Pattern Recognition Letters 32 (2011) 771–779

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Colour and rotation invariant textural features based on Markov random fields

Pavel Vácha, Michal Haindl *, Tomáš Suk

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

ARTICLE INFO

Article history:
Received 28 January 2010
Available online 7 January 2011
Communicated by Y.J. Zhang

Keywords:
Image modelling
Colour
Texture
Markov random field
Illumination invariance
Rotation invariance

ABSTRACT

A visual appearance of natural materials significantly depends on acquisition circumstances, particularly illumination conditions and viewpoint position, whose variations cause difficulties in the analysis of real scenes. We address this issue with novel texture features, based on fast estimates of Markovian statistics, that are simultaneously rotation and illumination invariant. The proposed features are invariant to in-plane material rotation and illumination spectrum (colour invariance), they are robust to local intensity changes (cast shadows) and illumination direction. No knowledge of illumination conditions is required and recognition is possible from a single training image per material. The material recognition is tested on the currently most realistic visual representation – Bidirectional Texture Function (BTF), using CUReT and ALOT texture datasets with more than 250 natural materials. Our proposed features significantly outperform leading alternatives including Local Binary Patterns (LBP, LBP-HF) and texton MR8 methods.

1. Introduction

Texture and colour are two major visual features used in most low level image processing applications. Although a notion of texture is tied to the human semantic meaning, there is no mathematically rigorous definition of texture that would be widely accepted by authors. We understand a textured image to be a realization of random field and our effort is simply to find rotation and illumination invariant statistics for its reliable representation.

Although textures were found "ill-suited for retrieval applications in which the user wants to use verbal descriptions of the image" at the end of the image retrieval early years (Smeulders et al., 2000), this was rather a direct consequence of inferior textural features than the textures themselves. Recent achievements (Filip and Haindl, 2009) in the high quality Bidirectional Texture Function (BTF) representation together with a steadily growth of an image sensor resolution quite contrary emphasize textural approach to the image retrieval. A promising method based on textural features was recently introduced by Shotton et al. (2009). Still, recent image retrieval methods suffer from a number of drawbacks and their error rate is too high for a routine operation. Other computer vision tasks, which can benefit from textural features, are object (image) recognition and real scene segmentation.

The invariance to image acquisition conditions is an important aspect of computer vision applications. Without sufficient classes of invariance the applications require multiple training images captured under a full variety of possible illumination and viewing conditions. The reason is that the appearance of natural materials significantly varies with changing illumination and viewing directions (see example images in Figs. 1 and 2 or Dana et al., 1999). An attempt to learn from all possible appearances is obviously clumsy, expensive and very often even impossible if the required measurements are not available. On the other hand, an undesired invariance to a broad range of conditions usually reduces the discriminability and aggravates the recognition. Diplaros et al. (2006) proposed illumination-viewpoint invariant recognition using combination of colour and shape. Burghouts and Geusebroek (2009b) extended the grey-scale rotation invariant SIFT descriptor (Lowe, 2004) to incorporate colour invariants. A texture segmentation can benefit from illumination invariants as well (Haindl et al., 2009), although it is not rotation invariant.

Rotation invariant textural features can be divided into two main groups. The first group contains filter based features such as modified Gabor features (Haley et al., 1999) and circular-Mellin features (Ravichandran and Trivedi, 1995), which have additional advantage of being scale insensitive. The second group composes of model based features, which were in the invariant form introduced by Kashyap et al. (1986), who used an autoregressive model of pixel value and averages on concentric circles around it. The disadvantage of this model and its multiresolution extension (Mao and Jain, 1992) is the insensitivity to anisotropic texture properties. This weakness was removed in Anisotropic Circular Gaussian Markov Random Field (ACGMRF) model (Deng et al., 2004), which computes the Fourier descriptors of estimated model parameters. However, all the previously mention textural features are illumination variant.

Another possibility to deal with rotated textures is a rotation normalisation (Jafari-Khouzani and Soltanian-Zadeh, 2005), where the principal texture direction is detected by the Radon transfor-

* Corresponding author. Tel.: +420 26605 2350; fax: +420 28468 3031.
E-mail addresses: vacha@utia.cas.cz (P. Vácha), haindl@utia.cas.cz (M. Haindl), suk@utia.cas.cz (T. Suk).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.01.002

Fig. 1. Example materials from the ALOT dataset and their appearance for different camera and light conditions.

Fig. 2. Example materials from the ALOT dataset and their appearance for different camera and light conditions.

mation. Still, a detection of the principal direction may be ambiguous and it can be influenced by the illumination direction.

