

# Kevert Markov modell alapú változások detekció nagy időkülönbséggel készült légi fotókon

Benedek Csaba<sup>1,2</sup> és Szirányi Tamás<sup>1</sup>

<sup>1</sup> Elosztott Események Elemzése Kutatócsoport, MTA Számítástechnikai és Automatizálási Kutatóintézet

H-1111, Budapest, Kende utca 13-17, Hungary

<http://web.eee.sztaki.hu>

<sup>2</sup> Ariana Project-Team INRIA/CNRS/UNSA

B.P. 93, 06902 Sophia Antipolis, France

<http://www-sop.inria.fr/ariana>

**Kivonat** A cikk nagy időkülönbséggel készült, előzetesen regisztrált légi felvételek összehasonlításával foglalkozik. A bemutatásra kerülő többrétegű kevert Markov modell alapú módszer célja a jelentős változáson átesett területek detekciója. A javasolt eljárás a képek szürkeségi értékeinek globális statisztikáját ötvözi képrészek korrelációjával és helyi kontraszt jellemzőivel. A végső változásmaszkot létrehozó energiaoptimalizációs eljárás párhuzamosan biztosít hatékony lokális jellemző-választást, valamint sima és a megfigyelt képi értékeknek megfelelő szegmentálást. A módszer hatékonyságát valódi légi képeken szemléltetjük.

## 1. Bevezetés

<sup>3</sup> Napjainkban fontos kutatási terület a légi fotókat tartalmazó adatbázisok automatikus kiértékelése, mivel a gyűjtemények nagy mérete és dinamikusan változó tartalma miatt a kézi feldolgozás gyakran túl drága és időigényes. Cikkünk olyan légi képek összehasonlításával foglalkozik, melyek több éves időkülönbséggel készültek, részben különböző évszakokban és eltérő megvilágítási körülmények között. Ezekben az esetekben az egyszerű képkivonás alapú módszerek [2,3] nem alkalmazhatók, mivel a két kép összetartozó pixeleinek színértékei a változatlan területeken belül is jelentősen különbözők lehetnek. Fontos megkötés, hogy a továbbiakban csak optikai képekre fókuszálunk eltérően a radar kép (SAR) alapú módszerektől (pl [4]), melyek könnyítésként kihasználhatják a felvételek megvilágítás-invarianciáját. Csupán azzal a feltételezéssel élünk, hogy a képi adatbázis előzetesen regisztrált ortofotókat tartalmaz.

A legtöbb korábbi módszer kizárólag lakatlan [5] vagy városi [6] területeket vizsgált, illetve egy-egy szűk feladatkörre összpontosított, mint új beépítések észlelése [7] vagy földrengések okozta károk felmérése [8]. Azonban ahogy láthatjuk az 1. és 3. ábrán, a fotóinkon megjelennek mind beépített mind lakatlan régiók, köztük erdők, mezők és mezőgazdasági területek, melyeken számos jellegükben különböző változás figyelhető meg. Célunk azoknak a változásoknak a szűrése, melyek statisztikailag szokatlanok.

<sup>3</sup> A cikk eredetileg angolul az ICPR 2008 konferencia kiadványában jelent meg [1].

A [6]-ban bemutatott független komponens analízis alapú (PCA) változásdetekciós modell szerzői feltételezik, hogy a lényegtelen különbségek forrása kizárólag a megvilágítás és a kamera beállítások eltérése. Mivel ezek a hatások additív vagy multiplikatív módon befolyásolják a képpontok értékeit, az összetartozó pixel-színek kapcsolatát az egész képen konstans lineáris transzformációval modellezték. Hasonló megközelítés olvasható [3]-ban is. Ugyanakkor fontos észrevenni, hogy több esetben a helyszín is változhat szabályszerűen - például az egyes évszakokra jellemző vegetáció különbségei miatt - és ezt a PCA-módszerek figyelmen kívül hagyják. Ugyanígy a vetésforgót használó mezőgazdasági területeken a szomszédos szántóföldek alakja és elrendeződése változhat jelentős mértékben. Cikkünkben közölt módszer példát mutat arra, hogy ezeknek a változásoknak a szabályszerűsége is mérhető statisztikai alapon, bár a pixel-párok mért értékei jelentős eltérést mutatnak a lineáris modelltől.

