Two Compound Random Field Texture Models

Michal Haindl[(✉)] and Vojtěch Havlíček

The Institute of Information Theory and Automation
of the Czech Academy of Sciences, Prague, Czech Republic
{haindl,havlicek}@utia.cz

Abstract. Two novel models for texture representation using parametric compound random field models are introduced. These models consist of a set of several sub-models each having different characteristics along with an underlying structure model which controls transitions between them. The structure model is a two-dimensional probabilistic mixture model either of the Bernoulli or Gaussian mixture type. Local textures are modeled using the fully multispectral three-dimensional causal autoregressive models. Both presented compound random field models allow to reproduce, compress, edit, and enlarge a given measured color, multispectral, or bidirectional texture function (BTF) texture so that ideally both measured and synthetic textures are visually indiscernible.

Keywords: Texture · Texture synthesis · Compound random field model · CAR model · Two-dimensional Bernoulli mixture · Two-dimensional Gaussian mixture · Bidirectional texture function

1 Introduction

Physically correct and visually convincing virtual models require object surfaces covered with realistic nature-like surface material textures to present realism in virtual scenes. The primary purpose of any synthetic texture approach is to reproduce and enlarge a given measured material texture so that ideally both natural and synthetic texture will be visually indiscernible. The appearance of real materials dramatically changes with illumination and viewing variations and its most advanced current texture representation is the seven-dimensional Bidirectional Texture Function (BTF) [9]. Unfortunately, measured texture data are nearly always too limited to reliable estimate these complex seven-dimensional models, thus their modeling requires some simplifying factorization [9], such as the presented compound random field models which serve as the the three-dimensional factor model in the complex overall BTF material model [9].

Compound random field models (CRF) consist of several sub-models each having different characteristics along with an underlying structure model which controls transitions between these sub models [11]. Compound Markov filed models were successfully applied to image restoration [2,3,11,12], segmentation [13], and modeling [6,7,10]. However, these models always require demanding numerical

solutions with all their well known drawbacks. Our exceptional CMRF [6] model allows analytical synthesis at the cost of a slightly compromised compression rate.

We propose two textural models - $\text{CRF}^{BM-3CAR}$, $\text{CRF}^{GM-3CAR}$, based on complex spatial probabilistic mixture models. These control field models are either probabilistic Bernoulli of Gaussian mixture models.

2   Compound Random Field Texture Models

Let us denote a multiindex $r = (r_1, r_2)$, $r \in I$, where I is a discrete 2-dimensional rectangular lattice and r_1 is the row and r_2 the column index, respectively. $X_r \in \{1, 2, \ldots, K\}$ is a random variable with natural number value (a positive integer), Y_r is multispectral pixel at location r and $Y_{r,j} \in \mathcal{R}$ is its j-th spectral plane component. Both random fields (X, Y) are indexed on the same lattice I. Let us assume that each multispectral or BTF observed texture $\tilde{Y}$ (composed of d spectral planes) can be modelled by a compound random field model, where the principal random field X controls switching to a regional local model $Y = \bigcup_{i=1}^{K} {}^{i}Y$. Single K regional sub-models ${}^{i}Y$ are defined on their corresponding lattice subsets ${}^{i}I$, ${}^{i}I \cap {}^{j}I = \emptyset$ $\forall i \neq j$ and they are of the same RF type. They differ only in their contextual support sets ${}^{i}I_r$ and corresponding parameters sets ${}^{i}\theta$. The CRF model has posterior probability $P(X, Y \,|\, \tilde{Y}) = P(Y \,|\, X, \tilde{Y})P(X \,|\, \tilde{Y})$ and the corresponding optimal MAP solution is: $(\hat{X}, \hat{Y}) = \arg\max_{X \in \Omega_X, Y \in \Omega_Y} P(Y \,|\, X, \tilde{Y}) \, P(X \,|\, \tilde{Y})$, where Ω_X, Ω_Y are corresponding configuration spaces for random fields (X, Y). To avoid an iterative MCMC MAP solution, we propose the following two step approximation [6]:

$$(\check{X}) = \arg\max_{X \in \Omega_X} P(X \,|\, \tilde{Y}), \tag{1}$$

$$(\check{Y}) = \arg\max_{Y \in \Omega_Y} P(Y \,|\, \check{X}, \tilde{Y}). \tag{2}$$

This approximation significantly simplifies the $\text{CRF}^{BM-3CAR}$, $\text{CRF}^{GM-3CAR}$ estimation because it allows us to take advantage of a straightforward analytical estimation of the regional RF models ${}^{i}Y$ in (2).

