

Kör alakú objektumok szegmentálása Markov mező segítségével

Blaskovics Tamás¹, Kató Zoltán¹ és Ian Jermyn²

¹ Szegedi Tudományegyetem, Informatikai Tanszékcsoport, Szeged, Magyarország

² Ariana (joint INRIA/I3S research group), INRIA, B.P. 93, Sophia Antipolis, France

Kivonat Mára igen népszerűek lettek a különböző energia-minimalizáláson alapuló képszegmentáló módszerek. Jelen cikkben egy olyan Markov mezős modellt dolgozunk ki, mellyel képesek vagyunk tetszőleges számú meghatározott sugarú kör alakú objektum detektálására. Ehhez egy korábbi phase field modellt, a „gas of circles” modellt vesszük alapul, melyet hasonló problémára dolgoztak ki. Megmutatjuk a két modell ekvivalenciáját, vizsgáljuk a Markov mezős modell előnyeit, lehetőségeit és kitérünk a modell egy lehetséges gyakorlati alkalmazására: fakoronák detektálására. A korábban publikált Markov mezős modellekkel szemben, melyek többnyire csak valamilyen homogenitást erőltettek, modellünk a kialakult régiók alakjára és méretére is ad megköztéseket, ami egy újszerű megközelítés.

1. Bevezetés

Amikor megfigyelünk egy képet nem csak a kép intenzitásértékeit, színeit, textúráit nézzük és döntjük el mit ábrázol a kép, hanem jelentős mennyiségű való világból vett tapasztalattal is rendelkezünk, ami megkönnyíti a feladatot. Egy tárgy felismerésénél a tárgy alakját és azt is nézzük, milyen környezetben van. A korai képszegmentáló módszerek még csak a képek intenzitását használták, és így nem lehettek hatékonyak például erősen zajos képeken. Manapság a kutatások afelé haladnak, hogy a képen található információk (intenzitás, szín, textúra, stb.) mellett használjunk olyan előzetes információkat is, mint például a keresett objektum alakja.

A korábbi aktív kontúros modellek egy előtér-háttér szegmentálásra koncentrálnak, adatmodellnek a kép gradiensét használják és egy előre definiált alak modellt. A kontúr mozgatását a gradiens nagysága és a kontúr alakja határozza meg. Lehetséges lenne több információt is felhasználni az adatmodellhez, de korántsem triviális például, hogy egy aktív kontúr esetén hogyan használjuk fel a kép textúráját. A másik oldalon pedig ott van a Markov modell, mely egy valószínűségi modell és könnyen kezeli a kép többi jellemzőit is, viszont nagyon kevés munka irányult arra, hogy előzetes alak információt építsenek be ilyen modellekbe.

Elsődleges célunk tehát egy olyan Markov mezős modell kifejlesztése, mely képes előzetes alak információkat felhasználni, és kipróbáljuk a modellt különböző képfeldolgozási problémákon.

1.1. A „gas of circles” phase field modell

Az [2] dolgozatban a magasabb rendű phase field modell megkonstruálásához egy már korábban definiált magasabb rendű aktív kontúr modellt [1] (HOAC) vettek alapul. Továbbá levezettek egy stabilitási analízist a HOAC energia paramétereire, mely egy lokális minimumot definiál a kontúr energiáján egy adott sugarú kör elérése esetén. A stabilitás jelen esetben azt jelenti, hogy az energia-minimalizálás során a kontúr a kör alakzat kicsiny változásai esetén visszaalakul körre.

