Natural Material Recognition with Illumination Invariant Textural Features

Pavel Vácha   Michal Haindl
Institute of Information Theory and Automation of the ASCR, 182 08 Prague, Czech Republic
{vacha,haindl}@utia.cz

***Abstract*—A visual appearance of natural materials fundamentally depends on illumination conditions, which significantly complicates a real scene analysis. We propose textural features based on fast Markovian statistics, which are simultaneously invariant to illumination colour and robust to illumination direction. No knowledge of illumination conditions is required and a recognition is possible from a single training image per material. Material recognition is tested on the currently most realistic visual representation - Bidirectional Texture Function (BTF), using the Amsterdam Library of Textures (ALOT), which contains 250 natural materials acquired in different illumination conditions. Our proposed features significantly outperform several leading alternatives including Local Binary Patterns (LBP, LBP-HF) and Gabor features.**

***Keywords*-texture; colour; Markov random field; illumination invariance;**

I. Introduction

Natural material recognition is an important subtask of many computer vision applications. Without description of material appearance structure (texture) the recognition is limited to diverse variants of colour histograms. A promising method for object/image recognition based on textural features was recently introduced [1]. Unfortunately, the appearance of natural materials is highly illumination and view angle dependent. As a consequence, most texture based classification or segmentation applications require multiple training images [2] captured under all available illumination and viewing conditions for each material class. Such learning is obviously clumsy and very often even impossible if the required measurements are not available. Although, unseen training images can be approximated to a certain extent using the photometric stereo [3], it requires at least three mutually registered images with different illumination direction for each material.

The normalisation cancelling lighting variations caused by the object geometry [4] completely wipes out rough texture structures with all their valuable discriminative information. It was demonstrated [5] that for a grey–scale image of an object with Lambertian reflectance there are no discriminative functions invariant to a change of illumination direction. Local Binary Patterns [6] (LBP) are popular illumination invariant features, but they are noise sensitive. This vulnerability was addressed [7], but used patterns are specifically selected according to the training set. Recently proposed LBP-HF [8] studies also relations between rotated patterns. Finally, the MR8 texton representation [2] was extended to be colour and illumination invariant [9].

We introduce efficient illumination invariant textural features based on Markov Random Fields (MRF). The proposed features extend the textural representation [10] with ten new illumination invariants. Moreover, we test the texture recognition on a very difficult, recently created ALOT texture library [9], where our method significantly outperforms alternative features.

II. Texture Representation

The texture is factorised into K levels of the Gaussian down-sampled pyramid and subsequently each pyramid level is modelled by a MRF type of model - either Causal Autoregressive Random (CAR) model or 2D Gaussian Markov Random Field model (GMRF). The model parameters are estimated and illumination invariants are computed from them. Finally, the illumination invariants from all the pyramid levels are concatenated into one feature vector.

A. Markov Random Field Models

Let us assume that each multispectral (colour) texture is composed of C spectral planes (usually $C = 3$), $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$ is the multispectral pixel at location r. The multiindex $r = (r_1, r_2)$ is composed of row index r_1 and column index r_2. The spectral planes are mutually decorrelated by the Karhunen-Loeve transformation (Principal Component Analysis) and subsequently modelled using either a 3-dimensional model or a set of C 2-dimensional models.

The 3-dimensional representation assumes that the multispectral texture pixel Y_r can be modelled as a linear combination of its neighbours:

$$Y_r = \gamma Z_r + \epsilon_r \ , \quad Z_r = [Y_{r-s}^T : \forall s \in I_r]^T \qquad (1)$$

where Z_r is the $C\eta \times 1$ data vector with multiindices r, s, t, $\gamma = [A_1, \ldots, A_\eta]$ is the $C \times C\,\eta$ unknown parameter matrix with square submatrices A_s. Similarly, the set of 2-dimensional models assumes that the j-th spectral plane of pixel at position r can be modelled as:

$$Y_{r,j} = \gamma_j Z_{r,j} + \epsilon_r \ , \quad Z_{r,j} = [Y_{r-s,j} : \forall s \in I_r]^T$$

where $Z_{r,j}$ is the $\eta \times 1$ data vector, $\gamma_j = [a_1, \ldots, a_\eta]$ is the $1 \times \eta$ unknown parameter vector. Some selected contextual neighbour index shift set is denoted I_r and $\eta =$

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.216

Figure 1. Example materials from ALOT dataset, rows differs in camera and light conditions.

$cardinality(I_r)$. GMRF and CAR models mutually differ in the correlation structure of the driving noise ϵ_r and in the topology of the contextual neighbourhood I_r (see [11] for details). As a consequence, all CAR model statistics can be efficiently estimated analytically [12], while the GMRF statistic estimates require either a numerical evaluation or some approximation.