Combined rotation and illumination invariance is the property of the following Local Binary Patterns (LBP) and texton MR8 features. The rotation invariant version LBP^{riu2} of popular LBP features was introduced by Ojala et al. (2002a). In general, LBP are histograms of thresholded micropatterns, which guarantee invariance to monotonous brightness changes. Unfortunately, they are sensitive to noise and illumination direction changes (Vacha and Haindl, 2007; Vacha and Haindl, 2008). The vulnerability to noise was addressed by Liao et al. (2009), who selected dominant LBP and combined it with circular Gabor features to include larger texture relations. However, this method specifically selects patterns according to the training set. A recently proposed LBP-HF (Ahonen et al., 2009) extended LBP^{riu2} with relations between rotated patterns. Varma and Zisserman (2005) characterize the texture by a histogram of textons, which are clusters of MR8 filter responses. Although the rotation invariance is provided by using the maximal response of each filter orientations, the method requires multiple training images with a full variety of illumination conditions for each class. This drawback was reduced by Burghouts and Geusebroek (2009a), who extended the MR8 texton method to include various degrees of illumination invariance – from colour invariance to cast shadows and shading.

We introduce novel textural features, which are simultaneously rotation and illumination invariant. They benefit from Markov random field (MRF) illumination invariants (Vacha and Haindl, 2007), which were derived with a restrictive assumption of Lambertian surfaces. In this article (Section 3), we extend the invariance to natural illumination models and local intensity changes (e.g. cast shadows). More importantly, we present two methods for adding rotation invariance and combine these methods to take advantage of their different approaches (Section 4). In Section 6, the proposed features are tested on two BTF databases: Columbia–Utrecht Reflectance and Texture Database (CUReT) (Dana et al., 1999) and recently created Amsterdam Library of Textures (ALOT) (Burghouts and Geusebroek, 2009a), which contains 250 natural materials acquired with varying viewpoint and illumination position.

2. Regression model

We use a textural representation based on fast estimates of Markovian statistics. A texture is modelled by an autoregressive random field model and the model parameters become the texture characterisation. The special wide sense Markov model is used, because it enables fast analytical estimate of its parameters and thus to avoid time-consuming Monte Carlo minimisation prevailing in most MRF models.

2.1. Causal autoregressive random field

Let us assume that each multispectral (e.g. colour) texture is composed of C spectral planes measured by the corresponding sensors (usually $C = 3$). $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$ is a multispectral pixel at location r, where $r = [r_1, r_2]$ is a multiindex with r_1 row and r_2 column index, respectively. The spectral planes are either modelled by a 3-dimensional causal autoregressive random (CAR) field model or mutually decorrelated by the Karhunen–Loève 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 a linear combination of its neighbours:

$$Y_r = \gamma Z_r + \epsilon_r, \quad Z_r = \left[Y_{r-s}^T : \forall s \in I_r^u \right]^T, \tag{1}$$

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 a general 3D CAR model. Some selected contextual causal or unilateral neighbour index shift set is denoted I_r^u and

$\eta = cardinality(I_r^u)$. The white noise vector ϵ_r has the normal density with zero mean and unknown constant covariance matrix, same for each pixel. For the 2D CAR model, we additionally 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 (Haindl and Šimberová, 1992):

$$\hat{\gamma}_{t-1}^T = V_{zz(t-1)}^{-1} V_{zy(t-1)}, \tag{2}$$

$$V_{t-1} = \tilde{V}_{t-1} + V_0,$$

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

$$\lambda_{t-1} = V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1} V_{zy(t-1)}, \tag{3}$$

where V_0 is a positive definite matrix, λ_{t-1} is statistic used in noise ϵ_r structure estimation.

3. Illumination invariance

Illumination conditions of an image acquisition can change due to various reasons. In our approach, we allow changes of brightness and spectrum of illumination sources. We assume that a textured surface is illuminated with several illumination sources and that positions of viewpoint and illumination sources remain unchanged. We start with the assumption of single illumination, which is far enough to produce uniform illumination, and planar Lambertian surfaces with varying albedo and surface texture normal. However, these restrictive assumptions will be further relieved to incorporate more illumination sources, nonuniform illumination, and surfaces with a natural reflectance model. Still, the assumptions of fixed illumination positions might sound limiting. Nevertheless, our experiments with natural surfaces show that the derived features are very robust even if the illumination positions changes dramatically.