A továbbiakban ismertetünk egy többretegű kevert Markov modell [9] alapú megközelítést a fenti változásdetekciós feladatra. Megmutatjuk, hogy a globális intenzitásstatisztika és a lokális korreláció egymást hatékonyan kiegészítő jellemző párt alkot a különbségek jelzésére. Az egyes képrégiókon belül kontraszt alapú folyamat választja ki megbízhatóbbnak becsült jellemzőt, a sima változásterképet pedig szomszédos pixelek közötti kölcsönhatásokkal biztosítjuk.

## 2. Képi reprezentáció és jellemzőkinyerés

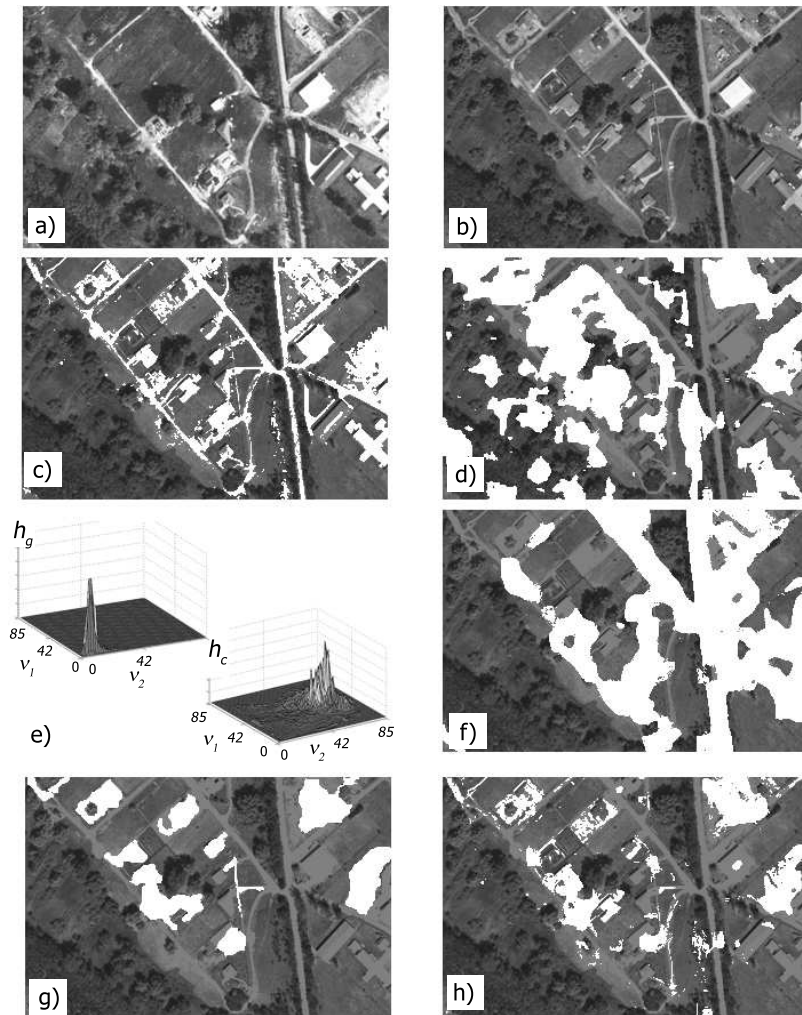
Legyen  $G_1$  és  $G_2$  az összehasonlítás előtt álló két regisztrált légi kép, közös  $S$  pixelráccsal. Jelöljük az  $s \in S$  képpont szürkességi (intenzitás) értékét  $g_1(s)$ -sel a  $G_1$  képen és  $g_2(s)$ -sel  $G_2$ -n. Ahhoz, hogy egy adott  $s \in S$  pixelt a változás (vs) vagy a háttér (hr) [azaz változatlan] osztályba soroljuk, mindenekelőtt releváns lokális jellemzőket kell kinyernünk a képeken. Valószínűségi megközelítést követve a vs/hr osztályokat véletlen folyamatoknak fogjuk tekinteni, melyek a megfigyelt jellemzőket különböző eloszlások alapján generálják.