Magasabb-rendű phase field modell. A ϕ phase field függvény egy Ω -tartományon értelmezett level set függvény. Egy adott z küszöbre ϕ definiál egy régiót (nem feltétlenül összefüggőt) az összes lehetséges régiók R halmaza közül a következő leképezés szerint: $\zeta_z(\phi) \in R = \{x \in \Omega : \phi(x) > z\}$. A phase field energia a következőképp írható le:

$$E_0(\phi) = \int_{\Omega} d^2x \left\{ \frac{D_f}{2} |\nabla \phi|^2 + \alpha_f \left(\phi - \frac{\phi^3}{3} \right) + \lambda_f \left(\frac{\phi^4}{4} - \frac{\phi^2}{2} \right) \right\} \quad (1)$$

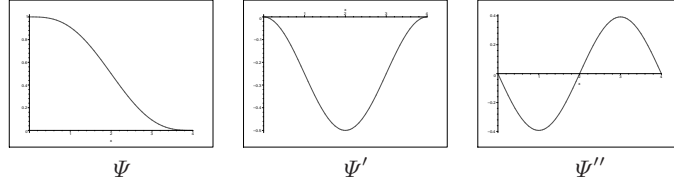
A phase field energia két tagból, az E_0 energiából és egy úgynevezett nem lokális energiából áll: E_{NL} (non-local), ami a HOAC modellben megadott kvadratikusan felel meg, és a következőképp van definiálva:

$$E_{NL}(\phi) = -\frac{\beta_f}{2} \int_{\Omega^2} \nabla \phi(x) \cdot \Psi(|x - x'|) \mathbf{I}_{2 \times 2} \cdot \nabla \phi(x') d^2x d^2x' \quad (2)$$

ahol Ψ az interakciós függvény a következő (1. ábra):

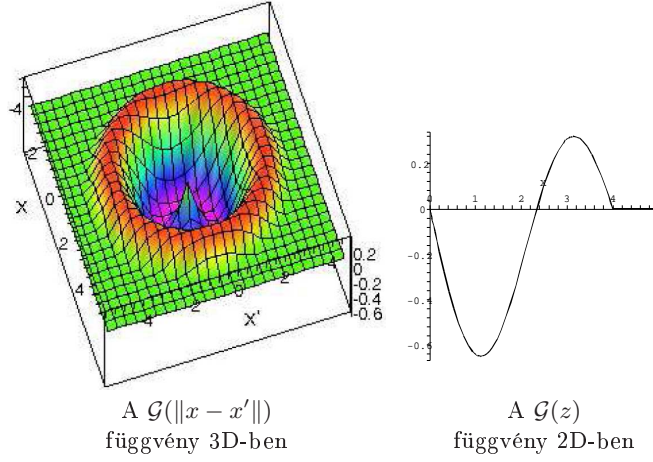
$$\Psi(z) = \begin{cases} \frac{1}{2} \left(2 - \frac{z}{d} - \frac{1}{\pi} \sin\left(\frac{\pi z}{d}\right) \right) & \text{ha } |z| < 2d, \\ 1 - H(z - d) & \text{egyébként.} \end{cases} \quad (3)$$

ahol d az interakció hatósugara és H a Heaviside függvény.



1. ábra. A Ψ interakciós függvény, és annak első és második deriváltja. ($d = 2$)

A „gas of circles” phase field modell. A HOAC modellben levezetett stabilitási analízis megadja a paraméterek helyes beállítását ahhoz, hogy az energiaminimalizálás folyamán adott sugarú, stabil körök alakuljanak ki. Ugyanakkor nem kell egy újabb stabilitási analízist végezni a phase field modellben (és később a Markov mezőben sem), mert létezik a HOAC energiának egy ekvivalens átírása a phase field modellbe. Feladatunk a phase field mezőben kapott, stabilitási analízisnek megfelelő paraméterek pontos átszámítása a Markov mezőbe.



2. ábra. A magasabb rendű interakciós függvény $\mathcal{G}(\|x - x'\|)$ $d = 2$ esetén.

2. A modell interpretációja Markov mezőként

2.1. Az energia diszkretizálása

A phase field modellben mind a tartomány, mind pedig a kontúr folytonos, míg a Markov mezős modell egy diszkrét modell, melyben a tartomány és az értékkészlet is diszkrét értékeket vehet fel. Ezért a Markov mezős modell megkonstruálása előtt diszkretizálnunk kell a phase field modellt.