2.1   Region Switching Model

The control RF $(P(X \,|\, \tilde{Y}))$ is supposed to be represented by two-dimensional Bernoulli or Gaussian distribution mixture model, respectively. The mixture distribution $P(Y_{\{r\}})$ has the form:

$$P(Y_{\{r\}}) = \sum_{m \in \mathcal{M}} P(Y_{\{r\}} \,|\, m) \, p(m) = \sum_{m \in \mathcal{M}} \prod_{s \in I_r} p_s(Y_s \,|\, m) \, p(m) \tag{3}$$

where $Y_{\{r\}} \in \mathcal{K}^{\eta}$, $\mathcal{M} = \{1, 2, \ldots, M\}$, $I_r \subset I$ is an index set, $\eta = cardinality\{I_r\}$, and $p(m)$ are probability weights $\sum_{m \in \mathcal{M}} p(m) = 1$. The maximum-likelihood parameter estimates $p(m)$ (probability weights),

μ_{ms}, σ_{ms} (Gaussian mixture component means and standard deviation), $\theta_{m,s}$ (Bernoulli mixture component parameters) are computed using the EM algorithm [1,4] $p_s^{(t+1)}(.\,|\,m)$ and

$$q^{(t)}(m\,|\,X_{\{r\}}) = \frac{p^{(t)}(m)\,P^{(t)}(X_{\{r\}}\,|\,m)}{\sum_{j \in \mathcal{M}} p^{(t)}(j) P^{(t)}(X_{\{r\}}\,|\,j)}, \tag{4}$$

$$p^{(t+1)}(m) = \frac{1}{|\mathcal{S}|} \sum_{X_{\{r\}} \in \mathcal{S}} q^{(t)}(m\,|\,X_{\{r\}}). \tag{5}$$

Bernoulli Distribution Mixture Model. We assume that control field pixel $X_r \in \mathcal{K}$ where $\mathcal{K}$ is the index set of K distinguished sub-models. The distribution $P(X_{\{r\}})$ is assumed to be multivariable Bernoulli mixture (BM) and the control field is further decomposed into separate binary bit planes of binary variables $\xi \in \mathcal{B}$, $\mathcal{B} = \{0,1\}$ which are separately modeled and can be learned from much smaller training texture than a multi-level discrete mixture model. We suppose that a bit factor of a control field can be fully characterised by a marginal probability distribution of binary levels on pixels within the scope of a window centered around the location r and specified by the index set $I_r \subset I$, i.e. $X_{\{r\}} \in \mathcal{B}^\eta$ and $P(X_{\{r\}})$ is the corresponding marginal distribution of $P(X\,|\,\tilde{Y})$. The component distributions $P(\cdot\,|\,m)$ are factorisable, and multivariable Bernoulli:

$$P(X_{\{r\}}\,|\,m) = \prod_{s \in I_r} \theta_{m,s}^{X_s} (1 - \theta_{m,s})^{1 - X_s} \qquad X_s \in X_{\{r\}}. \tag{6}$$

The mixture model parameters (6) include component weights $p(m)$ and the univariate discrete distributions of binary levels. They are simply defined by one parameter $\theta_{m,s}$ as a vector of probabilities:

$$p_s(\cdot\,|\,m) = (\theta_{m,s}, 1 - \theta_{m,s}). \tag{7}$$

The EM solution is (4), (5) and

$$p_s^{(t+1)}(\xi\,|\,m) = \frac{1}{|\mathcal{S}|\,p^{(t+1)}(m)} \sum_{X_{\{r\}} \in \mathcal{S}} \delta(\xi, X_s)\,q^{(t)}(m\,|\,X_{\{r\}}), \quad \xi \in \mathcal{B}. \tag{8}$$

The total number of mixture (3), (7) parameters is thus $M(1+\eta)$ – confined to the appropriate norming conditions. The advantage of the multivariable Bernoulli model (7) is a simple switch-over to any marginal distribution by deleting superfluous terms in the products $P(X_{\{r\}}\,|\,m)$.