Vizsgálatainkat a két kép együttes szürkességi terében kezdjük, ám ahelyett, hogy a háttér régiókra a hagyományos globális lineáris transzformációt íránk elő  $g_1(s)$  és  $g_2(s)$  között [6], a megfigyelt intenzitáspárok többmódusú leírását adjuk meg. A tanító háttér területeken észlelt  $\bar{g}(s) = [g_1(s), g_2(s)]^T$  vektorok két dimenziós (2-D) hisztogramját Gaussi sűrűségfüggvények súlyozott összegével közelítjük, amivel leírjuk, mely szürkességi értékek fordulnak elő együtt gyakran a képeken. Ezek után a  $\bar{g}(s)$  vektor háttérben történő megfigyelésének valószínűsége:

$$P(\bar{g}(s)|hr) = \sum_{i=1}^K \kappa_i \cdot \eta(\bar{g}(s), \bar{\mu}_i, \Sigma_i),$$

ahol  $\eta(\cdot)$  2-D Gaussi sűrűségfüggvény  $\bar{\mu}_i$  várható érték vektorral és  $\Sigma_i$  kovariancia mátrixszal, míg  $\kappa_i$  -k pozitív súlytényezők. Fix  $K = 5$ -öt használva az eloszlás paraméterei automatikusan becsülhetők az EM algoritmussal. Másrésztől feltételezzük, hogy tetszőleges  $\bar{g}(s)$  érték előfordulhat a változott területeken, ezért a 'vs' osztályt egyetlen eloszlással modellezzük:  $P(\bar{g}(s)|vs) = u$ .

A fenti intenzitások Gaussi keveréke (IGK) alapú megközelítés azonban sok helyen nem ad megfelelő eredményt. Tekintsük a 1(c) ábrán látható *maximum likelihood* (ML)



1. ábra. Jellemző választás: a)  $G_1$  kép, b)  $G_2$  kép, c) intenzitás alapú változásdetekció, d) korreláció alapú változásdetekció, e)  $h_g$  és  $h_c$  hisztogramok f) lokális kontraszt alapú szegmentáció, g) kézi referencia változásmaszk h) kinyert változásmaszk a jellemzők kombinált használatával, de térbeli simítás nélkül

alapon szegmentált képet, ahol  $s$  pixel címkéje  $\phi_g(s) = \operatorname{argmax}_{\psi \in \{vs, hr\}} P(\bar{g}(s)|\psi)$ . Összehasonlítva a referenciának használt kézi szegmentálással [1(g) ábra] megfigyelhető számos hamis „változás riasztás” is. Ugyanakkor a hibásan klasszifikált képrészek főként az erősen textúrált területekre korlátozódnak (pl építmények és utak), mivel az ott kinyert  $\bar{g}(s)$  értékek nem jelentkeznek dominánsan a képre vonatkozó globális intenzitás statisztikában.

A második jellemzőt,  $c(s)$ -t, az  $s$  pixel  $G_1$  illetve  $G_2$  képeken megfigyelt  $v \times v$  méretű négyzetes környezetének korrelációjaként definiáljuk (kísérleteinkben  $v = 17$ -et használtunk). Nagyobb  $c(s)$  értékkel rendelkező pixelek nyilván nagyobb valószínűséggel tartoznak a háttérhez, és kísérleteink azt mutatták, hogy a  $P(c(s)|hr)$  és  $P(c(s)|vs)$  valószínűségek jól közelíthetők Gaussi sűrűségfüggvényekkel. Magában a  $c(s)$ -alapú ML osztályozás is gyenge eredményt ad, amit az 1(d) ábra mutat ( $\phi_c$  térkép). Azonban megfigyelhetjük, hogy  $\bar{g}(s)$  és  $c(s)$  egymást hatékonyan kiegészítő jellemzők. Alacsony kontrasztú területeken ahol a zajos  $c(s)$  megtévesztő, a  $\bar{g}(s)$  alapú döntés általában megbízható. A textúrált képrészeket viszont célszerű  $c(s)$  szavazatát figyelembe venni  $\bar{g}(s)$  helyett.

