ILLUMINATION INVARIANT UNSUPERVISED SEGMENTER

Michal Haindl,   Stanislav Mikeš,   Pavel Vácha

Institute of Information Theory and Automation of the ASCR
Academy of Sciences CR, 182 08 Prague, Czech Republic

ABSTRACT

A novel illumination invariant unsupervised multispectral texture segmentation method with unknown number of classes is presented. Multispectral texture mosaics are locally represented by illumination invariants derived from four directional causal multispectral Markovian models recursively evaluated for each pixel. Resulted parametric space is segmented using a Gaussian mixture model based unsupervised segmenter. The segmentation algorithm starts with an over segmented initial estimation which is adaptively modified until the optimal number of homogeneous texture segments is reached. The performance of the presented method is extensively tested on the large illumination invariant benchmark from the Prague Segmentation Benchmark using 21 segmentation criteria and compares favourably with an alternative segmentation method.

Index Terms— Unsupervised image segmentation, illumination invariance.

1. INTRODUCTION

Segmentation is the fundamental process which affects the overall performance of any automated image analysis system. Image regions, homogeneous with respect to some usually textural or colour measure, which result from a segmentation algorithm are analysed in subsequent interpretation steps. Realistic applications of these segmenters often have to deal with variable illumination of the segmented scene. Texture-based image segmentation is an area of intense research activity in recent years and many algorithms were published in consequence of all this effort. These methods are usually categorised [1] as region-based, boundary-based, or as a hybrid of the two. Different published methods are difficult to compare because of lack of a comprehensive analysis together with accessible experimental data, however available results indicate that the ill-defined texture segmentation problem is still far from being satisfactorily solved. Spatial interaction models and especially Markov random fields-based models are increasingly popular for texture representation [1, 2, 3], etc. Several researchers dealt with the difficult problem of

unsupervised segmentation using these models see for example [4, 5, 6, 7, 8] or [12] which is also addressed in this paper. The outline of this paper is as follows. Section 2 presents our illumination invariant Markovian multispectral texture representation. Section 3 outlines the unsupervised segmenter, followed by the experimental verification in the subsequent section 4 and concluding section 5.

2. TEXTURE MODEL

Multispectral static smooth textures require three dimensional (3D) models for adequate representation. We assume that single multispectral textures can be locally modelled using a 3D simultaneous causal autoregressive random field model (AR3D). This model can be expressed as a stationary causal uncorrelated noise driven 3D autoregressive process [9]:

$$Y_r = \gamma X_r + e_r , \qquad (1)$$

where $\gamma = [A_1, \ldots, A_\eta]$ is the $d \times d\eta$ parameter matrix, d is the number of spectral bands, I_r^c is a causal neighborhood index set with $\eta = card(I_r^c)$ and e_r is a white Gaussian noise vector with zero mean and a constant but unknown covariance, X_r is a corresponding vector of the contextual neighbours Y_{r-s} and $r, r - 1, \ldots$ is a chosen direction of movement on the image index lattice I. The selection of an appropriate AR3D model support (I_r^c) is important to obtain good texture representation for realistic texture synthesis but less important for adequate texture segmentation. The optimal neighbourhood as well as the Bayesian parameters estimation of a AR3D model can be found analytically under few additional and acceptable assumptions using the Bayesian approach (see details in [9]). The AR3D model recursive Bayesian parameter estimator is [9]:

$$\hat{\gamma}_{r-1}^T = \hat{\gamma}_{r-2}^T + \frac{V_{x(r-2)}^{-1} X_{r-1} (Y_{r-1} - \hat{\gamma}_{r-2} X_{r-1})^T}{(1 + X_{r-1}^T V_{x(r-2)}^{-1} X_{r-1})} , \quad (2)$$

where $V_{x(r-1)} = \sum_{k=1}^{r-1} X_k X_k^T + V_{x(0)}$. Local texture for each pixel is represented by four parametric vectors. Each vector contains local estimations of the AR3D model invariants. These models have identical contextual neighbourhood

