

Illumination Invariants Based on Markov Random Fields

Pavel Vácha Michal Haindl

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

Abstract

We propose textural features, which are invariant to illumination spectrum and extremely robust to illumination direction. They require only a single training image per texture and no knowledge of illumination direction or spectrum. Hence, these features are suitable for content-based image retrieval (CBIR) of realistic scenes with colour textured objects and variable illumination. The illumination invariants are derived from Markov random field based texture representations. Our illumination invariant features are favourably compared with frequented features in this area - the Local Binary Patterns, steerable pyramid and Gabor textural features, respectively. The superiority of our new invariant features is demonstrated in the illumination invariant recognition of the most advanced representation for realistic real-world materials - Bidirectional Texture Function (BTF) textures.

1. Introduction

Textures are important clues to specify objects present in a visual scene. Unfortunately, the appearance of natural textures is highly illumination and view angle dependent. As a consequence, most recent realistic texture based classification or segmentation methods require multiple training images [18] captured under all possible illumination and viewing conditions for each class. Such learning is obviously clumsy, probably expensive and very often even impossible if required measurements are not available.

Authors [2] allow a single training image per class, but they require uniform albedo surfaces and the knowledge of illumination direction. The normalisation canceling lighting variations caused by the object geometry [4] completely wipes out rough texture structures with all its valuable discriminative information. It suffers also with instability due to involved nonlinear trans-

formations. The above mentioned drawbacks are inevitable, because there is no discriminative function for grey images of objects with Lambertian reflectance that is illumination direction invariant [1]. The quasi-invariants [19] compromise full invariance to be less noise sensitive. Local Binary Patterns [13] (LBP) are popular illumination invariant features, but very noise sensitive [17]. Other options are parameters of Weibull-distribution of image edges [6] (six-stimulus theory), the logarithm of Gabor filter responses together with new Gaussian colour model [5], which is a subclass of illumination model we use.

We introduce textural features, which are robust to illumination direction changes. This property is verified on University of Bonn BTF texture measurements [12], where illumination sources are spanned over 75% of possible illumination half-sphere. We require only a single training image per texture and no knowledge of illumination direction. Our features are also invariant to illumination brightness and spectrum changes, and robust to Gaussian noise degradation [17]. In contrast with the similar test setup [9], we employ newly derived illumination invariants, which result in significant improvement even in more difficult arrangement (more than twice textures, half resolution, and different viewpoints) presented hereafter.

2. Texture Representation

Let us assume that each multispectral texture (composed of C spectral planes) can be modeled either by a 3-dimensional Markov random field (MRF) model or spectral planes can be mutually decorrelated by the Karhunen-Loeve transformation (Principal Component Analysis) and subsequently modeled using a set of C 2-dimensional MRF models. The texture is factorised into K levels of the Gaussian pyramid, so we can use lower order MRF factor models.

Let us denote a multiindex $r = (r_1, r_2)$, where r_1 is the row and r_2 the column index, respectively.

Y_r is multispectral pixel at location r and $Y_{r,j}$ is its j -th spectral plane component.

2.1. MRF Models

A multispectral texture factor (for the k th Gaussian pyramid level) is represented using either adaptive 3D / 2D causal autoregressive random (CAR) field model [8] or 2D Gaussian Markov random field model (GMRF) [7]. All these models can be unified in the following matrix equation form:

$$Y_r = \gamma Z_r + \epsilon_r, \quad (1)$$

where

$$Z_r = [Y_{r-s}^T : \forall s \in I_r]^T \quad (2)$$

is the $C\eta \times 1$ data vector with multiindices r, s, t , $\gamma = [A_1, \dots, A_\eta]$ is the $C \times C\eta$ unknown parameter matrix with submatrices A_s . In the case of C 2D CAR / GMRF models stacked into the model equation (1) the parameter matrices A_s are diagonal otherwise they are full matrices for general 3D CAR models. The contextual neighbour index shift set is denoted I_r and $\eta = \text{cardinality}(I_r)$. GMRF and CAR models mutually differ in the correlation structure of the driving noise ϵ_r (1) and in the topology of the contextual neighbourhood I_r (see [7] for details). As a consequence, all CAR model statistics can be efficiently estimated analytically [8] while the GMRF statistics estimates require either numerical evaluation or some approximation.