A *kontraszt alapú jellemzőválasztás* elvét valószínűségi módon fogalmazzuk meg. Legyen  $\nu_i(s)$  az  $s$  pixel körül mért lokális kontraszt a  $G_i$  képen ( $i = 1, 2$ ), melyet a környező intenzitásértékek varianciájával írunk le. Szintén használni fogjuk az együttes kontraszt-vektort:  $\bar{\nu}(s) = [\nu_1(s), \nu_2(s)]^T$ . Ezután néhány szegmentált tanítókép segítségével származtatjuk a  $h_g$ -vel jelölt 2-D hisztogramot, ami a  $\bar{g}(s)$  alapú döntés statisztikai megbízhatóságát méri  $\bar{\nu}(s)$  függvényében. Hasonlóan definiáljuk  $h_c$ -t a  $c(s)$  jellemzőhöz. Az 1(e) ábra alapján a normalizált  $h_g$  és  $h_c$  hisztogramok hatékonyan közelíthetők 2-D Gaussi sűrűségfüggvényekkel:

$$P(\bar{\nu}(s)|h_\psi) = \eta(\bar{\nu}(s), \bar{\mu}_\psi, \Sigma_\psi), \quad \psi \in \{g, c\}.$$

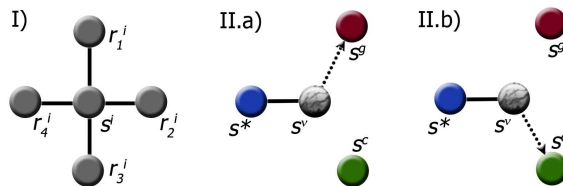
Így elkészíthető a kontraszt térkép:  $\phi_\nu(s) = \operatorname{argmax}_{\psi \in \{g, c\}} P(\bar{\nu}(s)|h_\psi)$  [1(f) ábra: ‘c’ osztály fehér], valamint a kombinált szegmentációs térkép  $\phi_*$  [1(h) ábra], ahol:

$$\phi_*(s) = \begin{cases} \phi_g(s) & \text{ha } \phi_\nu(s) = g \\ \phi_c(s) & \text{ha } \phi_\nu(s) = c \end{cases}$$

Bár látható, hogy  $\phi_*$  sokkal jobb közelítése a referenciamaszknak, mint az egyes jellemzők önálló térképei, még meglehetősen zajos. Ezért a fenti eredményekre alapozva bemutatunk egy robusztus szegmentációs eljárást a következő fejezetben.

### 3. Kevert Markovi képszegmentációs modell

A kevert Markov modellek [9] a hagyományos Markov véletlen mezők egy lehetséges kiterjesztését valósítják meg: lehetővé teszik mind statikus prior, mind dinamikus adatfüggő kapcsolatok alkalmazását a feldolgozó csomópontok között. Esetünkben hasznos kihasználni ezt a tulajdonságot, mivel a  $\bar{\nu}(s)$  jellemző szerepe különleges: lokálisan ki- és bekapcsolhatja a  $\bar{g}(s)$  illetve  $c(s)$  jellemzőket a szegmentációs folyamatba. Tekintsük a feladatot négy interaktív képszegmentáció együttesének, hasonlóan az 1(c), (d), (f) and (h) ábrákon látottakhoz. A problémát ezért egy négyrétegű  $\mathcal{G}$  gráfra vetítjük, ahol a



2. ábra. (I) rétegen belüli és (II.a,II.b) kívüli kapcsolatok  $\mathcal{G}$ -ben (az éleket folytonos, a cím-mutatókat szaggatott vonalak jelzik).

rétegeket  $S^g$ ,  $S^c$ ,  $S^\nu$  és  $S^*$  jelölik. Minden  $s \in S$  pixelhez kölcsönösen egyértelműen hozzárendelünk egy csomópontot valamennyi rétegeken: pl  $s^g$  csomópont tartozik  $s$ -hez az  $S^g$ -n. Hasonlóan jelöljük ezt a többi rétegen is:  $s^c \in S^c$ ,  $s^\nu \in S^\nu$  és  $s^* \in S^*$ .