I_r^c but they differ in their major movement direction (top-down, bottom-up, rightward, leftward), i.e.,

$$\tilde{\gamma}_r^T = \{\hat{\gamma}_r^t, \hat{\gamma}_r^b, \hat{\gamma}_r^r, \hat{\gamma}_r^l\}^T \ . \tag{3}$$

2.1. Illumination Invariants

We assume that the two images $\tilde{Y}, Y$ acquired with different illumination can be linearly transformed to each other: $\tilde{Y}_r = B Y_r$, where $\tilde{Y}_r, Y_r$ are multispectral pixel values at position r and B is a $d \times d$ transformation matrix. This linear relation holds for changes in brightness and illumination spectrum with Lambertian surface reflectance, or for a model which includes specular reflectance component. Using above assumption we derived [10] the illumination invariance of

1. trace: $\operatorname{tr} A_m, \ m = 1, \ldots, \eta$

2. eigenvalues: $\zeta_{m,j}$ of $A_m, \ m = 1, \ldots, \eta,$
   $j = 1, \ldots, d$.

The parametric space $\tilde{\gamma}$ is subsequently transformed into the illumination invariant parametric space $\bar{\gamma}$:

$$\begin{aligned}
\bar{\gamma}_r^T &= \begin{bmatrix} {}^t\psi, \ {}^b\psi, \ {}^r\psi, \ {}^l\psi \end{bmatrix}^T \ , \\
{}^\alpha\psi &= [{}^\alpha\zeta_{1,1}, \ldots, {}^\alpha\zeta_{\eta,d}, \operatorname{tr} {}^\alpha A_1, \ldots, \operatorname{tr} {}^\alpha A_\eta] \\
&\quad \alpha \in \{t, b, r, l\} \ .
\end{aligned} \tag{4}$$

Finally we add the local a_r, b_r components of the Lab colour coordinates to the resulting feature vector (Θ_r).

3. MIXTURE BASED SEGMENTATION

Multi-spectral texture segmentation is done by clustering in the AR3D parameter space Θ defined on the lattice I where

$$\Theta_r = [\bar{\gamma}_r, a_r, b_r]^T$$

is the modified illumination invariant parameter vector (3) computed for the lattice location r. We assume that this parametric space can be represented using the Gaussian mixture model (GM) with diagonal covariance matrices. The Gaussian mixture model for AR3D parametric representation is as follows:

$$p(\Theta_r) = \sum_{i=1}^{K} p_i \, p(\Theta_r \,|\, \nu_i, \Sigma_i) \ , \tag{5}$$

$$p(\Theta_r \,|\, \nu_i, \Sigma_i) = \frac{|\Sigma_i|^{-\frac{1}{2}}}{(2\pi)^{\frac{d}{2}}} \, e^{-\frac{(\Theta_r - \nu_i)^T \Sigma_i^{-1} (\Theta_r - \nu_i)}{2}} \ . \tag{6}$$

The mixture model equations (5),(6) are solved using a modified EM algorithm. The algorithm is initialised using ν_i, Σ_i statistics estimated from the corresponding rectangular subimages obtained by regular division of the input texture mosaic. An alternative initialisation can be random choice

of these statistics. For each possible couple of rectangles the Kullback Leibler divergence

$$D\left(p(\Theta_r \,|\, \nu_i, \Sigma_i) \,||\, p(\Theta_r \,|\, \nu_j, \Sigma_j)\right) = \int_\Omega p(\Theta_r \,|\, \nu_i, \Sigma_i) \, \log\left(\frac{p(\Theta_r \,|\, \nu_i, \Sigma_i)}{p(\Theta_r \,|\, \nu_j, \Sigma_j)}\right) \, d\Theta_r \tag{7}$$

is evaluated and the most similar rectangles, i.e.,

$$\{i, j\} = \arg\min_{k,l} D\left(p(\Theta_r \,|\, \nu_l, \Sigma_l) \,||\, p(\Theta_r \,|\, \nu_k, \Sigma_k)\right)$$