3.1. Illumination models

Let us assume that a textured Lambertian surface is illuminated with one uniform illumination. The value acquired by the jth sensor at the location r can be expressed as

$$Y_{rj} = \int_\Omega E(\omega) S(r,\omega) R_j(\omega) d\omega, \tag{4}$$

where ω is wavelength, $E(\omega)$ is the spectral power distribution of a single illumination, $S(r,\omega)$ is a Lambertian reflectance coefficient at the position r, $R_j(\omega)$ is the jth sensor response function, and the integral is taken over the visible spectrum Ω. The Lambertian reflectance term $S(r,\omega)$ depends on surface normal, illumination direction, and surface albedo.

Following the works of Finlayson (1995), Healey and Wang (1995), we approximate the surface reflectance $S(r,\omega)$ by a linear combination of a fixed basis $S(r,\omega) = \sum_{c=1}^{C} d_{r,c} s_c(\omega)$, where functions $s_c(\omega)$ are optimal basis functions that represent the data. The method for finding suitable basis was introduced by Marimont and Wandell (1992), they also concluded that, given the human receptive cones, a 3-dimensional basis set is sufficient to model colour observations. However, finding such basis set is not needed in our method, because the key assumption is only its existence. Provided that $j = 1, \ldots, C$ sensor measurements are available

$$Y_{rj} \approx \sum_{c=1}^{C} d_{r,c} \int_\Omega E(\omega) s_c(\omega) R_j(\omega) d\omega,$$

$$Y_r = B' d_r,$$

where $d_r = [d_{r,1}, \ldots, d_{r,C}]^T$ and B' is a $C \times C$ matrix. The two images $\tilde{Y}$, Y acquired with different illumination spectra can be transformed to each other by the linear transformation:

$$\tilde{Y}_r = B Y_r, \quad \forall r, \tag{5}$$

which is same for all the pixels. The formula (5) is valid even for several illumination sources with variable spectra provided that the spectra of all sources are the same and the positions of the illumination sources remain fixed.

Let us generalise the surface reflectance to the natural model of Bidirectional Texture Function (BTF) (Dana et al., 1999), where the reflectance is function of surface position, wavelength, incoming and outgoing light directions. Let $L(r,\omega,v_i,v_o)$ is the surface reflectance, v_i is illumination direction and v_o viewing direction then Eq. (4) becomes

$$Y_{rj} = \int_\Omega E(\omega) L(r,\omega,v_i,v_o) R_j(\omega) d\omega. \tag{6}$$

On the condition that Q is an arbitrary number of reflectance components in the reflectance model (e.g. Lambertian component, different isotropic or anisotropic spectacular components) and each component is separable in ω, the reflectance can be decomposed and approximated:

$$\begin{aligned} L(r,\omega,v_i,v_o) &= \sum_{q=1}^{Q} L^{(q)}(r,\omega,v_i,v_o) = \sum_{q=1}^{Q} \Lambda^{(q)}(r,v_i,v_o) S^{(q)}(r,\omega) \\ &\approx \sum_{q=1}^{Q} \Lambda^{(q)}(r,v_i,v_o) \sum_{c=1}^{C} d_{r,c}^{(q)} s_c(\omega) \\ &= \sum_{c=1}^{C} s_c(\omega) \sum_{q=1}^{Q} d_{r,c}^{(q)} \Lambda^{(q)}(r,v_i,v_o), \end{aligned}$$

where $\Lambda^{(q)}(r,v_i,v_o)$ is the qth reflectance component at position r dependent on the angles, while $S^{(q)}(r,\omega)$ is the reflectance dependent on ω, which is again approximated with optimal basis functions $s_c(\omega)$. Therefore, the images with a different illumination are expressed as

$$Y_r = B' \sum_{q=1}^{Q} d_r^{(q)} \Lambda^{(q)}(r,v_i,v_o) = B' d_r', \quad \forall r,$$

$$\tilde{Y}_r = \tilde{B}' \sum_{q=1}^{Q} d_r^{(q)} \Lambda^{(q)}(r,v_i,v_o) = \tilde{B}' d_r', \quad \forall r,$$

which is in accordance with the linear model (5). For a fixed position r, the sum of functions $\Lambda^{(q)}(r,v_i,v_o)$ becomes the well-known Bidirectional Reflectance Distribution Function (BRDF) (Nicodemus et al., 1977).

The linear model (5) also includes linear transformations to other colour models as CIE XYZ, Opponent colours and Gaussian colour model (Geusebroek et al., 2003) as well.