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 [8]:

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

where V_0 is a positive definite matrix (see [8]).

2.2. Illumination Invariant Features

We assume fixed positions of viewpoint and illumination sources, the illumination sources are supposed to be far enough to produce uniform illumination. Based on the work [3], assuming Lambertian surface reflectance, the two images \tilde{Y}, Y acquired with different illumination spectra can be transformed to each other:

$$\tilde{Y}_r = B Y_r \quad \forall r, \quad (3)$$

where \tilde{Y}_r, Y_r are multispectral pixel values at position r and B is a $C \times C$ transformation matrix. The linear formula (3) is also valid for more illumination sources provided that the spectra of all sources are the same and the positions of the illumination sources remain unchanged. More importantly, it can be proved that formula (3) is able to model specular reflectance component (e.g. dichromatic reflection model [15], which comprise well-known Phong model).

With the previous assumptions, the following illumination invariant features 1. 2. 6. 7. can be derived from the estimated MRF statistics for GMRF models with centered $Y_{r,j}$, while 1. – 5. are CAR invariants. In the case of 2D models, invariants 3. – 7. are computed for each spectral plane separately.

1. trace: $\text{tr } A_m, m = 1, \dots, \eta K$
2. eigenvalues: $\nu_{m,j}$ of $A_m, m = 1, \dots, \eta K, j = 1, \dots, C$
3. $\alpha_1: 1 + Z_r^T V_{zz}^{-1} Z_r$,
4. $\alpha_2: \sqrt{\sum_r (Y_r - \hat{\gamma} Z_r)^T \lambda^{-1} (Y_r - \hat{\gamma} Z_r)}$,
5. $\alpha_3: \sqrt{\sum_r (Y_r - \mu)^T \lambda^{-1} (Y_r - \mu)}$, μ is the mean value of vector Y_r ,
6. $\alpha_4: \sqrt{\sum_r \hat{\sigma}_j^{-2} (Y_{r,j} - \hat{\gamma}_j Z_{r,j})^2}$,
7. $\alpha_5: \sqrt{\sum_r \hat{\sigma}_j^{-2} (Y_{r,j})^2}$.

Feature vectors are formed from these illumination invariants, which are easily evaluated during the MRF parameters estimation process.

The distance between feature vectors of two textures T, S is computed using the Minkowski norms $L_1, L_{0.2}$, or alternatively with fuzzy contrast [14] in its symmetrical form FC_3 :

$$\begin{aligned} L_p(T, S) &= \left(\sum_{i=0}^M |f_i^{(T)} - f_i^{(S)}|^p \right)^{\frac{1}{p}}, \\ FC_p(T, S) &= M - \left\{ \sum_{i=1}^M \min \left\{ \tau(f_i^{(T)}), \tau(f_i^{(S)}) \right\} \right. \\ &\quad \left. - p \sum_{i=1}^M \left| \tau(f_i^{(T)}) - \tau(f_i^{(S)}) \right| \right\}, \\ \tau(f_i) &= \left(1 + \exp \left(- \frac{f_i - \mu(f_i)}{\sigma(f_i)} \right) \right)^{-1}, \end{aligned}$$

where M is the feature vector size and $\mu(f_i)$ and $\sigma(f_i)$ are average and standard deviation of the feature f_i computed over all database, respectively. The sigmoid function τ models the truth value of fuzzy predicate.

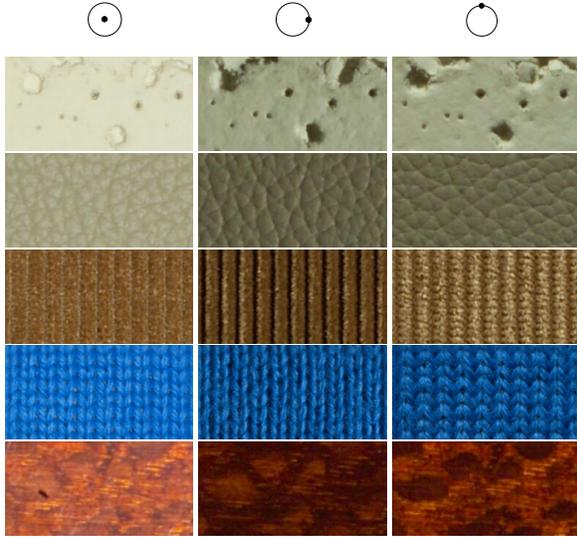


Figure 1. Effects of illumination direction changes on BTF material examples: ceiling, foil, corduroy, wool, lacquered wood.