Gaussian Mixture Model. The discrete control field can be alternatively modeled by a continuous RF if we map single indices into continuous random variables with uniformly separated mean values and small variance. The synthesis results are subsequently inversely mapped back into a corresponding synthetic discrete control field. We assume the joint probability distribution $P(X_{\{r\}})$,

$X_{\{r\}} \in \mathcal{K}^\eta$ in the form of a normal mixture and the mixture components are defined as products of univariate Gaussian densities

$$P(X_{\{r\}} \,|\, \mu_m, \sigma_m) = \prod_{s \in I_{\{r\}}} p_s(X_s \,|\, \mu_{ms}, \sigma_{ms}), \qquad (9)$$

$$p_s(X_s \,|\, \mu_{ms}, \sigma_{ms}) = \frac{1}{\sqrt{2\pi}\sigma_{ms}} \exp\left\{ -\frac{(X_s - \mu_{ms})^2}{2\sigma_{ms}^2} \right\},$$

i. e., the components are multivariate Gaussian densities with diagonal covariance matrices. The maximum-likelihood estimates of the parameters $p(m), \mu_{ms}, \sigma_{ms}$ can be computed by EM algorithm [1,4]. Anew we use a data set $\mathcal{S}$ obtained by pixel-wise shifting the observation window within the original texture image $\mathcal{S} = \{X_{\{r\}}^{(1)}, \ldots, X_{\{r\}}^{(K)}\}$, $X_{\{r\}}^{(k)} \subset X$. The corresponding log-likelihood function is maximized by the EM algorithm $(m \in \mathcal{M}, n \in \mathcal{N}, X_{\{r\}} \in \mathcal{S})$ and the iterations are $(4), (5)$ and

$$\mu_{m,n}^{(t+1)} = \frac{1}{\sum_{X_{\{r\}} \in \mathcal{S}} q^{(t)}(m \,|\, X_{\{r\}})} \sum_{X_{\{r\}} \in \mathcal{S}} X_n \, q(m \,|\, X_{\{r\}}), \qquad (10)$$

$$(\sigma_{m,n}^{(t+1)})^2 = -(\mu_{m,n}^{(t+1)})^2 + \frac{\sum_{X_{\{r\}} \in \mathcal{S}} X_n^2 \, q^{(t)}(m \,|\, X_{\{r\}})}{\sum_{X_{\{r\}} \in \mathcal{S}} q(m|X_{\{r\}})}. \qquad (11)$$

Control Field Synthesis. We can assume at a given position r of the contextual neighbourhood I_r to have some part of the pixel-wise synthesised control field $X_{\{r\}}$ already specified. If $X_{\{\rho\}}$ is a sub-vector of all of $X_{\{r\}}$ pixels previously specified within this window and $I_\rho \subset I_r$ the corresponding index subset, then the statistical properties of the remaining unspecified variables are fully described by the corresponding conditional distribution:

$$p_{n \,|\, \rho}(X_n \,|\, X_{\{\rho\}}) = \sum_{m=1}^{M} W_m(X_{\{\rho\}}) \, p_n(X_n \,|\, m), \qquad (12)$$

where $W_m(X_{\{\{\rho\}\}})$ are the a posteriori component weights corresponding to the given sub-vector $X_{\{\rho\}}$:

$$W_m(X_{\{\rho\}}) = \frac{p(m)P_\rho(X_{\{\rho\}} \,|\, m)}{\sum_{j=1}^{M} p(j)P_\rho(X_{\{\rho\}} \,|\, j)}, \qquad (13)$$

$$P_\rho(X_{\{\rho\}} \,|\, m) = \prod_{n \in \rho} p_n(X_n \,|\, m).$$

X_n can be randomly generated by the conditional distribution $p_{n \,|\, \rho}(X_n \,|\, X_{\{\rho\}})$ where by Eq. (12) can be applied to all the unspecified variables $n = \eta - \text{card}\{\rho\}$ given a fixed position of the control field. Each newly generated X_n is used to upgrade the conditional weights $W_m(X_{\{\rho\}})$.

