel sávszélessége. A gyors számításhoz uniform kerneleket használtunk. Ha $W_l(s)$ lefed egy épületet, a $\lambda_s(\vartheta)$ függvény két jellemző csúccsal rendelkezik, melyek 90° távolságban találhatóak a ϑ -térben (2. ábra). Ezt a tulajdonságot mérhetjük a $\lambda_s(\vartheta)$ függvény és egy megfelelően illesztett kétmódusú Gaussi keverék függvény korrelációjával:

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



2. ábra: Két választott pixel környezetéhez tartozó lokális gradiens orientációs hisztogramok: s épület középpont, r üres terület része.

ahol $\eta_2(\cdot)$ két Gauss függvény keveréke m illetve $m + 90^\circ$ várható értékkel és azonos d_λ szórással (d_λ a folyamat paramétere). A maximális korrelációhoz tartozó eltolás (m_s) korrelációs érték (α_s) így megkapható a következőképpen:

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

Nagy α_s értékkel rendelkező pixelek nagyobb valószínűséggel lehetnek épület középpontok, amit az α -születés-térképen kódolhatunk: $P_b^\alpha(s) = \alpha_s / \sum_{r \in S} \alpha_r$. Az elnevezés onnan ered, hogy az s pixelben a lokális $P_b(s)$ születési értékkel arányos frekvenciával fogunk generálni objektumokat az eljárás során.

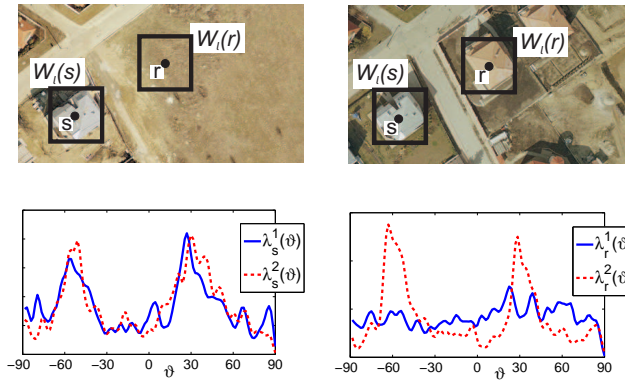
Másrésről az m_s eltolás értéke becslést ad a domináns gradiens irányra $W_l(s)$ -ben. Ezért ha s középponttal generálunk egy u objektumot, az orientációját $\theta(u) = m_s + \eta_s$ -ként modellezzük, ahol η_s zéró várható értékű Gaussi valószínűségi változó kis σ_θ szórással paraméterrel.

Kísérleteinkben megfigyeltük, hogy az α_s -gradiens jellemző általában képes a beépített területek durva becslésére. Ugyanakkor, számos esetben a detekció finomítható további jellemzők figyelembevételével, például tetőszín és árnyékok modellezésével [9]. A tetőszínek egy része jól elkülöníthető megvilágítás invariáns színreprezentációkban, mint a HSV színtér 'hue' csatornája. Tegyük fel, hogy ki tudunk nyerni egy $\mu_c(s) \in \{0, 1\}$ indikátor maszkot, ahol $\mu_c(s) = 1$ tartozik a tetőszíni s pixelhez. Ekkor s színleíróját a következőképpen számoljuk: $\Gamma_s = \sum_{r \in W_l(s)} \mu_c(r)$ és a szín alapú születés térképet szintén meghatározzuk: $P_b^c(s) = \Gamma_s / \sum_{r \in S} \Gamma_r$. Jegyezzük meg, hogy ez a jellemző nyilvánvalóan nem használható szürkeárnyalatú képekre, illetve színes felvételek esetén is általában csak a tetők egy részét azonosítja, melyek például tipikus 'piros' színűek ([9] és 5(b) ábra).

Az épületek jelenlétére vetett árnyékaikon keresztül is tehetünk becslést [5, 9]. Kihasználva, hogy az árnyékok sötétsége és iránya globális képi jellemzők, gyakran kinyerhetünk egy (zajos) $\mu_{sh}(s)$ árnyék maszkot, például a *sötét-kék* színtartományba eső pixelek kiszűrésével [9]. Ezután az épületeket tartalmazó területek azonosíthatók azon

képrészletekként, amelyek az árnyék foltok mellett nap irányban fekszenek (6. ábra). Konstans $P_b^{\text{sh}}(s) = p_0^{\text{sh}}$ születési értéket használtunk a kinyert beépített területeken belül, míg lényegesen kisebb ϵ_0^{sh} konstans kívül.