A nem-lokális energiatag átalakítása. Induljunk ki a [1]-ben bemutatott (és az előzőekben már említett) phase field energiából. A nem lokális energiatagot a következőképp tudjuk felbontani (2):

$$\begin{aligned}
 -\frac{\beta_f}{2} \int_{\mathcal{D}^2} \nabla \phi \mathbf{G}(x, x') \nabla \phi' d^2 x d^2 x' &= -\frac{\beta_f}{2} \int_{\mathcal{D}^2} \phi (\nabla \nabla' \Psi(\|x - x'\|) \mathbf{I}) \phi' d^2 x d^2 x' \quad (4) \\
 &= \frac{\beta_f}{2} \int_{\mathcal{D}^2} \underbrace{(\nabla^2 \Psi(\|x - x'\|) \mathbf{I})}_{\mathcal{G}} \phi \phi' d^2 x d^2 x' \quad (5)
 \end{aligned}$$

ahol $\mathcal{G}(\|x - x'\|)$ a magasabb rendű interakciós függvény (2. ábra).

A \mathcal{D} tartomány diszkretizálása. A \mathcal{D} tartomány diszkretizálásakor egyszerűen egy $\mathcal{S} \subset \mathbb{Z}^2$ rácshálót illesztünk \mathcal{D} -re, ami azt egységnégyzetekre bontja. Az $s \in \mathcal{S}$ rácspontok meghatároznak, egy-egy négyzetet - egy *cellát* - \mathcal{D} -n. Az $U(\phi)$ energia az \mathcal{S} rácshálón tehát a következőképp fog alakulni:

$$U(\phi) = \sum_{s \in \mathcal{S}} \left\{ \frac{D_f}{2} \int_{c_s} |\nabla \phi|^2 d^2 x + \alpha_f \int_{c_s} \left(\phi - \frac{\phi^3}{3} \right) d^2 x + \lambda_f \int_{c_s} \left(\frac{\phi^4}{4} - \frac{\phi^2}{2} \right) d^2 x \right\}$$

$$+ \frac{\beta_f}{2} \sum_{s \in \mathcal{S}} \sum_{r \in \mathcal{S}} \int_{c_s \times c_r} \mathcal{G}(\|x - x'\|) \phi \phi' d^2 x d^2 x' \quad (6)$$

Az energiatagok diszkretizálása. Ebben a részben kiszámítjuk a diszkrét energia funkcionálokat, amikor mind a \mathcal{D} tartomány, mind pedig a ϕ függvény diszkrét. Az eredményül kapott diszkrét mező a következőképp jelölhető: $\omega : \mathcal{S} \rightarrow \{-1, +1\}$. A mi esetünkben a diszkretizálás folyamán minden cellának az integrálját vesszük, ha ez pozitív lesz, akkor előtér pixelnek tekintjük egyébként háttérnek. Ezt a szabályt a következőképp formalizálhatjuk:

$$\omega_s := 2H \left(\int_{c_s} \phi d^2 x \right) - 1 \quad (7)$$

Ezt felhasználva elvégezhetjük az energiatagok diszkretizálást, és az alábbi diszkrét energiafüggvényhez jutunk, melynek segítségével megkonstruálhatjuk a Markov modellt:

$$U(\omega) = \lambda |\mathcal{S}| + \alpha \sum_s \omega_s + D \sum_{\|s-r\|=1} (\omega_s - \omega_r)^2 + \beta \sum_{\|s-r\| < 2d} \omega_s \omega_r F_{sr} \quad (8)$$

ahol F_{sr} a $\mathcal{G}(\|x - x'\|)$ operátor diszkrét megfelelője, amit a következőképp kapunk:

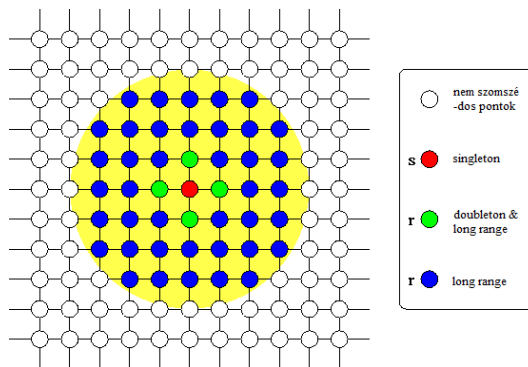
$$\begin{aligned} \frac{\beta_f}{2} \int_{\mathcal{D}^2} \mathcal{G}(\|x - x'\|) \phi \phi' d^2 x d^2 x' &= \beta_f \sum_{\|s-r\| < 2d} \int_{c_s \times c_r} \mathcal{G}(\|x - x'\|) \phi \phi' d^2 x d^2 x' \\ &\approx \beta_f \sum_{\|s-r\| < 2d} \omega_s \omega_r \underbrace{\int_{c_s \times c_r} \mathcal{G}(\|x - x'\|) d^2 x d^2 x'}_{F_{sr}} \end{aligned} \quad (9)$$

Az így kapott energia első tagja $U(\omega)$ minimalizálása alatt elhagyható, hiszen ez csak a kép méretétől függő konstans tag, a második tag fogja befolyásolni az előtér és háttér pixelek arányát, a harmadik tag a kontúrhosszt minimalizálja a régiók mentén, míg az utolsó tag felelős azért, hogy végeredményben körök alakuljanak ki. A diszkretizálás levezetése után a következő összefüggések adódtak a phase field és a diszkrét modell paramétereinek között:

$$\alpha = \frac{2}{3} \alpha_f; \quad D = \frac{1}{2} D_f; \quad \lambda = -\frac{1}{4} \lambda_f; \quad \beta = \beta_f \quad (10)$$

2.2. A Markov modell

A Markov mezős modellünk megkonstruálásakor a *singleton* és a *doubleton* klikkek mellett definiáljuk a magasabb rendű tag által megadott *long range* klikket, melynek érdekessége, hogy az eddigi MRF modellekkel ellentétben ez a szomszédsági viszony nem minden esetben a homogenitást, hanem d távolságon kívül a pontok ellentétes címkézését részesíti előnyben.

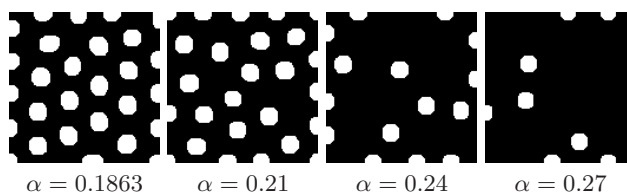


3. ábra. A $\mathcal{G}(\|x - x'\|)$ magasabb rendű interakciós függvény (2. ábra) által meghatározott MRF szomszédsági rendszer, $d = 2$ esetén (azaz $\|x - x'\| < 4$).

Singleton potenciál. A *singleton* potenciálnak a pontok aktuális címkéi felelnek meg. Az adatmodellünk, a kép intenzitását használva, minden osztályt egy normál eloszlással modellez. Egy előre megadott maszk alapján számoljuk az osztályokhoz tartozó μ és σ értékeket, majd az így kapott tagot egy γ súlyfaktortal hozzáadjuk a *singleton* energiához. Így a singleton potenciál a következő lesz:

$$V'_s = \alpha\omega_s + \gamma \left(\ln(\sqrt{(2\pi)}\sigma_\lambda) + \frac{(x - \mu_\lambda)^2}{2\sigma_\lambda^2} \right), \quad (11)$$

ahol γ az adattag súlyfaktora. Lehetőségünk van az α paraméter változtatásával a kép lefedettségét szabályozni, azaz hogy hány kör alakuljon ki. Amennyiben növeljük α értékét, úgy egyre kevesebb előtérpixel lesz a képen, ami kevesebb kört fog eredményezni (4. ábra). Valós képek esetén fontos lesz ennek a paraméternek a pontos beállítása, mert könnyen *fantomkörök* jöhetnek létre (azaz a geometriai modell oda is kört rajzol, ahol az nincs is).