3.2. Colour invariants

We assume that all pixels of textural images $\tilde{Y}$, Y can be transformed to each other by the linear model (5) and the transformation matrix B is not singular. We proved relations of the model parameter estimates (2) and (3): $\tilde{A}_s = B^{-1} A_s B$, $\tilde{\lambda} = B \lambda B^T$, where $\tilde{\cdot}$ denotes the corresponding statistics for the different illumination. Consequently, the following illumination invariant features were derived (Vacha and Haindl, 2007):

1. trace: $\mathrm{tr} A_s,\ \forall s \in I_r^u$,
2. eigenvalues: $v_{s,j}$ of $A_s,\ \forall s \in I_r^u$, $j = 1,\ldots,C$,
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.

These illumination invariants are easily evaluated during the CAR parameter estimation process. In the case of 2D models, the invariants 3–5 are computed for each spectral plane separately.

In this paper, we experiment with the less general illumination model, where B is restricted to a diagonal matrix. This illumination model still allows a change of illumination colour, but it forbids the mutual switch of image spectral planes. For the diagonal B, the multiplication $B^{-1} A_s B$ do not change the diagonal elements of A_s. Therefore we can alternatively define invariants $v_{s,j}$:

$$2.'\text{diagonals} : v_s = \mathrm{diag}(A_s), \quad \forall s \in I_r^u.$$

3.3. Local intensity changes

The previous invariants were derived with the uniform illumination assumption. We show that most of the invariants are also invariant to locally constant intensity changes, which can be caused by cast shadows or objects with more textured planar surfaces.

Let us start with an auxiliary construction without intensity changes. We assume that a textured image is composed of n copies of the same small texture tile S, which is homogeneous and homogeneously illuminated. The tiles are placed side by side to cover the whole image lattice I. Using the formula (2), the relation of tile parameter estimate $\hat{\gamma}^{(S)}$ and the estimate for the whole image $\hat{\gamma}^{(I)}$ is following

$$\begin{aligned}\hat{\gamma}^{(S)} &= \left(\sum_{r\in S} Z_r Z_r^T\right)^{-1} \left(\sum_{r\in S} Z_r Y_r^T\right),\\ \hat{\gamma}^{(I)} &= \left(\sum_{r\in I} Z_r Z_r^T\right)^{-1} \left(\sum_{r\in I} Z_r Y_r^T\right) \approx \left(n \sum_{r\in S} Z_r Z_r^T\right)^{-1} \left(n \sum_{r\in S} Z_r Y_r^T\right) = \hat{\gamma}^{(S)},\end{aligned} \tag{7}$$

where the approximation (7) discards statistics at the seams of tiles, if the tiles are seamless the equations are precise.

Provided that a modified image is composed of the same texture tiles S, where the ℓth texture tile is modified by the multiplication of all its pixels and spectral planes with some constant b_ℓ. This simulates locally constant intensity changes in the image. The parameter estimate become

$$\begin{aligned}\hat{\tilde{\gamma}}^{(I)} &= \left(\sum_{r\in I} \tilde{Z}_r \tilde{Z}_r^T\right)^{-1} \left(\sum_{r\in I} \tilde{Z}_r \tilde{Y}_r^T\right)\\ &\approx \left(\left(\sum_{\ell=1}^{n} b_\ell^2\right) \sum_{r\in S} Z_r Z_r^T\right)^{-1} \left(\left(\sum_{\ell=1}^{n} b_\ell^2\right) \sum_{r\in S} Z_r Y_r^T\right) = \hat{\gamma}^{(S)},\end{aligned} \tag{8}$$

where $\tilde{\hat{\gamma}}$, $\tilde{Z}_r$, $\tilde{Y}_r$ are related to the illumination modified image. The approximation again discards the seam statistics, which additionally includes local illumination changes. The previous assumption of the tile composed image can be further weaken. The image tiles could be even different on condition that the correlation statistics $\sum_{r\in S} Z_r Z_r^T$ and $\sum_{r\in S} Z_r Y_r^T$ remain the same, which is implicated by the homogeneous property of textures; natural examples are stochastic textures. Eq. (8) instantly implies that illumination invariants $\mathrm{tr} A_s$, $v_{s,j}$ are approximately invariant to local intensity changes. Analogically, it can be proved for the invariant α_2.

4. Rotation invariance

We propose two different methods for the rotation invariance of MRF features. The first method computes rotation invariant features before the estimation of MRF parameters, while the second method builds rotation invariants after the MRF parameter estimation by means of moment invariants.

Theoretically, the both methods are invariant to image rotation, which is imprecise approximation of rotation of rough materials (3D textures) with variable illumination direction (Chantler and August, 1995). However, our experiments show that the proposed methods succeeded in recognition of real materials including the rough ones.

4.1. Rotation autoregressive random model

The rotation autoregressive random (RAR) model is inspired by the model of Kashyap et al. (1986), who estimated a regression model of pixel values and averages on concentric circles around these pixels. Although, this model is suitable for modelling of isotropic textures, it has difficulties with anisotropic texture properties. Our model extends the regression data with maximum and minimum from circular samples, which enables model to capture some anisotropic texture properties.