3. Experiments

We have designed our experiments to test the method robustness to illumination direction changes, which drastically violates our previous theoretical assumptions. The experiments are performed on three different sets of 1215 BTF texture images. These BTF data are from the University of Bonn database [12] and contain BTF colour measurements such as ceiling, corduroy, two fabrics, walk way, foil, floor tile, pink tile, impalla, proposte, pulli, wallpaper, wool, and two lacquered wood textures. Each BTF material sample (Fig. 1) is measured in 81 illumination angles as a RGB image ($C = 3$). Each of our three test sets has fixed viewpoint direction, the declination angles of viewing direction from surface normal are 0° , 30° , and 60° , in plane rotation is not included. All images were cropped to the same size 256×256 .

3.1. Compared Features

Our proposed features are compared with the most popular features such as the Gabor features, steerable pyramid features and Local Binary Patterns (LBP).

The Gabor features [11], which are computed from responses of Gabor filters, were computed separately for each spectral plane and concatenated into the feature vector. The Opponent Gabor features [10], which are extension to colour textures, analyses relations be-

Table 1. Size of feature vectors.

method	size	method	size
Gabor features	144	2D CAR-KL	260
Opponent Gabor	252	GMRF-KL	248
Steerable pyramid	2904	3D CAR	236
$LBP_{8,1+8,3}$, grey	512	$LBP_{8,1+8,3}$	1536
$LBP_{16,2}^{u2}$, grey	243	$LBP_{16,2}^{riu2}$	54

tween spectral channels. We have also experimented with brightness normalisation prior to feature computation (denoted as “nm.”). Both Gabor feature vectors are compared with the author’s suggested norms ([11, 10]).

Steerable pyramid features are statistics of Steerable pyramid decomposition, which were proposed for texture synthesis in [16]. The feature vectors are compared with the same norm as Gabor features.

LBP [13] are histograms of texture micro patterns, which are thresholded values at each pixel neighbourhood. Subsequently, the histograms are compared using Kullback-Leibler divergence as suggested in [13]. We have tested features: $LBP_{8,1+8,3}$ and $LBP_{16,2}^{u2}$, which were published with best performance under different illuminations (Outex database) and rotation invariant features $LBP_{16,2}^{riu2}$. The features were computed either on grey-scale images or on each spectral plane separately and concatenated into the feature vector. Brightness normalisation is not necessary.

The proposed MRF features were computed at $K = 4$ levels of Gaussian pyramid, using the 6-th order hierarchical neighbourhood, which consists in $\eta = 14$ neighbours. The size of feature vectors used is listed in Tab. 1.

3.2. Results

In our experiments, single training image per each material was randomly chosen and the remaining images were classified using the nearest neighbour approach. The results in Tab. 2 are averages over 10^5 random choices of training images and the last column consists in averages of previous columns. We can see that the far best performance (90.3%) was achieved with 2D CAR-KL model and L_1 distance. The best alternative features were Opponent Gabor features with average performance 77.4%, the best of LBP features achieved 65.6%. Although the LBP feature are invariant to brightness changes, these results demonstrate their inefficiency to handle illumination direction variations. Rotation invariant LBP features are more capable, however rotating illumination cannot be modeled as a simple image rotation.