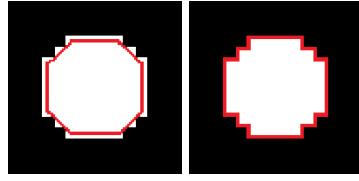
4. ábra. Az α paraméter változtatása. α növelése egyre több háttérpixel eredményez, azaz egyre kevesebb kör jelenik meg ($d = 8$; $\beta = 0.096$; $D = 0.1545$).

Doubleton potenciál. A *doubleton* potenciállal közelítjük a kontúrhosszt. A phase field modellben a kontúrhosszt folytonosan definiálják, így jelentősen rövidebb

lesz, mint a diszkrét modell esetén (lásd 5. ábra). Jól látható, hogy a kis sugarú körök esetén igen jelentős a torzulás, míg egy bizonyos méret felett egy stabil konstans értéket vesz fel a két modellben mért kontúrhosszak aránya. Ezen felül még azt is figyelembe kell vennünk, hogy a Markov mező esetén a kontúr a pixelpontok határán fut, így a *doubleton* miatt a kontúr energiáját kétszer számoljuk. Azaz a doubleton a következőképp írható fel:

$$\forall \{s, r\}, \|s - r\| = 1 : V_{\{s,r\}} = D'(\omega_s - \omega_r)^2 = \begin{cases} 4D' & \text{ha } \omega_s \neq \omega_r \\ 0 & \text{egyébként} \end{cases} \quad (12)$$

ahol $D' = \frac{0.82}{2}D = \frac{0.82}{4}D_f$.



| r | Phase field | MRF | PF/MRF |
|----|-------------|-----|--------|
| 2 | 8 | 12 | 0.66 |
| 5 | 27.3 | 36 | 0.75 |
| 10 | 62.6 | 76 | 0.82 |
| 20 | 128.5 | 156 | 0.82 |
| 50 | 326.4 | 396 | 0.82 |

5. ábra. **Képek:** A kontúr hossza a folytonos (bal) és a diszkrét (jobb) modellekben. **Táblázat:** A kontúrhossz alakulása különböző sugár esetén a phase field és a Markov mező esetén. Az első oszlop a sugárméretet, a következő kettő a kontúrhosszt mutatja a két modellnél, az utolsó oszlop a hosszok arányát. Jól látszik, hogy kis sugár esetén a diszkrétizálás miatt nagy a hiba.

Long range potenciál. A long range potenciált azok között a pontok között definiáljuk, amik legfeljebb $2d$ távolságra vannak egymástól.

$$\forall \{s, r\}, \|s - r\| < 2d : V'_{\{s,r\}} = \beta F_{sr} \omega_s \omega_r = \begin{cases} -\beta F_{sr} & \text{ha } \omega_s \neq \omega_r \\ +\beta F_{sr} & \text{egyébként} \end{cases} \quad (13)$$

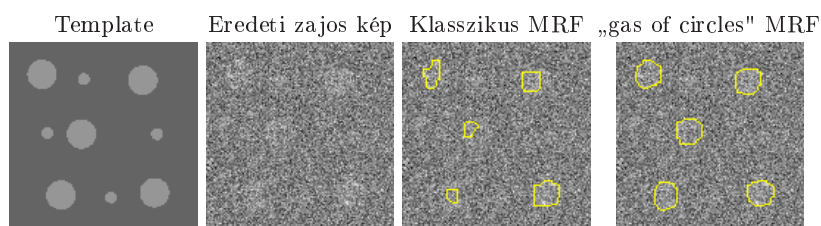
Megfigyelhetjük, hogy ez a tag két pontnak (s -nek és r -nek) ugyanazt a címkét javasolja, amennyiben $\|s - r\| < d'$ (*vonzó erő*) és ellentétes címkét, amennyiben $d' < \|s - r\| < 2d$ (*taszító erő*) és nincs semmilyen hatással, amennyiben $\|s - r\| = 0$ vagy $\|s - r\| = d'$, ahol $d' = d + \epsilon$, ugyanis a magasabb rendű interakciós függvényünk zéruspontja nem pont d -ben van (lásd 2. ábra).