Bevezetünk egy véletlen címkéző folyamatot, amely  $\omega(\cdot)$  jelű címkét rendel  $\mathcal{G}$  valamennyi csomópontjához. A gráf élei – a Markov mezős szegmentációk során ismert módon – fejezik ki a kölcsönhatásokat a kapcsolódó csomópontok címkéi között. Módszerünk itt használja fel, hogy a kevert Markov modellek kétféle feldolgozó egységet különböztetnek meg: *reguláris* és *cím* csomópontokat [9]. Esetünkben az  $S^g$ ,  $S^c$ , és  $S^*$  rétegek *reguláris* egységeket tartalmaznak, ahol a címke egy lehetséges vs/hr osztályt jelöl:  $\forall s \in S, i \in \{g, c, *\} : \omega(s^i) \in \{vs, hr\}$ . Az  $\omega(s^g)$  címke  $s$  pixel  $\bar{g}(s)$  alapú, míg  $\omega(s^c)$  a  $c(s)$  alapú szegmentációjához tartozik. Az  $S^*$  réteg címkéi adják a végső változástérképet. Ugyanakkor az  $S^\nu$  réteg elemei *cím* csomópontok, ahol  $s^\nu \in S^\nu$  esetén a címke  $\omega(s^\nu)$  valójában egy mutató a  $\mathcal{G}$  gráf valamelyik reguláris csomópontjára. Ellentétben a statikus élekkel, a cím mutatók adatfüggő összeköttetéseket létesítenek a csomópontok között.

Bevezetjük az alábbi jelöléseket:  $\tilde{\omega}(s^\nu) := \omega(\omega(s^\nu))$  legyen annak a (*reguláris*) csomópontnak a címkéje, amit  $s^\nu$  címez meg, és  $\underline{\omega} = \{\omega(s^i) | s \in S, i \in \{g, c, \nu, *\}\}$  jelöljön egy globális címkézést. Legyen  $\mathcal{F} = \{\mathcal{F}_s | s \in S\}$  a globális megfigyelés, ahol  $\mathcal{F}_s = \{\bar{g}(s), \bar{\nu}(s), c(s)\}$  a lokális jellemzők halmaza.

A kevert Markov modellek definíciója alapján [9] a statikus élek bármely két csomópontot összeköthetnek, és egy adott  $\underline{\omega}$  globális címkézés posterior valószínűsége az alábbi módon számítható:

$$P(\underline{\omega} | \mathcal{F}) = \alpha \prod_{C \in \mathcal{C}} \exp \left( -V_C(\omega_C, \omega_C^\nu, \mathcal{F}) \right), \quad (1)$$

ahol  $\mathcal{C}$   $\mathcal{G}$  klikkjeinek a halmaza, és tetszőleges  $C \in \mathcal{C}$  klikk esetén:  $\omega_C = \{\omega(q) | q \in C\}$  és  $\omega_C^\nu = \{\tilde{\omega}(s^\nu) | s^\nu \in S^\nu \cap C\}$ .  $V_C$  a  $\mathcal{C} \rightarrow \mathbb{R}$  *klikk potenciál*, ami „alacsony” értéket vesz fel, ha a  $\omega_C \cup \omega_C^\nu$  halmazon belüli címkék konzisztensek a címkézési szabályokkal, egyébként „magas” értéket kell mérnünk.