The texture is analysed in a chosen direction, where multiindex t changes according to the movement on the image lattice I. Given the known CAR process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1, Z_t, Z_{t-1}, \dots, Z_1\}$ the parameter estimation $\hat{\gamma}$ can be accomplished using fast, numerically robust and recursive statistics [12]:

$$\begin{aligned} V_{t-1} &= \begin{pmatrix} \sum_{u=1}^{t-1} Y_u {Y_u}^T & \sum_{u=1}^{t-1} Y_u {Z_u}^T \\ \sum_{u=1}^{t-1} Z_u {Y_u}^T & \sum_{u=1}^{t-1} Z_u {Z_u}^T \end{pmatrix} + V_0 \\ &= \begin{pmatrix} V_{yy(t-1)} & V_{zy(t-1)}^T \\ V_{zy(t-1)} & V_{zz(t-1)} \end{pmatrix} , \\ \lambda_{t-1} &= V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1} V_{zy(t-1)} , \end{aligned}$$

where the positive definite matrix V_0 represents prior knowledge.

The parameter estimate of GMRF is approximated by pseudo-likelihood estimator, because either Bayesian or Maximum Likelihood estimate is computationally demanding. The noise ϵ_r covariance estimate is denoted $\hat{\Sigma}$.

B. Illumination Invariant Features

We assume that two images $\tilde{Y}, Y$ of the same texture and view position differing only in 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 transformation matrix. This linear formula is valid for changes in brightness and illumination spectrum, with surfaces including both Lambertian and specular reflectance.

Consequently, the model data vectors are related by $\tilde{Z}_r = \Delta\, Z_r$, where Δ is the $C\eta \times C\eta$ block diagonal matrix with blocks B on the diagonal. We derived [13] that $\tilde{A}_s = B\, A_s\, B^{-1}$, $\tilde{\lambda}_r = B\, \lambda_r B^T$, $\tilde{\hat{\Sigma}} = B\, \hat{\Sigma} B^T$, where matrices $\tilde{Z}_r, \tilde{A}_s, \tilde{\lambda}_r, \tilde{\hat{\Sigma}}$ are related to the model of the same texture with different illumination. Illumination invariants features were derived as well [13].

In this paper, we propose ten new illumination invariant features for the CAR model (the derivation will occur in follow-up paper):

1) $\beta_1 = \log\left(\frac{1}{r-t}|\lambda_r||\lambda_t|^{-1}\right)$,
2) $\beta_2 = \log\left(\frac{1}{r-t}|V_{zz(r)}||V_{zz(t)}|^{-1}\right)$,
3) $\beta_3 = \log\left(|V_{zz(r)}||\lambda_r|^{-\eta}\right)$,
4) $\beta_4 = \mathrm{tr}\left\{V_{yy(r)}\lambda_r^{-1}\right\}$.

Moreover, the prediction probability $p(Y_r|Y^{(r-1)})$ and the probability $p(Y^{(r)}|M)$ used in optimal model selection M (both derived in [12]) can be utilized. We define:

5) $\beta_5 = \log \sum_r \frac{1}{|I|} p(Y_r|Y^{(r-1)})$,
6) $\beta_6 = \log\left|\log p(Y^{(r)}|M)\right|$,

Instead of logarithm, we can use an alternative normalisation of invariants 1. – 4. based on geometric means:

7) $\beta_7 = \left(\frac{1}{r-t}|\lambda_r||\lambda_t|^{-1}\right)^{-2C}$,
8) $\beta_8 = \left(\frac{1}{r-t}|V_{zz(r)}||V_{zz(t)}|^{-1}\right)^{-2C\eta}$,
9) $\beta_9 = \left(|V_{zz(r)}||\lambda_r|^{-\eta}\right)^{-2C}$,
10) $\beta_{10} = \sqrt{|V_{yy(r)}||\lambda_r|^{-1}}$.

For the GMRF model, the invariants 1. – 4. and 7. – 10. can be defined using the same statistics with the following differences: λ have to be substituted with $\hat{\Sigma}$, $V_0 = O$ (zero matrix) and the absolute value of determinant $|V_{zz}|$ is used, because V_{zz} is not always positive definite in the GMRF model. The invariants 6. and 7. do not have GMRF counterparts.

In the case of 2D models, all the invariants are computed for each spectral plane separately. Feature vectors are formed from the introduced illumination invariants with $t = 0$, r equal to the last pixel position. Together with these invariants, we use the illumination invariants [13] with redefined $\nu_s = \mathrm{diag}(A_s)$, $\forall s \in I_r$ (the original definition as eigenvalues would cancel effect of the Karhunen-Loeve transformation). Although the features β_1 – β_{10} are rotation invariant, the overall representation is not, due to the included [13] features.