The basic modelling equation is similar to (1):

$$Y_r = \gamma Z_r + \epsilon_r, \quad Z_r = \left[Y_{r,\max}^T, Y_{r,\min}^T, Y_{r,\mathrm{mean}}^T : \forall s \in I_r^{\odot}\right]^T$$

with the difference in data vector Z_r. The vector Z_r is now composed of averages $Y_{r,\mathrm{mean}}$ of points, which are sampled on concentric circles from the circular contextual neighbourhood $I_r^{\odot}$. The bilinear interpolation is used for the interpolation of sampled points. Additionally, we extend Z_r by maximum $Y_{r,\max}$ and minimum $Y_{r,\min}$ of the sampled points for each circle, which distinguish among isotropic and anisotropic neighbourhood. The parameter estimate $\hat{\gamma}$ cannot be computed using the analytical Bayesian estimate (2) anymore, therefore we use the corresponding least square approximation. This LS estimate leads formally to the same equations as (2) with $V_0 = O$ (zero matrix).

4.2. Rotation moment invariants

The rotation moment invariants are used to describe anisotropic texture properties, which are only briefly captured by the previous model. The CAR model parameters are estimated (Section 2.1) and the rotation moment invariants are computed from the illumination invariants $\mathrm{tr} A_s$, $v_{s,j}$ (Section 3.2), according to their position in the neighbourhood I_r^u. Since the unilateral neighbourhood I_r^u covers only the upper half plane, the values are duplicated in the central symmetry to cover all the plane, which is advantageous for the rotation invariance of moments. The moments are computed separately for each spectral plane of $v_{s,j}$. We also add moment invariants that describe inter spectral relations of $v_{s,j}$.

It is advantageous to compute the rotation invariants from complex moments, because they change more simply in rotation than other types of moments. The complex moment of the order $p + q$ of the function $f(r_1, r_2)$ is defined

$$c_{pq}^{(f)} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (r_1 + i r_2)^p (r_1 - i r_2)^q f(r_1, r_2) dr_1\, dr_2, \tag{9}$$

where i is an imaginary unit. We omit the superscript $^{(f)}$ if there is no danger of confusion. It follows from the definition that only the indices $p \geqslant q$ are meaningful because $c_{pq} = c_{qp}^*$ (the asterisk denotes complex conjugate).

The complex moment c'_{pq} after a rotation of $f(r_1,r_2)$ equals

$$c'_{pq} = e^{-i(p-q)\alpha} \cdot c_{pq}, \tag{10}$$

where α is the rotation angle. Therefore a product of complex moments

$$\prod_{\ell=1}^{n} c_{p_\ell q_\ell}^{k_\ell} \tag{11}$$

is invariant to rotation, if the sum of the first indices equals the sum of the second indices, i.e.

$$\sum_{\ell=1}^{n} k_\ell p_\ell = \sum_{\ell=1}^{n} k_\ell q_\ell \quad \text{or} \quad \sum_{\ell=1}^{n} k_\ell (p_\ell - q_\ell) = 0.$$

The total number of moment invariants equals $m - t$, where m is the number of real values of moments (complex moment with $p \neq q$ has two real values – real and imaginary parts) and t is the number of transform parameters. In our case, the rotation has one parameter (the angle), i.e. $t = 1$ and the number of real values of the sth order moments is $s + 1$. Moment c_{00} is an exception (there should be no zeroth-order invariant according to this rule, but the transform parameter is dependent in this case and it is not counted). The set of invariants should be chosen to be independent, see Flusser et al. (2009), Flusser and Suk (2006) for more details and additional references.

In our case, we must consider the behaviour of the complex moments of symmetric functions. A function has so-called N-fold rotation symmetry (N-FRS), if it repeats itself when it rotates around its centroid by $2\pi\ell/N$ for all $\ell = 1,\dots,N$. The central symmetry is a special case of N-FRS, where $N = 2$. If $f(r_1,r_2)$ has N-FRS and $(p-q)/N$ is not an integer, then $c_{pq} = 0$. If f' is the rotated version of f by the angle $2\pi/N$, then, because of its symmetry, it must be the same as the original. Therefore, it must hold $c'_{pq} = c_{pq}$ for any p and q. At the same time, it follows from Eq. (10) that

$$c'_{pq} = e^{-2\pi i(p-q)/N} \cdot c_{pq}.$$

Since $e^{2\pi i\zeta} = 1$ for an integer ζ and $(p-q)/N$ is assumed not to be an integer, this equation can be fulfilled only if $c_{pq} = 0$.