Mivel a kombinált születéstérkép célja az, hogy valamennyi lehetséges beépített területre ráirányítsa a figyelmet, a jellemző térképekből a maximum operátorral származtatjuk: $P_b(s) = \max \{P_b^\alpha(s), P_b^c(s), P_b^{\text{sh}}(s)\} \forall s \in S$. Olyan képi bemenetekre, melyek nem tartalmaznak szín- vagy árnyékinformációt, egyszerűen el kell hagyni a megfelelő komponenst. Ismét megjegyezzük, hogy a születési és orientációs térképeket mindkét képre külön generáljuk, szükség esetén az időfüggést az alábbi módon jelöljük: $P_b^{(i)}(s), m_s^{(i)}, i \in \{1, 2\}$.



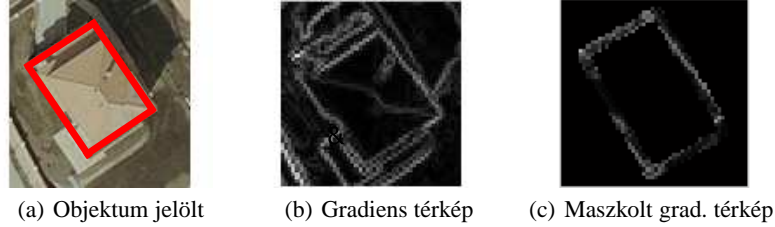
3. ábra: A $\lambda(\cdot)$ függvények összehasonlítása a két különböző idejű képen két választott pixel környezetében. Az s pixel változatlan régió eleme, míg r pixel körül egy új megjelenő épület figyelhető meg.

3.2. Alacsony szintű hasonlóságvizsgálat

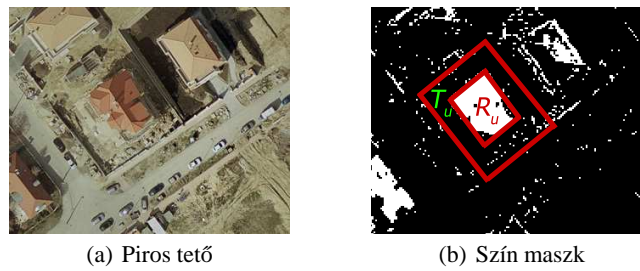
A gradiens orientációs statisztika hatékony eszközt biztosít régiók alacsony szintű összevetésére is. A $\lambda_s^1(\cdot)$ és $\lambda_s^2(\cdot)$ függvények illesztése felfogható az s körüli területek összehasonlításaként, épületspecifikus textúra jellemzők alapján (3. ábra). Mi több, ezek a leírók függetlenek a megvilágítástól és fehéregyensúlytól, valamint robusztusak a parallaxis és regisztrációs hibákra. A texturális különbségeket az eloszlások Bhattacharyya távolságával mérjük:

$$b(s) = -\log \int \sqrt{\lambda_s^1(\vartheta) \cdot \lambda_s^2(\vartheta)} d\vartheta$$

A bináris hasonlóságtérkép B a következő módon számítható: $B(s) = 1$ akkor és csak akkor ha $b(s) < b_0$, egyébként $B(s) = 0$.



4. ábra: A gradiens jellemző bemutatása



5. ábra: A tetőszín jellemző bemutatása

3.3. Objektum szintű jellemzők

Ebben a fejezetben különböző objektum szintű jellemzőket mutatunk be. Felhasználva ezeket, $\varphi^{(i)}(u)$ energia tagokat definiálunk, melyek kiértékelik az u épület-hipotézist az i -edik képen (később ismét elhagyjuk az i indexet). $\varphi(u)$ negatív fitness értéként értelmezhető, míg egy téglalapot $\varphi(u) < 0$ értékkel attraktív objektumnak hívunk. Mivel a populáció bővítése attraktív objektumokkal csökkentheti a globális konfigurációs energiát [4], ezek az alakzatok hatékony épület-jelöltek.