are merged together in each step. This initialization results in K_{ini} subimages and recomputed statistics ν_i, Σ_i. $K_{ini} > K$ where K is the optimal number of textured segments to be found by the algorithm. Two steps of the EM algorithm are repeating after initialisation. The components with smaller weights than a fixed threshold ($p_j < \frac{0.1}{K_{ini}}$) are eliminated. For every pair of components we estimate their Kullback Leibler divergence (7). From the most similar couple, the component with the weight smaller than the threshold is merged to its stronger partner and all statistics are actualised using the EM algorithm. The algorithm stops when either the likelihood function has negligible increase ($\mathcal{L}_t - \mathcal{L}_{t-1} < 0.05$) or the maximum iteration number threshold is reached.

The parametric vectors representing texture mosaic pixels are assigned to the clusters according to the highest component probabilities, i.e., Y_r is assigned to the cluster ω_j if

$$\pi_{r,j} = max_j \sum_{s \in I_r} w_s \, p(\Theta_{r-s} \,|\, \nu_j, \Sigma_j) \ ,$$

where w_s are fixed distance-based weights, I_r is a rectangular neighbourhood and $\pi_{r,j} > \pi_{thre}$ (otherwise the pixel is unclassified). The area of single cluster blobs is evaluated in the post-processing thematic map filtration step. Regions with similar statistics are merged. Thematic map blobs with area smaller than a given threshold are attached to its neighbour with the highest similarity value.

4. EXPERIMENTAL RESULTS

The algorithm was tested on natural colour textures mosaics from the Prague Texture Segmentation Data-Generator and Benchmark [11]. The benchmark test mosaics layouts and each cell texture membership are randomly generated and filled with colour textures from the large Prague colour texture database. The benchmark ranks segmentation algorithms according to a chosen criterion. There are implemented twenty seven most frequented evaluation criteria [11] categorised into four groups: region-based (5+5), pixel-wise (12), consistency measures (2) and clustering comparison criteria (3). Tab.1 compares the overall benchmark performance of the proposed algorithm (AR3D+EM ii) with its non illumination invariant version (AR3D+EM [12]) and the HGS

	Benchmark – Colour (Illum. Inv.)			
	AR3D+EM ii	AR3D+EM [12]	HGS E [13]	HGS C [13]
↑ CS	40.70	34.14	9.55	9.17
↓ OS	53.02	53.33	19.30	12.80
↓ US	16.76	13.29	30.05	37.48
↓ ME	13.96	20.12	39.72	38.41
↓ NE	14.85	20.57	39.64	35.36
↓ O	35.17	31.53	56.44	68.87
↓ C	91.72	95.34	60.20	51.63
↑ CA	59.15	57.87	40.20	35.81
↑ CO	65.72	64.76	53.61	50.70
↑ CC	86.36	87.17	62.45	60.67
↓ I.	34.28	35.24	46.39	49.30
↓ II.	3.83	3.52	12.11	16.15
↑ EA	68.26	68.15	51.44	46.22
↑ MS	56.91	57.23	34.80	28.32
↓ RM	5.89	4.78	12.93	16.63
↑ CI	71.32	71.40	54.22	50.03
↓ GCE	14.34	16.99	25.36	21.31
↓ LCE	7.62	8.64	16.69	12.23
↓ dM	16.58	14.64	29.18	38.39
↓ dD	19.82	20.27	29.21	29.82
↓ dVI	15.80	16.75	13.98	12.61

Table 1. Benchmark criteria (see details in http://mosaic.utia.cas.cz): CS = correct segmentation; OS = over-segmentation; US = under-segmentation; ME = missed error; NE = noise error; O = omission error; C = commission error; CA = class accuracy; CO = recall - correct assignment; CC = precision - object accuracy; I. = type I error; II. = type II error; EA = mean class accuracy estimate; MS = mapping score; RM = root mean square proportion estimation error; CI = comparison index; GCE = Global Consistency Error; LCE = Local Consistency Error; dM = Mirkin metric; dD = Van Dongen metric; dVI = variation of information;