Since our neighbourhood is centrally symmetric, we cannot use any odd-order moment. That is why we use these even-order rotation moment invariants

- zeroth order: c_{00},
- second order: c_{11}, $c_{20}c_{02}$,
- fourth and mixed order: c_{22}, $c_{40}c_{04}$, $c_{31}c_{13}$, $\mathcal{R}e(c_{40}c_{02}^2)$, $\mathcal{R}e(c_{31}c_{02})$.

We can utilise the fact that all colour channels are rotated together, by the same angle and construct joint colour rotation invariants

- second order: $c_{20}^{(\ell)}c_{02}^{(j)}$,
  where $\ell = 1$, $j = 2,\dots,C$ are the individual colour channels.

We have the matrix of discrete values instead of a continuous function here, therefore we must use a discrete approximation of the complex moments (9):

$$\hat{c}_{pq}^{(f)} = \sum_{r_1} \sum_{r_2} (r_1 + ir_2)^p (r_1 - ir_2)^q f(r_1,r_2). \tag{12}$$

Bilinear interpolation of $f(r_1,r_2)$ is used to enhance its resolution and precision of computed moments.

All single spectral plane moments $\hat{c}_{pq}$ were computed for invariants tr A_s and $v_{s,j}$, $j = 1,\dots,C$ according to their position in the neighbourhood I_r^u. Each matrix $A_{s=(s_1,s_2)}$ is associated with the position (s_1,s_2) in neighbourhood I_r^u, therefore the input function f is defined from the matrix traces and made symmetric as

$$f(r_1,r_2) = \begin{cases} \mathrm{tr}A_{(r_1,r_2)}, & (r_1,r_2) \in I_r^u, \\ \mathrm{tr}A_{(-r_1,-r_2)}, & (-r_1,-r_2) \in I_r^u. \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$

Similar construction applies for each spectral plane j of $v_{s,j}$. The interspectral moments are computed only for invariants $v_{s,j}$, with the same construction of f. Altogether, it makes $34 = 8 + 3 * 8 + 2$ moment invariants for $C = 3$, which we denote as set m_1. Since the high order moments tend to be numerically unstable, especially for roughly defined f, we also work with the reduced set of invariants denoted as m_2, which only includes invariants c_{00}, c_{11}, $c_{20}c_{02}$, c_{22} and $c_{20}^{(1)}c_{02}^{(j)}$.

The illumination invariants α_1, α_2, α_3 (Section 3.2) are not associated with the position in the neighbourhood, therefore the rotation invariant transformation is not necessary and they can be added directly into the feature vector.

5. Texture analysis algorithm

The texture analysis algorithm starts with the factorisation of texture into K levels of the Gaussian downsampled pyramid (it captures larger spatial relations), followed by modelling with two different MRF models. At first, each pyramid level is modelled by the RAR model and the illumination invariants are computed from its parameters. Secondly, each pyramid level is modelled by the CAR model. After the estimation of CAR model parameters, the illumination invariants and subsequently the rotation moment invariants are computed from the parameters. Finally, the features from all the models are concatenated into one feature vector. The algorithm scheme is depicted in Fig. 3.

More precisely, we used $K = 4$ levels of Gaussian pyramid and the CAR models with the sixth order hierarchical neighbourhood (cardinality $\eta = 14$), which corresponds to maximum radius 3 used in the RAR models. Actually, the optimal size of neighbourhood and pyramid depends on the size of input images, because the models require enough data for a reliable parameter estimation. The optimal neighbourhood can be determined automatically (Haindl and Šimberová, 1992), nevertheless we used the same parameters K, η in all our experiments in order to make the results comparable. The moment based features are composed of either a full or re-

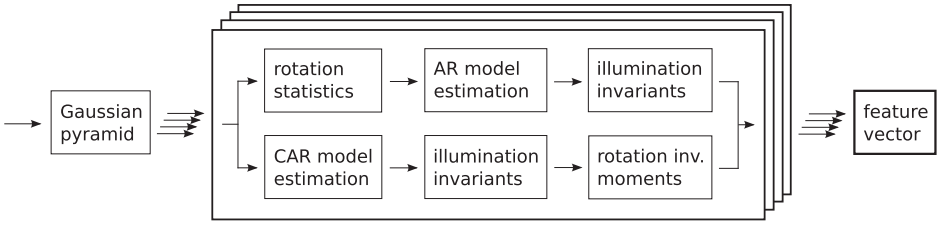

Fig. 3. Texture analysis algorithm which combines illumination invariants with two approaches to rotation invariance. It is either a autoregressive model of rotation invariant statistics (RAR) in the upper line, or a causal autoregressive model followed by the computation of rotation moment invariants (m(CAR-KL)) in the lower line.

duced set of invariants, in the results they are denoted as "m_1(*model*)" or "m_2(*model*)", respectively.