Lokális energia. Összegezve az eddigieket, minden $s \in \mathcal{S}$ ponthoz definiáltunk egy lokális energiát:

$$U_s(\omega) = V_s + \sum_{\{r:\|s-r\|=1\}} V_{\{s,r\}} + \sum_{\{r:\|s-r\|<2d\}} V'_{\{s,r\}} \quad (14)$$

MRF energia-minimalizálás. A Markov mező energiafüggvényének minimalizálásához egy klasszikus globális minimalizáló módszert, a szimulált hűtést [5] használtuk. Így az energiaminimalizálás végére - a phase field és az aktív kontúr modellekkel ellentétben - megközelítjük az energiafüggvényünk globális minimumát. A kezdeti hőmérséklet 3 volt és a hőmérsékletet exponenciálisan csökkenttük: $T_{n+1} = 0.97T_n$, a következő megállási feltétellel: $T = 0.1$.

3. Eredmények



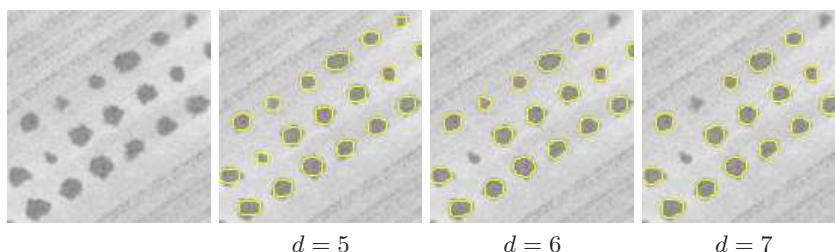
6. ábra. Eredmények erősen zajos képre (-16dB). A „gas of circles” MRF modell sokkal szebben szegmentálja a köröket az előzetes alakinformáció segítségével.

| MRF | | | gocMRF | | |
|-----|--------|--------|--------|--------|--------|
| Zaj | FP (%) | FN (%) | Zaj | FP (%) | FN (%) |
| 0 | 0.3 | 0.3 | 0 | 0.2 | 1.0 |
| -5 | 1.8 | 2.0 | -5 | 1.2 | 1.7 |
| -10 | 2.5 | 7.2 | -10 | 2.1 | 4.3 |
| -16 | 4.2 | 24.5 | -16 | 3.9 | 8.5 |
| -20 | 9.8 | 39.6 | -20 | 6.4 | 16.5 |

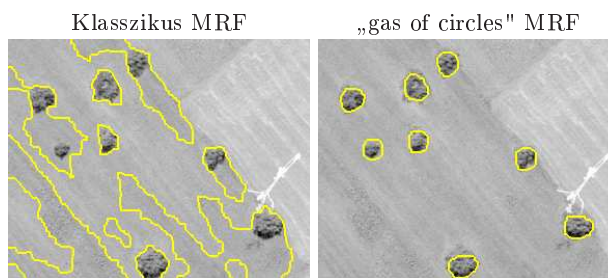
1. táblázat. Eredmények szintetikus, zajos képeken.

A klasszikus Markov mezős modell az erősen zajos képekre nem találja meg a körök határát, míg a mi modellünk az alakinformáció segítségével sokkal jobb eredményt ér el (lásd 6. ábra). Teszteltük a klasszikus MRF és a gocMRF modelleket különböző zajszintek mellett. A tesztelés folyamán minden képre a legjobb eredményt vettük. Az 1. táblázat egy 20 különböző szintetikus képet – mind egyiket 8 különböző zajszinttel (0 és -20 dB között), azaz összesen 160 képet – tartalmazó adatbázis alapján készült kiértékelést mutat. A gocMRF modell által adott eredményeket összehasonlítjuk a klasszikus MRF modell által adottakkal, ami nem tartalmaz előzetes alakinformációt. A korrekt összehasonlítás érdekében, a *fals-pozitív* (FP) és a *fals-negatív* (FN) hibát nem vesszük figyelembe a klasszikus MRF esetében a kis körökön. Ugyanis definíció szerint a klasszikus

MRF mindent szegmentál, ami eltér objektum, de jelen esetben mi csak az előre beállított sugarú köröket szeretnénk detektálni. Az értékek jól mutatják, hogy a bemutatott modell sokkal kevésbé érzékeny a zajra.