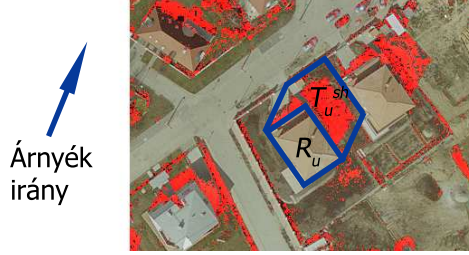
A vizsgálatokat a gradienstérkép analízisével kezdjük ismét. Egy releváns téglalap-jelölt R_u határai alatt elvárhatjuk, hogy a képen a gradiens értékek magasak legyenek, és a gradiens irány legyen merőleges az R_u legközelebbi oldalára (4. ábra). A Λ_u jellemzőt a következőképpen számítjuk:

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

ahol $\tilde{\partial}R_u$ az R_u dilatált él maszkja, $\Theta_u^s \in \{\theta(u), \theta(u) + 90^\circ\}$ az R_u él orientációja $s \in \tilde{\partial}R_u$ körül, és q_u a $\tilde{\partial}R_u$ pixeleinek száma. Az adat-energia tagot az alábbi módon számítjuk: $\varphi_\Lambda(u) = \mathcal{Q}(\Lambda_u, d_\Lambda, D_\Lambda)$ ahol a következő nemlineáris \mathcal{Q} átviteli függvényt használjuk [4]:

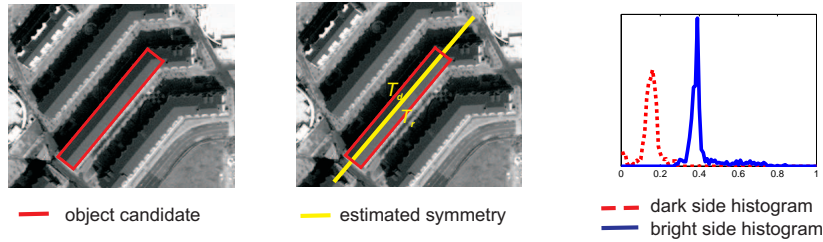
$$\mathcal{Q}(x, d_0, D) = \begin{cases} \left(1 - \frac{x}{d_0}\right) & \text{ha } x < d_0 \\ \exp\left(-\frac{x-d_0}{D}\right) - 1 & \text{ha } x \geq d_0 \end{cases}$$

A *tetőszín* jellemző számítását az 5. ábra szemlélteti. Definiáljuk a T_u objektum környezetét és kiszámítjuk a $\mathcal{C}_R(u) = \frac{1}{\#R_u} \cdot \sum_{s \in R_u} \mu_c(s)$ belső és $\mathcal{C}_o(u) = \frac{1}{\#T_u} \cdot \sum_{s \in T_u} [1 - \mu_c(s)]$ külső kitöltési tényezőket (# pixelekben mért területet jelöl). Végül származtatjuk az energiát: $\varphi_C(u) = \max [\mathcal{Q}(\mathcal{C}_R(u), d_R^c, D_R^c), \mathcal{Q}(\mathcal{C}_o(u), d_o^c, D_o^c)]$



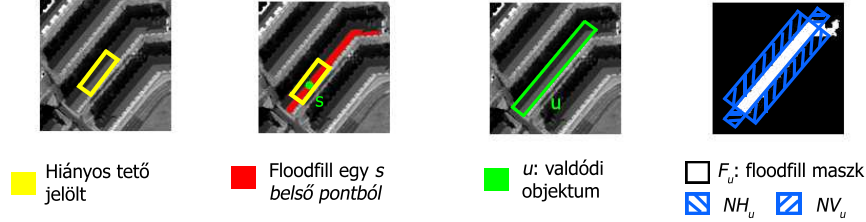
6. ábra: Az árnyék jellemző bemutatása

Az *árnyék tényezőt* a színével analóg módon definiáljuk, de a vizsgált objektum környezetet, T_u^{sh} -t az árnyék irányban helyezzük el (6. ábra). Ezután a belső illetve külső kitöltési tényezőket számítjuk $\chi_R(u) = \frac{1}{\#R_u} \sum_{s \in R_u} [1 - \mu_{\text{sh}}(s)]$ és $\chi_o(u) = \frac{1}{\#T_u^{\text{sh}}} \sum_{s \in T_u^{\text{sh}}} \mu_{\text{sh}}(s)$, míg a $\varphi_\chi(u)$ energiát meghatározása azonos módon történik, mint a $\varphi_C(u)$ esetén. Megjegyezzük, hogy kísérleteinkben a $\varphi_\chi(u)$ tag akkor is robusztusnak bizonyult, amikor az árnyékos területek mérete erősen különböző volt az épületek magasságainak változatosságából adódóan.