It is possible to decorrelate image spectral planes by means of the Karhunen–Loève transformation before the estimation of the texture model. If the decorrelation is employed, the illumination invariant features $v_{s,j}$ have to be implemented using diagonals instead of eigenvalues ($v_{s,j}$ definition denoted as 2.′ in Section 3.2), otherwise the eigenvalues would cancel the decorrelation. The decorrelation is necessary for the 2D model, since this model is not able to incorporate interspectral relations. In the results, the decorrelation is denoted by "-KL" suffix in model names.

The distance between two feature vectors is computed using the fuzzy contrast (Santini et al., 1999) in its symmetrical form:

$$FC_\beta(T,S) = P - \left\{ \sum_{p=1}^{P} \min\left\{ \tau(f_p^{(T)}), \tau(f_p^{(S)}) \right\} - \beta \sum_{p=1}^{P} \left| \tau(f_p^{(T)}) - \tau(f_p^{(S)}) \right| \right\},$$

$$\tau(f_p) = \left(1 + \exp\left(-\frac{f_p - \mu(f_p)}{\sigma(f_p)} \right) \right)^{-1},$$

where P is the feature vector size and $\mu(f_p)$ and $\sigma(f_p)$ are average and standard deviation of the feature f_p, which were estimated on special parameter tuning sets. The sigmoid function τ models the truth value of a fuzzy predicate. This feature normalisation is necessary since the moments have different scales.

6. Experiments

In the first experiment, we focused on the robustness of textural features under varying illumination and viewpoint direction, which resembles real world scenes with natural materials. In the second experiment, we tested features under varying illumination spectrum and texture rotation, which simulates different day light or artificial illuminations. In the third experiment, our results were compared with other recently published features.

The proposed features were compared with the following illumination and rotation invariant features:

- MR8-$*$. These representatives of filter based methods characterise a texture as the histogram of textons (Varma and Zisserman, 2005), which are clusters of MR8 filter responses. The extension by Burghouts and Geusebroek (2009a) introduced four different degrees of illumination invariance – from colour invariance to cast shadows and shading. We compare our results with variants MR8-NC and MR8-LINC, which were reported with the best performance.
- $LBP^{riu2}_{P,R}$. Local Binary Patterns (Ojala et al., 2002b) are the histograms of micro patterns, which are thresholded values sampled at each pixel neighbourhood (P is the number of samples with radius R). $LBP^{riu2}_{P,R}$ consider only uniform patterns (a subset of simpler patterns), regardless their orientations. The features are computed either on grey-scale images or separately on all spectral planes of colour images and concatenated into a common feature vector (denoted with "RGB" suffix).
- LBP-HF. This extension of $LBP^{riu2}_{P,R}$ (Ahonen et al., 2009) analyses also the relations of micropattern orientations. The Fourier transformation is computed from the histogram of single pattern orientations and the amplitudes of Fourier coefficients are rotation invariant features. The authors' implementation is provided in MATLAB.

6.1. Experiment 1

We follow the experimental setup of Burghouts and Geusebroek (2009a) and evaluate the texture recognition accuracy on CUReT (Dana et al., 1999) and ALOT (Burghouts and Geusebroek, 2009a) datasets.

The ALOT library is a BTF database containing an extraordinary collection of 250 natural materials, each acquired with varying viewpoint and illumination positions (Figs. 1, 2). Most of the materials have rough surfaces, so the movement of light source changes the appearance of materials. Moreover, the significant height variation of some materials (e.g. leaves) causes a large and variable cast shadows, which makes recognition even more difficult. The dataset (Burghouts and Geusebroek, 2009a) consists of images of the first 200 materials divided into tune, train, and test parts, each with 2400 samples. Let c stands for camera, l for light, i for reddish illumination, and r for material rotation. The parameter tuning set consists in samples with setup $c\{1,4\}l\{1,4,8\}r\{60°,180°\}$; the training set is defined as $c\{1,4\}l\{1,4,8\}r\{0°,120°\}$ and finally, the test set contains images with $c\{2,3\}l\{3,5\}r\{0°,120°\}$, $c3l2r\{0°,120°\}$,$c2l2r0°$, and $c1ir0°$. Additionally, we cropped all the images to the same size 1536×660 pixels.