7. ábra. Nagyobb d esetén, a kisebb fakoronákat nem szegmentáljuk.



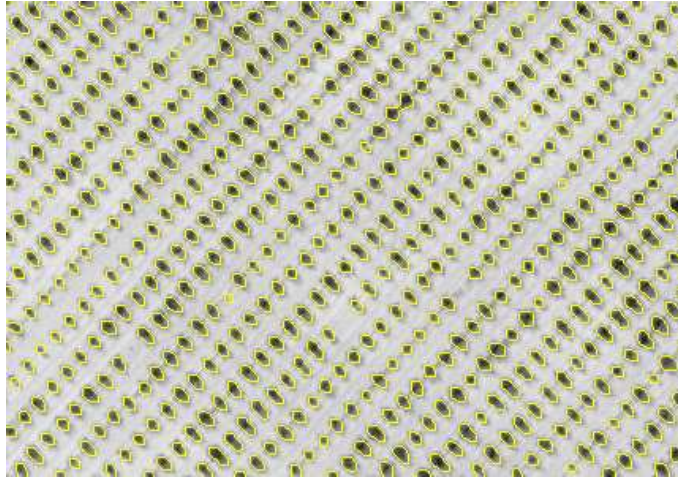
8. ábra. A klasszikus Markov mező nem tudja jól elkülöníteni a fákat, mert azok intenzitása hasonlít a háttér egyes részeihez, így ezeket összemosza.

4. Konklúzió

A legfőbb eredmény, hogy sikerült megkonstruálni egy olyan Markov modellt, ami képes előzetes alak információkat kezelni, és külön érdekessége a modellnek, hogy a korábbi MRF modellekhez képest, egyes szomszédos pontok között taszító hatás lép fel, ami végülis az előzetesen definiált forma kialakulását eredményezi. Megmutattuk, hogy a phase field és HOAC modellekben megkonstruált energiafüggvények átültethetőek MRF-be és sikerült kísérleti úton is bizonyítani ekvivalenciájukat. Bemutattuk a *gocMRF* modell viselkedését egy valós problémán is, melyben fakoronákat detektáltunk légi felvételekről.

A paraméterbeállítások módosításával elérhetjük, hogy a modell úthálózatot alakítson ki, így lehetőség van nem csak kör alakú objektumok, hanem útszerű előterek szegmentálására is (hasonlóan a phase field modellhez [1]).

További lehetőség lehet másfajta, egyszerűbb alakzatokra is kiterjeszteni a modellt (téglalap, rúd, stb.).



9. ábra. Szabályosan ültetett fenyőerdő, eredmény a gocMRF modellel.

Hivatkozások

1. M. Rochery, I. H. Jermyn, and J. Zerubia, „Phase field models and higher-order active contours,” in *Proc. IEEE International Conference on Computer Vision*, Beijing, China, October 2005.
2. P. Horvath, „The 'gas of circles' model and its application to tree crown extraction, Thesis for the degree of Doctor of Philosophy, 2007.
3. Z. Kato and Ting-Chuen Pong, „A Markov random field image segmentation model for color textured images,” *Image and Vision Computing* 24, 1103-1114, 2006.
4. P. Horvath, I. H. Jermyn, J. Zerubia, and Z. Kato, „A higher-order active contour model of a gas of circles' and its application to tree crown extraction,” *Pattern Recognition*, Sept. 2008, article in press
5. S. Geman and D. Geman, „Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, 1984.