The dissimilarity of two feature vectors is computed using fuzzy contrast [14] in its symmetrical form FC_3 (see details in [10]). The fuzzy contrast includes normalisation

Table I
CORRECT CLASSIFICATION [%] ON THE ALOT TEXTURE LIBRARY. THE LAST COLUMN IS SIZE OF FEATURE VECTOR.

	experiment		
method	1	2	size
2D CAR-KL, β_1 – β_6	52.6	68.6	415
2D CAR-KL, β_5 – β_{10}	**53.9**	**69.0**	**415**
2D CAR-KL, β_5, β_6	51.6	68.1	355
2D CAR-KL, β_1 – β_{10}	**54.0**	68.9	475
3D CAR-KL, β_1 – β_{10}	50.3	68.0	345
3D CAR-KL, β_5 – β_{10}	50.2	68.1	325
GMRF-KL β_1 – β_4,β_7 – β_{10}	46.8	57.8	430
GMRF-KL β_7 – β_{10}	44.8	56.2	370
2D CAR-KL	**48.5**	**67.2**	325
3D CAR-KL	47.5	65.1	295
GMRF-KL	36.6	52.2	310
Gabor features, RGB	**44.6**	34.0	144
Opponent Gabor f.	41.8	**53.1**	252
$\text{LBP}_{8,1+8,3}$	32.8	39.8	512
$\text{LBP}_{8,1+8,3}$, RGB	**41.2**	45.6	**1536**
$\text{LBP}^{u}_{16,2}$, RGB	38.6	43.4	729
$\text{LBP}^{riu2}_{8,1+24,3}$, RGB	34.2	42.6	108
$\text{LBP-HF}_{8,1+24,3}$	32.6	**50.0**	340

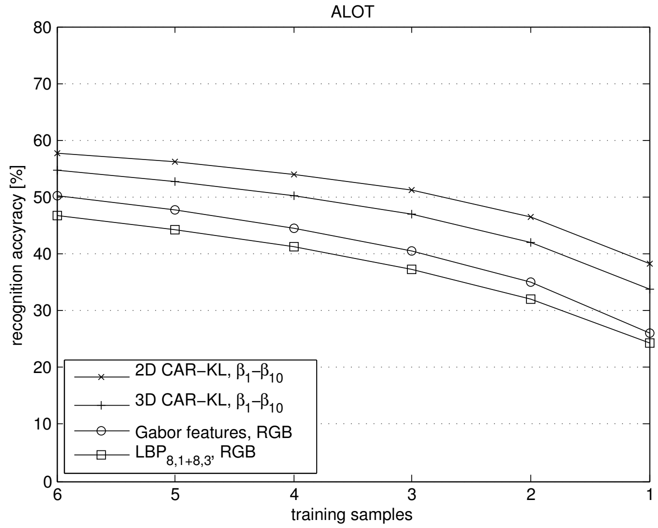

Figure 2. Correction classification [%] for different number of random training images per material in experiment 1.

of features, which is necessary for features with different scales. It requires estimates of mean and standard deviation of all features. The proposed MRF features were computed at $K = 5$ levels of the Gaussian pyramid, using the 6-th order hierarchical neighbourhood, which consists in $\eta = 14$ neighbours.

III. EXPERIMENTS

In the experiments, we tested the proposed features on the recognition of natural materials, which is needed in real scene analysis applications. We focused on the feature robustness under changing illumination spectrum and direction. Viewpoint changes are limited to the slant angle variation.

The material recognition was tested on the recently created Amsterdam Library of Textures (ALOT) [9]. The ALOT library is a BTF database containing an extraordinary collection of 250 natural materials, each acquired with varying viewpoint and illumination positions (Fig. 1). Most of the materials have rough surfaces, so the movement of light source changes the appearance of materials.

In the first experiment, we used one half of the dataset [9] to exclude multiple texture rotations. It contains images of the first 200 materials divided into parameter tuning, training, and test sets (3×1200 images). Let c stands for camera, l for light, i for reddish illumination, and r for optional material rotation. The tuning set consists in samples $c\{1,4\}l\{1,4,8\}r60°$; the training set is defined as $c\{1,4\}l\{1,4,8\}$ and the test set contains setups $c\{2,3\}l\{3,5\}$, $c3l2$, and $c1i$. We cropped all the images to the same size 1536×660 pixels. The classification was evaluated on the test set images, where the nearest neighbour (1-NN) classifier was trained on 4 images per material randomly selected from the training set. The features were computed from whole images.

In the second experiment, we used images of all 250 materials, with all light setups, no rotations and cameras 1 and 3, which is 14 images per material. One training image per material was randomly selected and the others were classified with the 1-NN classifier. This test was performed separately for images from camera 1 and 3, the results were averaged (2×1750 images in total). As a consequence this experiment do not include a recognition under viewpoint variation, which is in contrast with the experiment 1.