A következőkben meghatározzuk  $\mathcal{G}$  klikkjeit és a hozzájuk tartozó  $V_C$  klikk potenciál függvényeket. Az  $\mathcal{F}_s$  megfigyelés úgynevezett szingleton potenciálokon – egyes csomópontokon értelmezett energiatagodon – keresztül befolyásolja a modellt. Ahogy korábban leszögeztük, az  $S^g$  és  $S^c$  rétegek címkéi közvetlen kapcsolatban állnak a  $\bar{g}(\cdot)$  illetve  $c(\cdot)$  jellemzőkkel, míg a címkék az  $S^*$  rétegen nincsenek közvetlen kapcsolatban a

mért adatokkal. Ezért az alábbi definíciókat használjuk:  $V_{\{s^g\}} = -\log P(\bar{g}(s)|\omega(s^g))$ ,  $V_{\{s^c\}} = -\log P(c(s)|\omega(s^c))$  és  $V_{\{s^*\}} \equiv 0$ . Az előbbi sűrűségfüggvények megegyeznek a 2 fejezetben definiáltakkal és  $V_{\{s^\nu\}}$ -t később adjuk meg.

A szegmentációk simaságát úgy biztosítjuk, hogy az egyes rétegeken belül éleket húzunk az  $S$  pixelrács szomszédos képpontjaihoz tartozó csomópontok közé (2-I ábra). Jelöljük az így nyert „rétegen belüli klikkek” halmazát  $\mathcal{C}_2$ -vel. Egy adott  $\mathcal{C}_2$ -beli klikk potenciál függvénye bünteti ha az így szomszédos csomópontok címkéi különböznek. Legyenek  $r$  és  $s$  az  $S$  rács szomszédos pixelei, ekkor a  $\mathcal{C}_2 = \{r^i, s^i\} \in \mathcal{C}_2$   $i \in \{g, c, \nu, *\}$  klikk potenciálja egy  $\varphi^i > 0$  konstans segítségével számítható:

$$V_{\mathcal{C}_2}(\omega(s^i), \omega(r^i)) = \begin{cases} -\varphi^i & \text{ha } \omega(s^i) = \omega(r^i) \\ +\varphi^i & \text{ha } \omega(s^i) \neq \omega(r^i) \end{cases}$$

A leírást a rétegek közötti interakciók bemutatásával folytatjuk. Korábbi vizsgálataink alapján  $\omega(s^*)$  általában  $\omega(s^g)$  vagy  $\omega(s^c)$  értékét veszi fel, a megfigyelt  $\nu(s)$  jellemző függvényében. Ezért élet húzunk  $s^*$  és  $s^\nu$  közé, valamint előírjuk, hogy az  $s^\nu$  mutató az  $s^g$  vagy  $s^c$  csomópontot címezze meg (2-II.a és II.b ábrák). Így az  $S^\nu$  réteg singleton potenciáljai meghatározhatók: ha  $s^\nu$  az  $s^\psi|_{\psi \in \{g, c\}}$  pontba mutat, legyen  $V_{\{s^\nu\}} = -\log P(\bar{\nu}(s)|h_\psi)$ . Ezek után egy tetszőleges „rétegek közötti klikk”  $\mathcal{C}_3 = \{s^*, s^\nu\}$  potenciálját fix  $\rho > 0$ -t használva az alábbi módon számíthatjuk

$$V_{\mathcal{C}_3}(\omega(s^*), \tilde{\omega}(s^\nu)) = \begin{cases} -\rho & \text{ha } \omega(s^*) = \tilde{\omega}(s^\nu) \\ +\rho & \text{különben} \end{cases}$$

Végül (1) alapján az  $\hat{\omega}$  optimális címkézésre maximum a posteriori becslést adunk, mely maximalizálja  $P(\hat{\omega}|\mathcal{F})$ -t (így minimalizálja  $-\log P(\hat{\omega}|\mathcal{F})$ -t) és a következőképpen kapható meg:

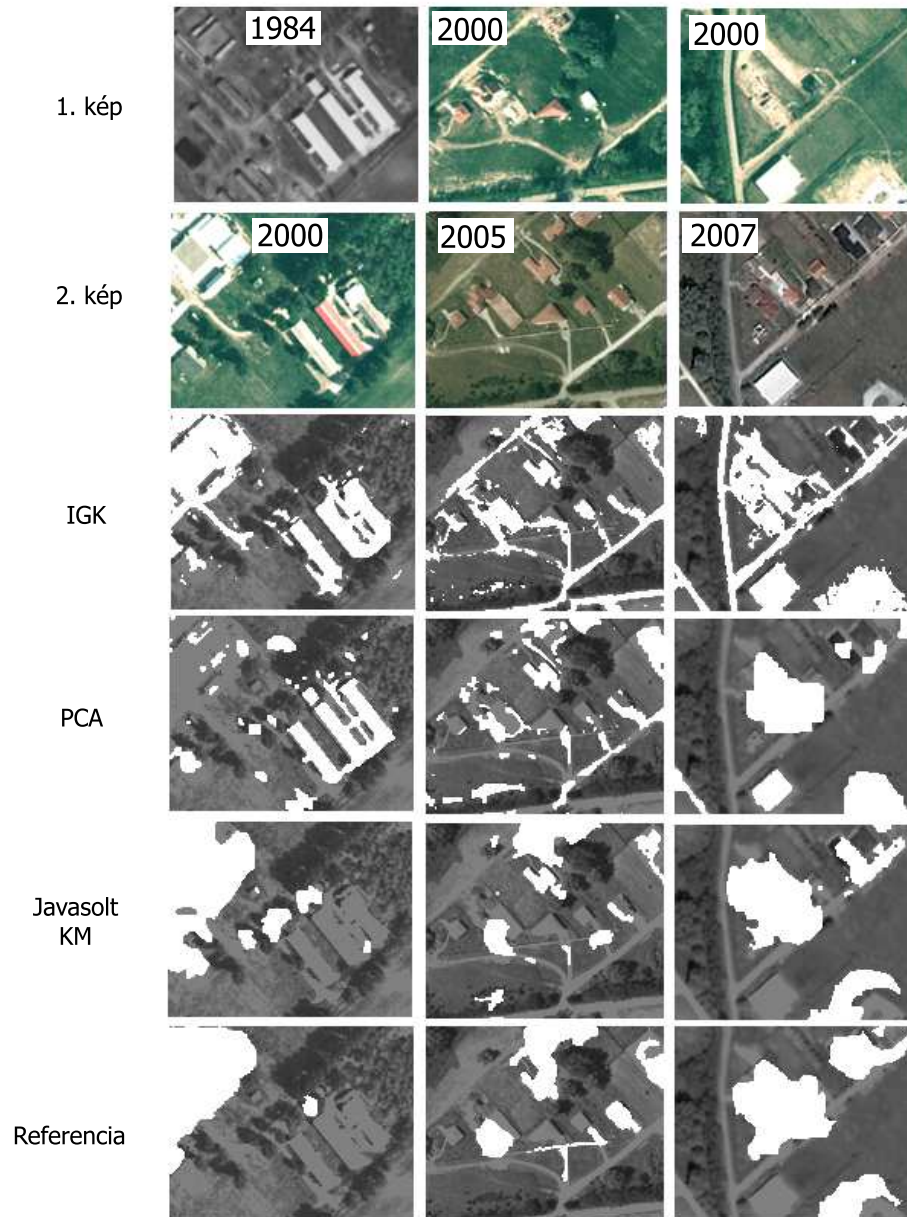
$$\begin{aligned} \hat{\omega} = \operatorname{argmin}_{\omega \in \Omega} & \sum_{s \in S; i} V_{\{s^i\}}(\omega(s^i), \mathcal{F}_s) + \\ & + \sum_{\{s, r\} \in \mathcal{C}_2; i} V_{\mathcal{C}_2}(\omega(s^i), \omega(r^i)) + \sum_{s \in S} V_{\mathcal{C}_3}(\omega(s^*), \tilde{\omega}(s^\nu)) \end{aligned} \quad (2)$$

ahol  $i \in \{g, c, \nu, *\}$  és  $\Omega$  valamennyi lehetséges címkézés halmazát jelöli. Az eljárás kimenete az  $S^*$  réteg címkézésével azonos.