Table 2. Correct classification [%], using one random training image per texture.

method	viewpoint declination angle			
	0°	30°	60°	avg.
Gabor features.	71.7	64.6	60.1	65.5
Opponent Gabor	82.5	77.7	71.7	77.3
Steerable pyramid	72.3	65.5	60.4	63.1
Opponent Gabor, nm.	80.5	77.6	74.2	77.4
LBP _{8,1+8,3} , grey	61.2	61.1	65.4	62.6
LBP _{16,2} ^u , grey	55.7	56.3	60.7	57.6
LBP _{8,1+8,3}	65.7	64.2	67.0	65.6
LBP _{16,2} ^{riu}	68.4	60.7	57.4	62.2
2D CAR-KL, L₁	92.4	91.1	87.5	90.3
3D CAR, L ₁	87.4	84.3	78.9	83.5
GMRF-KL, L ₁	89.6	86.3	81.0	85.6
GMRF-KL, FC ₃	86.5	82.6	78.7	82.6
GMRF-KL, L _{0,2}	87.1	83.7	79.6	83.5
2D CAR-KL, FC ₃	92.3	89.6	85.7	89.2
2D CAR-KL, L _{0,2}	91.8	89.5	85.8	89.0
3D CAR, FC ₃	89.8	86.1	80.2	85.4
3D CAR, L _{0,2}	89.2	85.7	81.0	85.3

4. Conclusions

We have introduced MRF based features which are invariant to brightness and illumination spectrum variations and they are considerably robust to illumination direction changes. Moreover, we have demonstrated their superiority over Gabor or LBP features, which have significant difficulties to correctly recognize textures under changing illumination direction. In further research we will target rotation invariance and integration into an image retrieval system.

5. Acknowledgements

This research was supported by grants GAAV 1ET400750407, GAČR 102/08/0593 and partially by the MŠMT grants 1M0572 DAR, 2C06019.

References

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

[2] O. Drbohlav and M. Chantler. Illumination-invariant texture classification using single training images. In

Texture 2005: Proceedings of the 4th international workshop on texture analysis and synthesis, pages 31–36, 2005.

[3] G. D. Finlayson. *Coefficient color constancy*. PhD thesis, Simon Fraser University, 1995.

[4] G. Finlyason and R. Xu. Illuminant and gamma comprehensive normalisation in $\log r g b$ space. *Pattern Recognition Letters*, 24:1679–1690, 2002.

[5] J.-M. Geusebroek, R. v. d. Boomgaard, A. W. Smeulders, and T. Gevers. Colour constancy from physical principles. *Pattern Recognition Letters*, 24:1653–1662, 2003.

[6] J.-M. Geusebroek and A. W. Smeulders. A six-stimulus theory for stochastic texture. *International Journal of Computer Vision*, 62:7–16, 2005.

[7] M. Haindl. Texture synthesis. *CWI Quarterly*, 4(4):305–331, December 1991.

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

[9] M. Haindl and P. Vacha. Illumination invariant texture retrieval. In *International Conference on Pattern Recognition*, pages III: 276–279, 2006.

[10] A. Jain and G. Healey. A multiscale representation including opponent colour features for texture recognition. *IEEE Transactions on Image Processing*, 7(1):125–128, January 1998.

[11] W. Y. Ma and B. S. Manjunath. Texture features and learning similarity. *IEEE*, pages 425–430, 1996.

[12] J. Meseth, G. Müller, and R. Klein. Preserving realism in real-time rendering. In D. Reiners, editor, *OpenGL Symposium*, pages 89–96. Eurographics Association, Switzerland, April 2003.

[13] 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.*, 24(7):971–987, 2002.

[14] S. Santini and R. Jain. Similarity measures. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(9):871–883, 1999.

[15] S. A. Shafer. Using color to separate reflection components. *COLOR research and application*, 10(4):210–218, Winter 1985.

[16] E. Simoncelli and J. Portilla. Texture characterization via joint statistics of wavelet coefficient magnitudes. In *Fifth IEEE Int’l Conf on Image Proc.*, volume I, Chicago, 4-7 1998. IEEE Computer Society.

[17] P. Vacha and M. Haindl. Image retrieval measures based on illumination invariant textural MRF features. In N. Sebe and M. Worring, editors, *CIVR*, pages 448–454. ACM, 2007.

[18] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

[19] J. Weijer, T. Gevers, and J. Geusebroek. Edge and corner detection by photometric quasi-invariants. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 27(4):625–630, 2005.