7. ábra: A tető homogenitás jellemző bemutatása

Szürkeárnyalatos műhold képeken a tető *homogenitása* szintén hasznos jellemzőt ad. A 7. ábra bemutat egy példát két komponensű homogén tetők leírására. Miután kinyertük az u objektum jelölt fő szimmetria tengelyét, a világos és sötét tetőrészlet histogramjainak csúcosságát a $\kappa_d(u)$ illetve $\kappa_b(u)$ kurtózisokkal jellemezzük. Vegyük észre azonban (8. ábra), hogy a kurtózis maximuma alapján gyakran töredék-tetőkhöz juthatunk (homogén tető része is homogén), ezért a tető teljességét is mérni kell egyidejűleg. Ezért létrehozuk u elárasztás-maszkját (floodfill mask), F_u -t, ami az R_u



8. ábra: Elárasztás alapú tető teljesség jellemző

belsejéből a bemeneti képen elárasztás művelettel elérhető pixeleket tartalmazza. Ha a homogén tető jelölt teljes, F_u alacsony átlapolódási értékkel rendelkezik az R_u horizontális (NH_u) valamint vertikális (NV_u) szomszédossági területeivel (8. ábra). Végül, a $\varphi_\kappa(u)$ energiatagot a kurtózis és teljesség jellemzőkből az előzőekben ismertetett módon hozzuk létre.

A keretrendszer lehetővé teszi a jellemzők rugalmas integrációját a bemeneti képek tulajdonságaitól függően. Minden épület prototípushoz előírhatunk egy vagy több jellemző kényszert, melyek φ energiatagjait a max operátorral csatoljuk össze a prototípus összesített energiatagjában (logikai és művelet a negatív fitnessz értékek terében). Egy adott képpáron több prototípust detektálhatunk párhuzamosan, ha a különböző prototípusok energiatagjait a min (logikai vagy) operátorral csatoljuk. Például, a Budapest képpáron (11. ábra) két prototípus használunk: az első előírja az élekre és az árnyékokra vonatkozó kényszert, míg a másik kizárólag a tető színre vonatkozót. Így a teljes energiatag az alábbi módon számítható:

$$\varphi(u) = \min \{ \max \{ \varphi_A(u), \varphi_X(u) \}, \varphi_C(u) \}.$$

4. Jelölt pontfolyamat modell

Jelöljük \mathcal{H} -val az u objektumok terét. Definiáljunk egy $H \in \mathcal{H}$ Borel halmazt, és az Ω konfigurációs teret az alábbi módon határozzuk meg [4]:

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

Jelöljön ω egy tetszőleges objektum konfigurációt $\{u_1, \dots, u_n\} \in \Omega$. Definiáljuk ezenfelül a \sim szomszédossági relációt \mathcal{H} -ban: $u \sim v$ ha a téglalapjaik, R_u és R_v metszik egymást.

Bevezetünk egy nem stacionárius adatfüggő Gibbs eloszlást a konfigurációs téren: $P_{\mathcal{D}}(\omega) = 1/Z \cdot \exp[-\Phi_{\mathcal{D}}(\omega)]$, ahol Z normalizáló konstans, és

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

$A_{\mathcal{D}}(u)$ és $I(u, v)$ az adatfüggő singleton, illetve a prior interakciós potenciálok és γ súlyozó tényező a két energiatag között. A legvalószínűbb, Maximum Likelihood (ML) konfiguráció $P_{\mathcal{D}}(\omega)$ alapján becsülhető:

$$\omega_{\text{ML}} = \arg \min_{\omega \in \Omega} [\Phi_{\mathcal{D}}(\omega)]$$

A *singleton* potenciálok egy adott $u = \{c_x, c_y, e_L, e_l, \theta, \xi\}$ épület szegmens jelöltet jellemeznek a környezetbeli képi adat függvényében, azonban függetlenül a populáció többi objektumától:

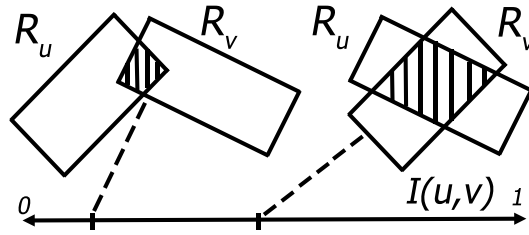
$$A_{\mathcal{D}}(u) = \mathbb{I}_{[\xi(u) \in \{1, *\}]} \cdot \varphi^{(1)}(u) + \mathbb{I}_{[\xi(u) \in \{2, *\}]} \cdot \varphi^{(2)}(u) + \frac{\gamma_{\xi}}{\#R_u} \left\{ \mathbb{I}_{[\xi(u) = *]} \sum_{s \in R_u} (1 - B(s)) + \mathbb{I}_{[\xi(u) \in \{1, 2\}]} \sum_{s \in R_u} B(s) \right\}$$

ahol $\mathbb{I}_{[E]} \in \{0, 1\}$ az E esemény indikátor függvénye, és ahogy korábban definiáltuk, $\varphi^{(1)}(u)$ és $\varphi^{(2)}(u)$ épület illeszkedési energiák az első, illetve második képen (3.3. fejezet). $B(\cdot)$ az alacsony szintű hasonlósági maszk a két időréteg között (3.2. fejezet). Az utolsó tag bünteti a populációban található változatlan jelölt objektumokat ($\xi(u) = *$) melyek nagy textúraváltozást jelző területeken helyezkednek el, illetve újnak/leromboltnak jelzett épületeket ($\xi(u) \in \{1, 2\}$) a *változatlan* területeken.

Az *interakciós* potenciálok prior geometriai kényszereket írnak le: büntetik az azonos időréteget osztó objektum téglalapok nagy mértékű átlapolódását (4. ábra):

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

ahol $\xi(u) \simeq \xi(v)$ fennáll, ha $\xi(u) = \xi(v)$, vagy $\xi(u) = *$, or $\xi(v) = *$.



9. ábra: Interakciós jellemző

5. Optimalizáció

Az optimális objektum konfigurációt a többszörös születés és halál (Multiple Birth and Death) algoritmussal [4] becsüljük, mely a következő lépéseket tartalmazza:

Inicializáció: határozzuk meg a $P_b^{(i)}(s)$ és $m_s^{(i)}$ ($i \in \{1, 2\}$) születési térképeket, és induljunk ki üres populációból $\omega = \emptyset$.

Fő program: állítsuk be a kezdeti inverz hőmérséklet paramétert $\beta = \beta_0$ és a diszkrétizációs lépcsőt $\delta = \delta_0$, ezután alternáljuk a születés és halál lépéseket.

1. *Születés:* sorban vizsgáljuk meg az $s \in S$ pixeleket, és ha nincs s középpontú objektum a jelenlegi ω konfigurációban, vegyünk fel véletlenszerűen egy $\xi \in \{1, 2, *\}$ értéket, legyen $\hat{P}_b = P_b^{(\xi)}(s)$ ha $\xi \in \{1, 2\}$, $\hat{P}_b = \max\{P_b^{(1)}(s), P_b^{(2)}(s)\}$ ha $\xi = *$; és futtassuk a születés folyamatot s -ben $\delta \hat{P}_b$ valószínűséggel. Születés folyamat s -ben: generáljunk egy új u objektumot s középponttal, ξ időindexszel, állítsuk be a $e_L(u)$, $e_r(u)$ oldalhossz értékeket véletlenszerűen előre meghatározott maximális és minimális oldalhosszak között, és állítsuk be a $\theta(u)$ orientációt $\eta(\cdot, m_s^{(\xi)}, \sigma_\theta)$ Gaussi sűrűségfüggvényű véletlen változó alapján, ahogy azt a 3.1. fejezetben leírtuk. Végül, adjuk u -t az aktuális ω konfigurációhoz.
2. *Halál:* tekintsük az aktuális $\omega = \{u_1, \dots, u_n\}$ objektum konfigurációt és rendezzük az objektumokat $A_{\mathcal{D}}(u)$ alapján, csökkenő irányban. Az objektumokat vizsgáljuk meg ebben a sorrendben, és minden egyes u -ra számítsuk ki az alábbi értéket $\Delta\Phi_\omega(u) = \Phi_{\mathcal{D}}(\omega/\{u\}) - \Phi_{\mathcal{D}}(\omega)$, amely az u törlésével létrejövő potenciális energia változást határozza meg. A halálzási arány így:

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

Ezután töröljük u -t ω -ból $d_\omega(u)$ valószínűséggel.