#### 4. Kísérletek

A módszer kiértékeléséhez különböző 1.5m/pixel felbontású légi fotó párokat használtunk, melyekhez referenciának kézzel is generáltunk változásmaszkokat. A felvételeket részben a Földmérési és Távérzékelési Intézetől (FÖMI) vásároltuk, részben a Google Earth képein dolgoztunk. A modellparamétereket felügyelt módszerekkel tanítottuk és az eljárást külön teszt adatbázison validáltuk. Az (2) energiatag minimumának jó közelítését a ‘Modified Metropolis’ optimalizációs eljárással származtattuk hasonlóan korábbi munkáinkhoz [2,7].

A kidolgozott kevert Markov modell (KM) alapú eljárást összehasonlítottuk az intenzitások Gaussi keveréke (IGK) módszerrel (részletezve 2 fejezetben) és a korábban



3. ábra. Detektált változástérkép (fehér régiók) három mintakép-páron az IGK, PCA [6] és a javasolt kevert Markov model (KM) módszerekkel, valamint a kézi referencia maszk

ismertett PCA alapú modellel [6]. A szegmentálás minőségének kvantitatív jellemzéséhez az  $F$  mértéket használtuk, ami a változásdetektáció pontosságának és detekciós rátájának a harmonikus átlagaként számítható. Kvantitatív eredmények a 1 táblázatban szerepelnek, míg az 3 ábra kvalitatívan mutatja három választott képpár változásterképét a különböző módszerekkel. Az eredmények modellünk előnyeit mutatják.

	IGK	PCA	Jav. KM
$F$ -mérték	0.478	0.605	0.844

1. táblázat. Kiértékelés – kvantitatív eredmények

## 5. Konklúzió

Cikkünk szokatlan változások automatikus szűrésével foglalkozott jelentős időkülönbséggel készített légi képpárokon. Bevezettünk egy új kevert Markov modell alapú megoldást, ami három különböző jellemző robosztus integrációját végzi. A módszer hatékonyságát valódi légi felvételeken teszteltük, valamint kvalitatív és kvantitatív összehasonlítást végeztünk két korábbi módszerrel.

## Hivatkozások

1. Benedek, Cs., Szirányi, T.: A Mixed Markov model for change detection in aerial photos with large time differences. In: Proc. International Conference on Pattern Recognition, Tampa, Florida, USA (2008) In press.
2. Benedek, Cs., Szirányi, T., Kato, Z., Zerubia, J.: A multi-layer MRF model for object-motion detection in unregistered airborne image-pairs. In: Proc. IEEE International Conference on Image Processing. (2007)
3. Radke, R.J., Andra, S., Al-Kofahi, O., Roysam, B.: Image change detection algorithms: A systematic survey. IEEE Trans. IP **14**(3) (2005) 294–307
4. Gamba, P., Dell'Acqua, F., Lisini, G.: Change detection of multitemporal SAR data in urban areas combining feature-based and pixel-based techniques. IEEE Trans. GRS **44**(10) (2006) 2820–2827
5. Perrin, G., Descombes, X., Zerubia, J.: 2D and 3D vegetation resource parameters assessment using marked point processes. In: Proc. ICPR, Hong-Kong (2006)
6. Wiemker, R.: An iterative spectral-spatial bayesian labeling approach for unsupervised robust change detection on remotely sensed multispectral imagery. In: Proc. CAIP. Volume LNCS 1296. (1997) 263–270
7. Benedek, Cs., Szirányi, T.: Markovian framework for structural change detection with application on detecting built-in changes in airborne images. In: Proc. SPPRA. (2007)
8. Kosugi, Y., Sakamoto, M., Fukunishi, M., Wei Lu Doihara, T., Kakumoto, S.: Urban change detection related to earthquakes using an adaptive nonlinear mapping of high-resolution images. IEEE GRSL **1**(3) (2004) 152–156
9. Fridman, A.: Mixed Markov models. Proc. National Academy of Sciences of USA **100**(14) (July 2003) 8092–8096