method [13] in it both fully illumination invariant version C and the non illumination invariant version E, respectively. The HGS segmenter combines the K-means clustering with region merging step. It uses a Gabor-Gaussian spatial-colour texture representation and its illumination invariant C version uses features derived from the Gabor filters applied to log-transformed images. Our results demonstrate very good performance on all criteria with the exception of oversegmentation tendency and slightly worse variation of information criterion. The important correct region segmentation criterion is four times better than for the HGS method, undersegmentation is low as well as missed and noise errors. Our illumination invariant segmenter outperforms its non–invariant counterpart as expected, however the same conclusion cannot be claimed for the HGS method. Fig.1 shows three selected 512×512 experimental benchmark mosaics created from three to eleven natural colour textures. The last four columns demonstrate comparative results from two alternative methods, both in illumination invariant and non-invariant versions, respectively. Hard natural textures were chosen for comparison rather than synthesised (for example using the generative AR3D model or some other Markov random field model) ones because they are expected to be more difficult for the underlying segmentation model. The third column demonstrates robust behaviour of our algorithm but also infrequent algorithm failures producing the oversegmented thematic map for some textures. Such failures can be reduced by a more elaborate postprocessing step. The HGS-C [13], HGS-E [13] algorithms on these data performed steadily worse as can be seen in the last two columns of Fig.1, some areas are undersegmented while other parts of the mosaics are oversegmented. Resulting segmentation results are promising even if we could compare only one illumination invariant alternative method. However we did extensive verification and comparison of our non-invariant AR3D+EM method with 22 other leading unsupervised segmenters with very good results (see details in http://mosaic.utia.cas.cz). Our results can be easily further improved by an appropriate more elaborate postprocessing.

5. CONCLUSIONS

We proposed novel efficient and robust method for illumination invariant unsupervised texture or image segmentation with unknown number of classes based on the underlying AR3D local image representation and and Gaussian mixture parametric space models. Although the algorithm uses the random field type model it is very fast because it uses efficient recursive parameter estimation of the model and therefore is much faster than the usual Markov chain Monte Carlo estimation approach. Segmentation methods typically suffer with lot of application dependent parameters to be experimentally estimated. Our method requires only a contextual neighbourhood selection and two additional thresholds. The algorithm's performance is demonstrated on the extensive benchmark tests on natural texture mosaics. It performs favourably compared with the alternative HGS segmentation algorithm and it is faster than our previously published [8] GMRF-GM method. These test results are encouraging and we proceed with more elaborate postprocessing and some modification of the texture representation model.

