ILLUMINATION INVARIANT AND ROTATIONAL INSENSITIVE TEXTURAL REPRESENTATION

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 an illumination invariant and rotation insensitive texture representation based on a Markovian textural model. A texture is aligned with its dominant orientation and textural features are derived from fast analytical estimates of Markovian statistics. We do not require any knowledge of illumination direction or spectrum. This makes our method suitable for computer analysis of real scenes, where appearance of materials depends on their orientation towards the illumination source. Our method is tested on the most realistic visual representation of natural materials - the bidirectional texture function (BTF), using data from the CUReT database, where it outperforms the alternative leading illumination invariant Local Binary Patterns (LBP) and texton MR8 methods, respectively.

Index Terms— Illumination invariance, Markov random fields, texture features

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 real materials texture based classification or segmentation applications require multiple training images [1] captured under all possible illumination and viewing conditions for each material class. Such learning is obviously clumsy, expensive and very often even impossible if required measurements are not available.

Even though Drbohlav and Chantler [2] allow a single training image per class, they require uniform albedo surfaces and the knowledge of illumination direction. The normalisation canceling lighting variations caused by the object geometry [3] completely wipes out rough texture structures with all its valuable discriminative information. It was demonstrated [4] that for a grey–scale image of an object with Lambertian reflectance and missing surface material interreflections there

are no discriminative functions that are invariant to change of illumination direction. Local Binary Patterns [5] (LBP) are popular illumination invariant features, but too noise sensitive [6]. The rotation invariant texton representation [1] based on MR8 filter responses have been extended to incorporate colour information and to be illumination invariant [7]. Another approach [8] generates unseen training images using the photometric stereo, which requires three mutually registered images with different illumination direction for each material. Rotation invariance [9] and normalisation [10] was proposed without illumination invariance.

We introduce the efficient illumination invariant multispectral texture representation, which is simultaneously insensitive to texture rotation. The recognition accuracy is tested on the CUReT texture database [11], which includes samples with varying illumination direction and also limited viewpoint changes. The employed features are simultaneously invariant to illumination brightness and spectrum changes, and robust to Gaussian noise degradation [6].

2. TEXTURE ANALYSIS

The scheme of texture analysis algorithm is depicted in Fig. 1. We start with the estimation of the dominant texture orientation. If the texture is significantly directional, it is rotated according to its dominant orientation. The texture is factorised into K levels of the Gaussian pyramid and subsequently each pyramid level is modelled by a Markov Random Field type of model - the Causal Autoregressive Random (CAR) model. The CAR model parameters are estimated and illumination invariants are subsequently computed from them. Finally, the illumination invariants from all the models are concatenated into one feature vector.

2.1. Orientation normalisation

The dominant texture orientation is estimated from the histogram of gradient orientations. Similar algorithm is used to determine orientation of SIFT keypoints [12]. The purpose of this preprocessing is detection of textures with strong orientations and their rotation alignment. The rotation of tex-
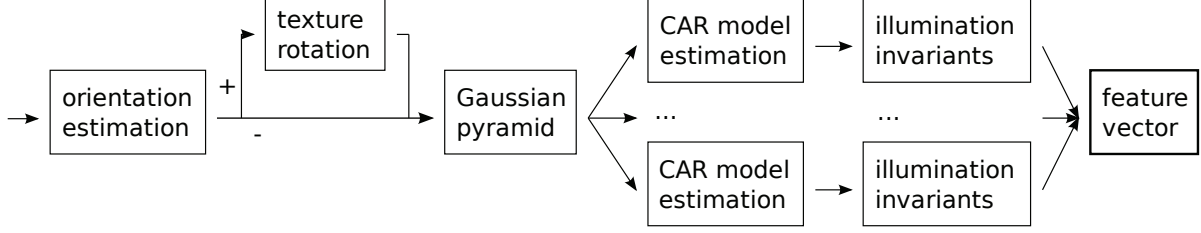
Fig. 1. Texture analysis algorithm

tures with less significant directions or undirectional textures is not required, because they are aptly represented by the CAR model.

The input texture is converted into greyscale image G and its gradient $\nabla G(r_1, r_2) = [G(r_1 + 1, r_2) - G(r_1 - 1, r_2), G(r_1, r_2+1) - G(r_1, r_2-1)]$ is computed at each pixel $r = (r_1, r_2)$ where the first component r_1 of the multiindex is the row and r_2 is the column index, respectively.

Subsequently, histogram of gradient orientations is computed. Each gradient orientation is weighted by its magnitude and bilinear interpolation is used to assign gradient weight into two adjacent bins. All orientations are fitted into the interval $[0°, 180°)$. The reason is homogeneity of textures, where gradients in some direction are followed by gradients in the opposite direction. We use histogram with 36 bins and 6 pass histogram smoothing (averaging of three adjacent bins).