2.2   Local Markov Models

Local i-th texture region (not necessarily continuous) is represented by the adaptive 3D causal auto-regressive random (3DCAR) field model [5,8]. This model can be analytically estimated as well as easily synthesised. The model can be defined in the following matrix equation (i-th model index is further omitted to simplify notation):

$$Y_r = \gamma\, Z_r + \epsilon_r, \tag{14}$$

where $Z_r = [Y_{r-s}^T : \forall s \in I_r]^T$ is the $\eta\, d \times 1$ data vector with multiindices r, s, t, $\gamma = [A_1, \ldots, A_\eta]$ is the $d \times d\, \eta$ unknown parameter matrix with parametric sub-matrices A_s. The model functional contextual neighbour index shift set is denoted I_r and $\eta = cardinality(I_r)$. All CAR model statistics can be efficiently estimated analytically [8]. Given the known 3DCAR process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ the parameter estimation $\hat{\gamma}$ can be accomplished using fast, numerically robust and recursive statistics [8]:

$$\hat{\gamma}_{t-1}^T = V_{zz(t-1)}^{-1} V_{zy(t-1)}, \qquad\qquad V_{t-1} = \tilde{V}_{t-1} + V_0,$$

$$\tilde{V}_{t-1} = \begin{pmatrix} \sum_{u=1}^{t-1} Y_u Y_u^T & \sum_{u=1}^{t-1} Y_u Z_u^T \\ \sum_{u=1}^{t-1} Z_u Y_u^T & \sum_{u=1}^{t-1} Z_u Z_u^T \end{pmatrix} = \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}_{zy(t-1)}^T \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{zz(t-1)} \end{pmatrix},$$
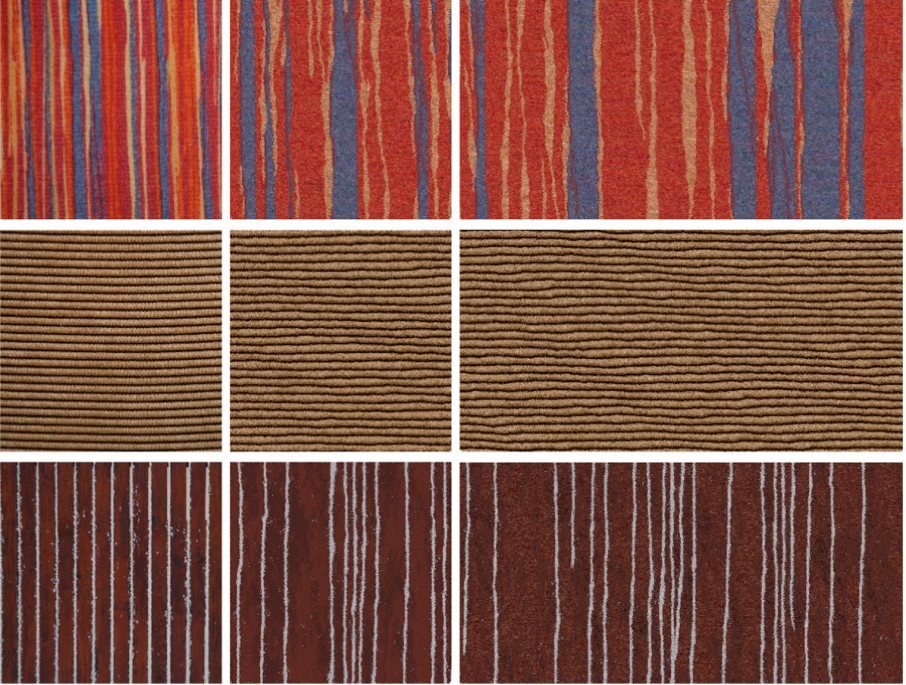

Fig. 1. Cloth and frosted planks (left column) synthesis and enlargement (right column) using the $CRF^{BM-3CAR}$ model.