6. REFERENCES

[1] Todd R. Reed and J. M. Hans du Buf, "A review of recent texture segmentation and feature extraction techniques," *CVGIP–Image Understanding*, vol. 57, no. 3, pp. 359–372, May 1993.
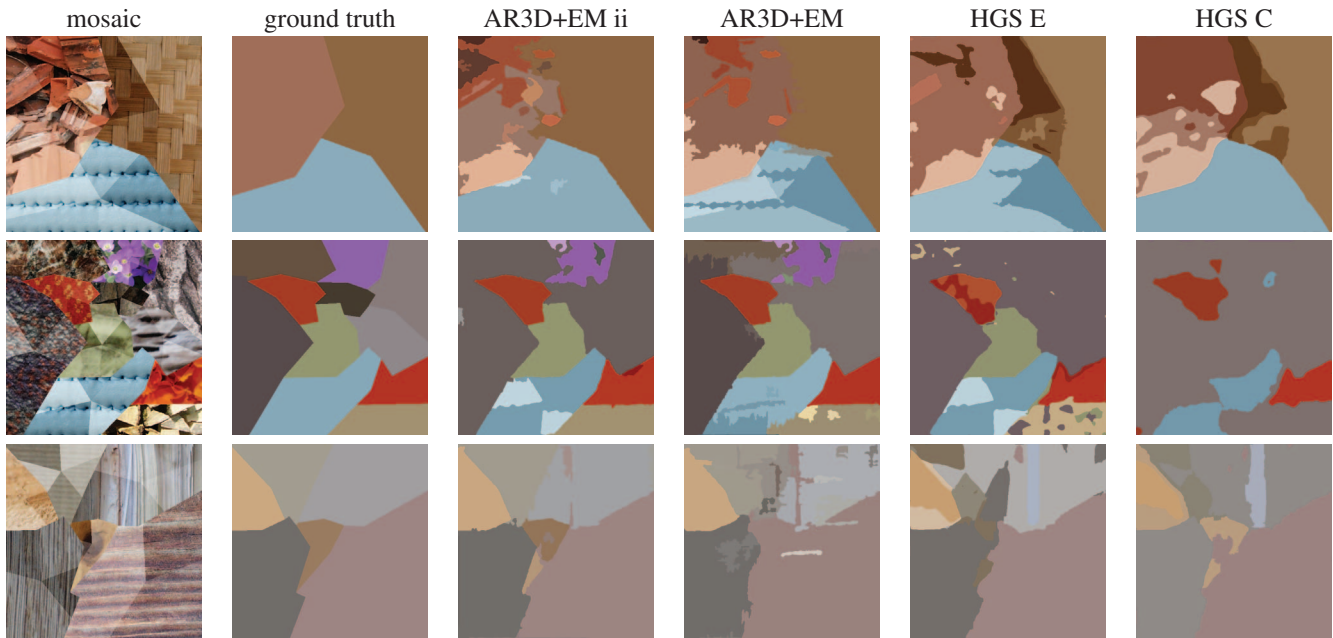
[2] R.L. Kashyap, "Image models," in *Handbook of Pattern*

| mosaic | ground truth | AR3D+EM ii | AR3D+EM | HGS E | HGS C |

Fig. 1. Selected experimental texture mosaics, ground truth from the benchmark and the corresponding segmentation results.

Recognition and Image Processing, K.S. Fu T.Y. Young, Ed. Academic Press, New York, 1986.

[3] M. Haindl, "Texture synthesis," *CWI Quarterly*, vol. 4, no. 4, pp. 305–331, December 1991.

[4] D.K. Panjwani and G. Healey, "Markov random field models for unsupervised segmentation of textured color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 10, pp. 939–954, 1995.

[5] B.S. Manjunath and R. Chellapa, "Unsupervised texture segmentation using markov random field models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 478–482, 1991.

[6] P. Andrey and P. Tarroux, "Unsupervised segmentation of markov random field modeled textured images using selectionist relaxation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 252–262, March 1998.

[7] Michal Haindl, "Texture segmentation using recursive markov random field parameter estimation," in *Proceedings of the 11th SCIA*, Lyngby, Denmark, June 1999, pp. 771–776, Pattern Recognition Society of Denmark.

[8] M. Haindl and S. Mikeš, "Model-based texture segmentation," *Lecture Notes in Computer Science*, , no. 3212, pp. 306 – 313, 2004.

[9] M. Haindl and S. Šimberová, *Theory & Applications of Image Analysis*, chapter A Multispectral Image Line Reconstruction Method, pp. 306–315, World Scientific Publishing Co., Singapore, 1992.

[10] Pavel Vacha and Michal Haindl, "Image retrieval measures based on illumination invariant textural mrf features," in *CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval*, New York, NY, USA, 2007, pp. 448–454, ACM Press.

[11] M. Haindl and S. Mikeš, "Texture segmentation benchmark," in *Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008*, Los Alamitos, December 2008, pp. 1–4, IEEE Computer Society.

[12] M. Haindl and S. Mikeš, "Unsupervised texture segmentation using multispectral modelling approach," in *Proceedings of the 18th International Conference on Pattern Recognition, ICPR 2006*, Los Alamitos, August 2006, vol. II, pp. 203–206, IEEE Computer Society.

[13] Minh A. Hoang, Jan-Mark Geusebroek, and Arnold W.M. Smeulders, "Color texture measurement and segmentation," *Signal Processing*, vol. 85, no. 2, pp. 265–275, 2005.