Konvergencia teszt: amíg a folyamat nem konvergál, növeljük a β inverz hőmérsékletet, csökkentünk a diszkrétizációs lépcsőt δ geometriai sémával, és lépünk vissza a születés lépésre. A konvergencia tényét akkor állapíthatjuk meg, ha minden olyan objektum, melyet az aktuális születés lépés generált, és csak azok tűnnek el a következő halál lépésben.

1. táblázat: A tesztadatok fő jellemzői.

Adathalmaz	Típus	Szín	árnyék	Gradiens	Kurtózis
BUDAPEST	légi	igen	igen	jó	részben
BEIJING	QBird	nem	igen	gyenge	részben
SZADA	légi	igen	nem	gyenge	nem
ABIDJAN	Ikonos	nem	nem	éles	igen

6. Kísérletek

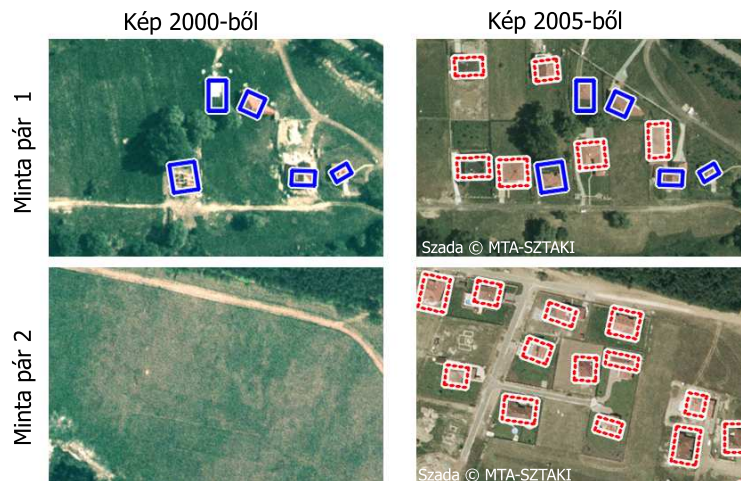
Módszerünket négy jelentősen különböző adathalmazon értékeltük ki¹, melyek fő jellemzőit az 1. táblázat tartalmazza. Kvalitatív eredmények a 10–12 ábrákon láthatók.

¹ A szerzők köszönetet mondanak az adatszolgáltatóknak: Görög András, Budapest, Francia Védelmi Hivatal (DGA) és Liama Laboratórium, Kína

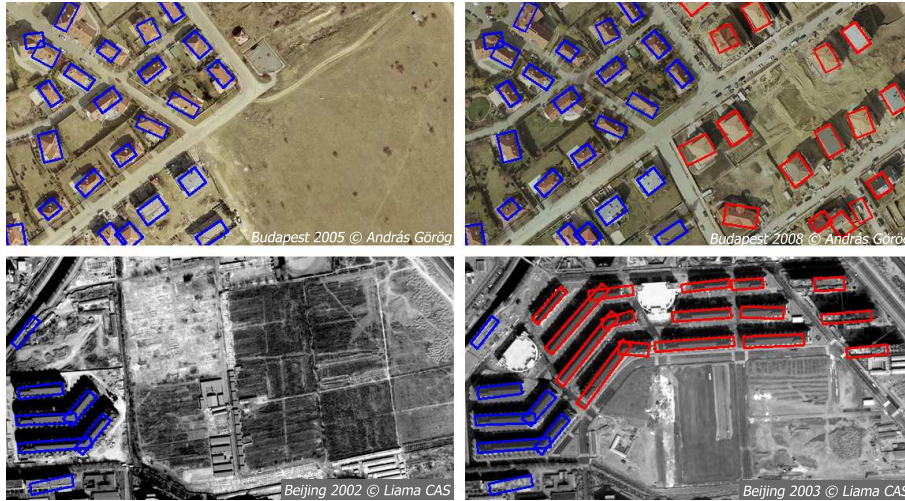
Azért, hogy alátámasszuk az objektumdetekció és a változsfelismerés együttes kezelésének előnyeit, a javasolt modellt (együttes detekció, ED) összehasonlítottuk egy konvencionális megoldással, ahol az épületeket külön nyerve ki az egyes képeken, a változásinformációt pedig utólag becsüljük a kinyert objektumok elhelyezkedésének és geometriai paramétereinek összevetésével (szekvenciális detekció, SD). A 12. ábrán látható, hogy a SD módszer hamis riasztásokat okoz, mivel alacsony kontrasztú területeken egy-egy objektum hiányozhat valamelyik időrétegen, valamint zaj miatt hamis találatok is gyakrabban megjelennek a kevésbé robusztus egynézetű információra való hagyatkozással.