where V_0 is a positive definite matrix (see [8]). Although, an optimal causal functional contextual neighbourhood I_r can be solved analytically by a straight-forward generalisation of the Bayesian estimate in [8], we use faster approximation which does not need to evaluate statistics for all possible I_r configurations. This approximation is based on spatial correlations. Starting from the causal part
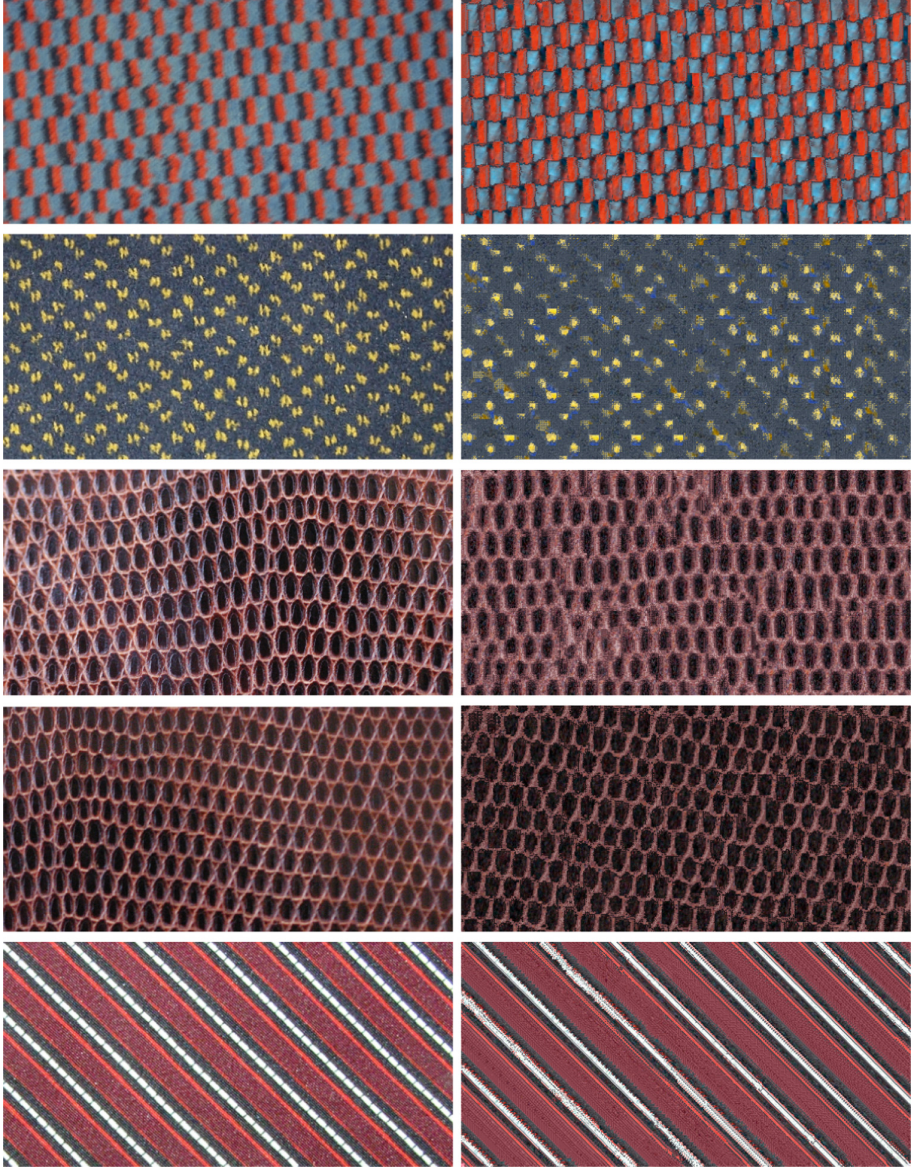

Fig. 2. Measured three textile and two cobra skin textures (left column) and their synthesis using the $CRF^{GM-3CAR}$ model.

of a hierarchical non-causal neighbourhood, neighbours locations corresponding to spatial correlations larger than a specified threshold (>0.6) are selected. The i-th model pixel-wise synthesis is simple direct application of (14) for all 3DCAR models. 3DCAR models provide better spectral modelling quality than the alternative spectrally decorrelated 2D models for motley textures at the cost of small increase of number of parameters to be stored.