The texture is considered to have a dominant orientation if the height of the second highest peak in the histogram is lower than 80% of the highest peak. If there is no second highest peak, the sum of the highest peak bin and its two adjacent bins have to be grater than 150% of the expected value for three bins.

2.2. CAR Model

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

The CAR representation assumes that the multispectral texture pixel Y_r can be modelled as linear combination of its neighbours:

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

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 submatrices A_s. In the case of C 2D CAR models stacked into the model equation (1) the parameter matrices A_s are diagonal otherwise they are full matrices for general 3D CAR models. Some selected contextual

causal or unilateral neighbour index shift set is denoted I_r and $\eta = cardinality(I_r)$. The white noise vector ϵ_r has normal density with zero mean and unknown constant covariance matrix, same for each pixel. Additionally for 2D CAR model, we assume uncorrelated noise vector components. Given the known CAR 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 [13]:

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

where V_0 is a positive definite matrix (see [13]).

2.3. Illumination Invariant Features

We assume that the 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 Lambertian surface reflectance, or with model including specular reflectance component (e.g. dichromatic reflection model [14]). With the previous assumptions, the following illumination invariant features were derived [6]:

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

2. eigenvalues: $\nu_{m,j}$ of $A_m, \ m = 1, \ldots, \eta\, K,$ $j = 1, \ldots, C$

3. α_1: $1 + Z_r^T V_{zz}^{-1} Z_r$,

4. α_2: $\sqrt{\sum_r (Y_r - \hat{\gamma} Z_r)^T \lambda^{-1} (Y_r - \hat{\gamma} Z_r)}$,

5. α_3: $\sqrt{\sum_r (Y_r - \mu)^T \lambda^{-1} (Y_r - \mu)}$, μ is the mean value of vector Y_r,
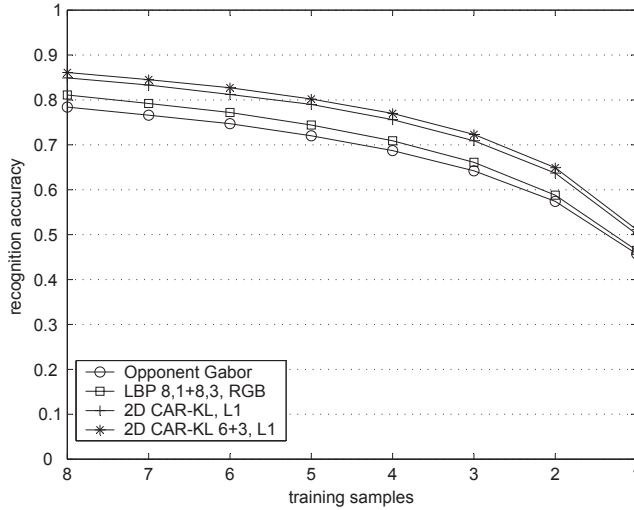
Fig. 2. Classification accuracy for different number of random training images, mean values were computed over 1000 repetitions.

Table 1. Classification accuracy [%], using 4 random training images per texture.

method	performance	features
[7] MR8	58	600
[7] MR8-NC	54	600
[7] MR8-INC	60	600
[7] MR8-LINC	**67**	600
[7] MR8-SLINC	57	600
Gabor features	61.7	144
Opponent Gabor features	**68.7**	252
$LBP_{8,1+8,3}$, grey	66.9	512
$LBP_{8,1+8,3}$, RGB	**70.9**	**1536**
$LBP_{16,2}^{u}$, RGB	68.7	729
$LBP_{16,2}^{riu2}$, RGB	64.2	54
$LBP_{8,1+8,3}$, opponent	57.4	1536
$LBP_{16,2}^{u}$, opponent	69.7	729
2D CAR-KL, L_1	**75.6**	260
2D CAR-KL, FC_3	75.1	260
2D CAR-KL 6+3, L_1	**77.0**	**392**
2D CAR-KL 6+3, FC_3	77.6	392
3D CAR, L_1	69.7	236
3D CAR, FC_3	67.6	236
3D CAR 6+3, L_1	**72.4**	344
3D CAR 6+3, FC_3	72.6	344

Feature vectors are formed from these illumination invariants, which are easily evaluated during the CAR parameters estimation process. In the case of 2D models, invariants $3. - 5.$ are computed for each spectral plane separately. We have also experimented with combination of two models, where each factor of Gaussian pyramid is modelled by two models with different neighbourhood I_r. In that case, illumination invariants for both models are included in the final feature vector.

The distance between two feature vectors is computed using the L_1 norm or alternatively with fuzzy contrast [15] in its symmetrical form FC_3 (see details in [16]). However, the fuzzy contrast requires mean and standard deviation of each feature, which was estimated on the whole image database.