Our proposed features are compared with the most frequented features such as Gabor features [15], Opponent Gabor features [16], Local Binary Patterns [6] (LBP) and also recently published extension LBP-HF [8]. The grey level features as Gabor features and LBP were computed for each spectral plane separately and concatenated (denoted with “RGB” suffix in Tab. I), because the performance was better. The mean and standard deviation of features, which are required by the dissimilarity of our features and both Gabor features, were estimated on the parameter tuning set of experiment 1 and on all images in experiment 2.

A. Results

Both experiments were computed for 1000 random selections of training images and average classification results are shown in Tab. I. Standard deviations were below 0.5% and 1.4% for experiment 1 and 2, respectively. The best results were achieved with “2D CAR-KL, β_5 – β_{10}” model, which are 5%, 2% improvements to our previous model and 9%, 16% to alternative features. Tab. I also displays performance of “2D CAR-KL” model with different groups of the illumination invariants β_ℓ. Fig. 2 shows progression of correct classification in experiment 1, where the classifier was trained with descending number of random training images per material. Advantage of the proposed features was maintained for all numbers of training images.

Although the experiment 2 used only a single training image per material, the results are about 15% better than in the experiment 1. The reason is that the experiment 1 includes a viewpoint variation, which is even grater in the test set than in the training set. Moreover, the methodology in the experiment 2 produces an upper bound of correct classification, while the hold-out used in the experiment 1 yields a lower bound.

The proposed features were approximately $1.5\times$ slower than $\mathrm{LBP}^{riu2}_{8,1+24,3}$ and $4\times$ faster than Gabor features.

IV. Conclusion

We have introduced new illumination invariants derived from a MRF textural representation. The proposed textural features are fast to compute and they are invariant to illumination colour and brightness changes. Superiority of our features over LBP, LBP-HF and Gabor features was demonstrated in the recognition test with 250 natural materials acquired under varying viewpoint, illumination colour and direction. This make the proposed features suitable for Content Based Image Retrieval (CBIR) applications. [1]

In future research we will target the rotation invariance and integration into a CBIR system.

Acknowledgement

This research was supported by grants GAČR 102/08/0593 and partially by the MŠMT grants 1M0572 DAR, 2C06019.

References

[1] J. D. J. Shotton, J. Winn, C. Rother, and A. Criminisi, “Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context,” *International Journal of Computer Vision*, vol. 81, no. 1, pp. 2–23, Jan. 2009.

[2] M. Varma and A. Zisserman, “A statistical approach to texture classification from single images,” *International Journal of Computer Vision*, vol. 62, no. 1-2, pp. 61–81, 2005.

[3] A. T. Targhi, J.-M. Geusebroek, and A. Zisserman, “Texture classification with minimal training images,” in *Proceedings of the 19th International Conference on Pattern Recognition*, 2008.

[4] G. Finlyason and R. Xu, “Illuminant and gamma comprehensive normalisation in logrgb space,” *Patterm Recognition Letters*, vol. 24, pp. 1679–1690, 2002.

[5] H. F. Chen, P. N. Belhumeur, and D. W. Jacobs, “In search of illumination invariants,” in *IEEE Computer Vision and Pattern Recognition or CVPR*, 2000, pp. I: 254–261.

[6] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 24, no. 7, pp. 971–987, 2002.

[7] S. Liao, M. W. K. Law, and A. C. S. Chung, “Dominant local binary patterns for texture classification,” *IEEE Trans. Image Processing*, vol. 18, no. 5, pp. 1107–1118, May 2009.

[8] T. Ahonen, J. Matas, C. He, and M. Pietikäinen, “Rotation invariant image description with local binary pattern histogram fourier features,” in *SCIA*, ser. Lecture Notes in Computer Science, A.-B. Salberg, J. Y. Hardeberg, and R. Jenssen, Eds., vol. 5575. Springer, 2009, pp. 61–70.

[9] G. J. Burghouts and J. M. Geusebroek, “Material-specific adaptation of color invariant features,” *Pattern Recognition Letters*, vol. 30, pp. 306–313, 2009.

[10] P. Vacha and M. Haindl, “Illumination invariants based on markov random fields,” in *Proceedings of the 19th International Conference on Pattern Recognition*, dec. 2008, pp. 1 –4.

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

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

[13] P. Vacha and M. Haindl, “Image retrieval measures based on illumination invariant textural MRF features,” in *CIVR*, N. Sebe and M. Worring, Eds. ACM, 2007, pp. 448–454.

[14] S. Santini and R. Jain, “Similarity measures,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 21, no. 9, pp. 871–883, 1999.

[15] B. S. Manjunath and W. Y. Ma, “Texture features for browsing and retrieval of image data,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837–842, August 1996.

[16] A. Jain and G. Healey, “A multiscale representation including opponent colour features for texture recognition,” *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 125–128, January 1998.

[1] The interactive demonstration is available at http://cbir.utia.cas.cz/.