A több jellemzőt együttesen kihasználó épületleíró modell hatékonyságát összehasonlítjuk egy főként éleken alapuló módszerrel (Él Verifikáció, ÉV). Az ÉV módszerben hasonlóan [9]-hez, az árnyék és tetőszín információt csak a beépített területek hozzávetőleges becsléséhez használjuk, míg az objektum verifikáció kizárólag az élkép illesztésén alapul.

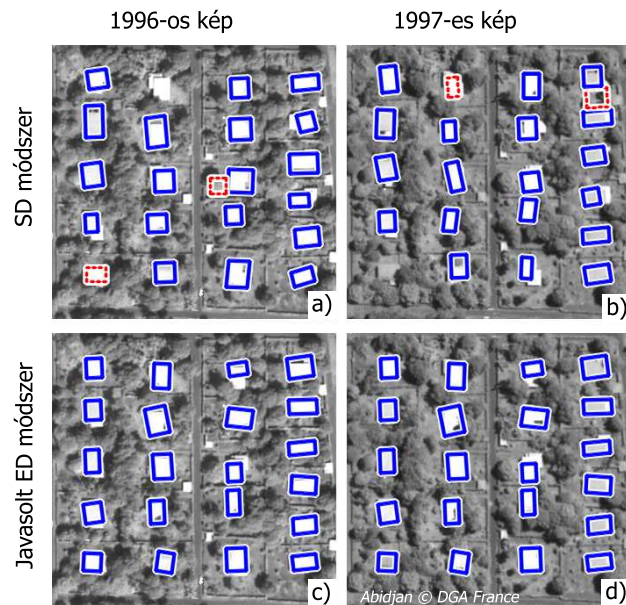
Kvantitatív tesztheinkben a hiányzó (HO) és hamisan (FO) detektált objektumainkat valamint a hiányzó (HV) és hamis (FV) változásokat számoltuk meg, ezenfelül a detekció pixel szintű pontosságát (DP) is megadtuk. A DP számításához összehasonlítottuk az eredményül kapott épület alaprajz maszkjainkat a kézzel készített referencia (*ground truth*) maszkhoz, és meghatároztuk a detekció F-mértékét (a precizitás és visszahívási arány harmonikus közepe). A 2. táblázatban található eredmények megerősítik a javasolt ED modell általánosságát és előnyeit az SD és ÉV módszerekhez képest (alacsonyabb objektum szintű hibák, magasabb DP tényező). A kiértékelés további részletei a szerzők kutatási riportjában olvashatók [1].



10. ábra: Eredmények két képrészleten a SZADA adathalmazból. A kék téglalapok változatlan objektumokat jelölnek, a piros téglalapok változottakat (új, lerombolt vagy átépített).



11. ábra: Az ismertetett módszer eredményei két képpáron. Fenn: BUDAPEST képrészlet (forrás: Görög András ©). Lenn: BEIJING (Liama Laboratórium CAS © Kína). A változatlan (kék) és a megváltozott (piros) épületeket megkülönböztetjük.



12. ábra: Eredmények az ABIDJAN képeken (DGA © Franciaország). Fenn: szekvenciális detekció (SD), valamennyi jelzett változás hamis riasztás. Lenn: az ismertetett együttes detekciós (ED) modell.

2. táblázat: Kvantitatív eredmények. #VT and #VN a változatlan és változott épületek számát jelöli az adott tesztalomban. ED a saját módszer, míg az SD és ÉV referenciamódszerek valamint a HO, FO, HV, FV és DP kiértékelési tényezők a 6. fejezetben kerülnek részletes bemutatásra.

Adathalmaz	#VT	#VN	HO			FO			HV			FV			DP		
			ÉV	SD	ED	ÉV	SD	ED	ÉV	SD	RD	ÉV	SD	ED	ÉV	SD	ED
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

Irodalom

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 7143, INRIA, Sophia Antipolis, December 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. Nicolas Champion, Leena Matikainen, Xinlian Liang, Juha Hyypä, and Franz 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. Sanjiv Kumar and Martial 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.A. 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.