3. RESULTS

In the experiments, we focus on robustness of proposed texture representation under varying illumination direction and also limited viewpoint changes, which are compensated by texture rotation. These conditions are quite close to real world, where materials in a scene are observed under different viewpoint and illumination conditions.

We evaluated the texture recognition accuracy on the Columbia-Utrecht Reflectance and Texture Database (CUReT) [11] to be able to compare our results with alternative approaches even if this database is already overcome by BTF measurements from the University of Bonn. This first BTF database consists of 61 real-world materials captured under different combinations of viewing and illumination directions. Dataset provided by Varma and Zisserman [1] consists of 61 materials, each with 92 samples with resolution

200×200 pixels. We follow the experimental setup [7], where the classification accuracy is tested with randomly selected training samples and the SVM classifier. The number of training samples per material decreases from 8 to 1. The mean and standard deviation of classification accuracy is computed over 1000 repetitions (random selections). On the contrary, we use only simple nearest neighbour classifier.

Additionally to rotation and illumination invariant MR8 variants [7], we have compared in Tab.1 the performance of some other most frequented features. The Gabor features [17], which are rotation variant statistics of Gabor filters responses, were computed separately for each spectral plane and concatenated into the feature vector. The Opponent Gabor features [18] are extension that analyse relations between spectral channels. The distances for Gabor features requires standard deviation of features, which have been estimated on the whole image database. Local Binary Patterns [5] (LBP) are histograms of thresholded micro patterns, we have tested variants $LBP_{8,1+8,3}$, $LBP_{16,2}^{u2}$ and rotation invariant $LBP_{16,2}^{riu2}$. They were computed either on grey-scale images or on each spectral plane of RGB colour space or opponent colour space.

The CAR features were computed for $K = 4$ levels of Gaussian pyramid, using the 6-th order hierarchical neighbourhood. For combination of two models we use 6-th and 3-th order neighbourhoods, which consist in $\eta = 14$ and $\eta = 6$ neighbours, respectively.

Fig. 2 shows mean recognition rates over 1000 random repetitions. It is directly comparable to results [7], where recognition rate for MR8-LINC monotonously decreases, approximately, from 75% to 45% for 8 to 1 training samples. Our best performance, which is combination of two models (2D CAR-KL 6+3), goes from 86% to 51% with standard deviations from 0.6% to 1.5% .

The exact mean recognition rates together with size of feature vectors are displayed in Tab. 1, standard deviations are below 1%. The best performance 77% was again achieved with 2D CAR-KL 6+3 models, followed by 2D CAR-KL model with 75.6%, both with L_1 distance. The best alternative features were $LBP_{8,1+8,3}$ with average performance 70.9% and 4 times larger feature vector.

4. CONCLUSIONS

We have compared the proposed illumination invariant textural representation on the classification of textures captured under varying illumination directions.[1] These BTF textures represent visual properties of sixty different real-world materials. Our illumination invariants are efficiently computed Markovian texture statistics. The overall method is insensitive to texture rotation and uses low number of features. The classifier can be learned from only one training image per texture and requires no knowledge of illumination direction or the spectrum. The classifier was tested on the database CUReT, where it outperforms alternative leading illumination invariant LBP and texton MR8 methods.

5. REFERENCES

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

[2] O. Drbohlav and M. Chantler, "Illumination-invariant texture classification using single training images," in *Texture 2005: Proc. of the 4th international workshop on texture analysis and synthesis*, 2005, pp. 31–36.

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

[4] 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.

[5] 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.

[6] P. Vacha and M. Haindl, "Image retrieval measures based on illumination invariant textural MRF features," in *CIVR*, Nicu Sebe and Marcel Worring, Eds. 2007, pp. 448–454, ACM.

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

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

[9] R. L. Kashyap and A. Khotanzad, "A model-based method for rotation invariant texture classification," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 8, pp. 472–481, 1986.

[10] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Radon transform orientation estimation for rotation invariant texture analysis," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 27, no. 6, pp. 1004–1008, 2005.

[11] K. J. Dana, B. Van-Ginneken, S.K. Nayar, and J. J. Koenderink, "Reflectance and Texture of Real World Surfaces," *ACM Trans. on Graphics (TOG)*, vol. 18, no. 1, pp. 1–34, Jan 1999.

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[13] 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.

[14] S. A. Shafer, "Using color to seperate reflection components," *COLOR research and application*, vol. 10, no. 4, pp. 210–218, Winter 1985.

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

[16] P. Vacha and M. Haindl, "Illumination invariants based on markov random fields," in *Proc. of the 19th International Conference on Pattern Recognition*, 2008.

[17] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 18, no. 8, pp. 837–842, August 1996.

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

---

[1]Demonstrations are at http://cbir.utia.cas.cz, http://ro.utia.cz/dem.html