3 Experiments

Both presented compound random field models ($CRF^{BM-3CAR}$, $CRF^{GM-3CAR}$) are well suited for near-regular textures such as textile materials which are notoriously difficult for Markov random field type of textural models [6,9]. The dimension of the estimated control field model distribution is not too high ($\eta \approx 10^1 - 10^2$) and the number of the training data vectors is relatively large ($|\mathcal{S}| \approx 10^4 - 10^5$). Nevertheless the window should always be kept reasonably small and the sample size as large as possible. In our experiments we have used a regular left-to-right and top-to-down shifting of the generating window. Figure 1 illustrates the $CRF^{BM-3CAR}$ model applied to a frosted planks and two textile textures synthesis and enlargement, while Fig. 2 shows three textile materials and two skin samples synthesized using the $CRF^{GM-3CAR}$ model.

4 Conclusion

Both presented CRF ($CRF^{GM-3CAR}, CRF^{BM-3CAR}$) methods show good visual performance on selected real-world measured materials. The appearance of such materials should consist of several types of relatively small regions with fine-granular inner structure such as fabric, skin, or wood. The models offer large data compression ratio (only tens of parameters per BTF) easy simulation and fast seamless synthesis of any required texture size. The methods can be easily generalised for colour or BTF texture editing by estimating some local models from different target materials. The model does not compromise spectral correlation thus it can reliably model motley textures. A drawback of the presented CRF models is the need to have sufficiently large learning data for both mixture sub-models.

References

1. Dempster, A., Laird, N., Rubin, D.: Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Stat. Soc., B **39**(1), 1–38 (1977)
2. Figueiredo, M., Leitao, J.: Unsupervised image restoration and edge location using compound Gauss - Markov random fields and the MDL principle. IEEE Trans. Image Process. **6**(8), 1089–1102 (1997)

3. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions and Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. **6**(11), 721–741 (1984)
4. Grim, J., Haindl, M.: Texture modelling by discrete distribution mixtures. Comput. Stat. Data Anal. **41**(3–4), 603–615 (2003)
5. Haindl, M., Havlíček, V.: A multiscale colour texture model. In: Kasturi, R., Laurendeau, D., Suen, C. (eds.) Proceedings of the 16th International Conference on Pattern Recognition, pp. 255–258. IEEE Computer Society, Los Alamitos, August 2002. http://dx.doi.org/10.1109/ICPR.2002.1044676
6. Haindl, M., Havlíček, V.: A compound MRF texture model. In: Proceedings of the 20th International Conference on Pattern Recognition, ICPR 2010, pp. 1792–1795. IEEE Computer Society CPS, Los Alamitos, August 2010. http://doi.ieeecomputersociety.org/10.1109/ICPR.2010.442
7. Haindl, M., Remeš, V., Havlíček, V.: Potts compound Markovian texture model. In: Proceedings of the 21st International Conference on Pattern Recognition, ICPR 2012, pp. 29–32. IEEE Computer Society CPS, Los Alamitos, November 2012
8. Haindl, M.: Visual data recognition and modeling based on local Markovian models. In: Florack, L., Duits, R., Jongbloed, G., Lieshout, M.C., Davies, L. (eds.) Mathematical Methods for Signal and Image Analysis and Representation, Computational Imaging and Vision, vol. 41, chap. 14, pp. 241–259. Springer, Heidelberg (2012). http://dx.doi.org/10.1007/978-1-4471-2353-8_14
9. Haindl, M., Filip, J.: Visual Texture. Advances in Computer Vision and Pattern Recognition. Springer-Verlag, London (2013)
10. Haindl, M., Havlíček, V.: A plausible texture enlargement and editing compound Markovian model. In: Salerno, E., Çetin, A.E., Salvetti, O. (eds.) MUSCLE 2011. LNCS, vol. 7252, pp. 138–148. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32436-9_12. http://www.springerlink.com/content/047124j43073m202/
11. Jeng, F.C., Woods, J.W.: Compound Gauss-Markov random fields for image estimation. IEEE Trans. Sig. Process. **39**(3), 683–697 (1991)
12. Molina, R., Mateos, J., Katsaggelos, A., Vega, M.: Bayesian multichannel image restoration using compound Gauss-Markov random fields. IEEE Trans. Image Proc. **12**(12), 1642–1654 (2003)
13. Wu, J., Chung, A.C.S.: A segmentation model using compound Markov random fields based on a boundary model. IEEE Trans. Image Process. **16**(1), 241–252 (2007)