The CUReT database also consists of real-world materials acquired with different combinations of viewing and illumination directions. The dataset provided by Varma and Zisserman (2005) consists of 61 materials, each with 92 samples and resolution 200×200 pixels. We defined the parameter tuning set as the subset the training set which contains the first four samples for each material.

In the setup of Burghouts and Geusebroek (2009a), the classification accuracy is tested with randomly selected training samples and an ensemble of classifiers. The number of training samples per material decreases from 8 to 1. The mean and standard deviation of

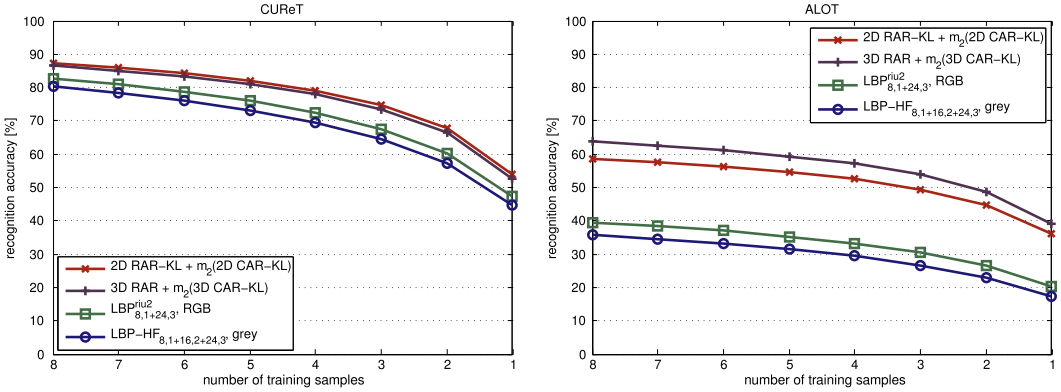

Fig. 4. Accuracy of material recognition [%] for CUReT and ALOT datasets, using different numbers of random training images per material. The values are averages over 10^3 random selections of training images.

classification accuracy is computed over 10^3 repetitions (random selections). We differ only in the classifier, where the simple nearest neighbour (1-NN) is employed.

The results of correct classification and the progression for different number of training samples are shown in Fig. 4. The graphs in Fig. 4 are directly comparable to the results of Burghouts and Geusebroek (2009a), where the best classification accuracy monotonously decreases from 75% to 45% for MR8-LINC on the CUReT and from 40% to 20% for MR8-NC on the ALOT dataset. Standard deviations for the CUReT is below 0.5%, 0.7% and 1.2% for 8, 4 and 1 samples, respectively and for the ALOT dataset, it is below 0.4%, 0.5% and 0.6% for the same number of samples. The more detailed comparison is displayed in Table 1, where also the separate results of two rotation invariant approaches are shown. The best results were achieved with the combination of these two approaches "3D RAR + m_1(3D CAR-KL)", which performed significantly better than LBP and MR8-$*$ alternatives on both datasets. On the ALOT dataset, the proposed features surpassed the best alternative by more than 20%. This remarkable improvement was probably achieved by the combination of colour invariance and robustness to local intensity changes. The performance difference were maintained for all numbers of training images. Moreover, the 3-dimensional model outperformed its 2-dimensional counterpart on the ALOT dataset, since large textures provided enough training data for a precise estimation of interspectral relations.

Table 1
Accuracy of material recognition [%] for CUReT and ALOT datasets, using 4 random training images per material. The values are averages over 10^3 random selections of training images. The bold values highlight the best results in groups and the last column consists of feature vector sizes.

Method	CUReT	ALOT	Size
2D RAR-KL	63.6	45.3	180
m_1(2D CAR-KL)	75.6	38.8	172
m_2(2D CAR-KL)	76.4	37.1	108
2D RAR-KL + m_1 (2D CAR-KL)	**79.5**	**53.4**	352
2D RAR-KL + m_2 (2D CAR-KL)	79.0	52.6	288
3D RAR	62.5	46.8	156
m_1(3D CAR)	57.0	26.0	148
m_1(3D CAR-KL)	71.7	41.1	304
m_2(3D CAR-KL)	72.6	39.2	84
3D RAR + m_1(3D CAR-KL)	**78.1**	**58.3**	304
3D RAR + m_2(3D CAR-KL)	**78.2**	57.1	240
$LBP_{8,1+8,3}$, RGB	71.2	32.0	1536
$LBP^{riu2}_{8,1+24,3}$, RGB	**73.0**	**33.2**	108
$LBP^{riu2}_{8,1+24,3}$, grey	67.3	24.3	36
$LBP\text{-}HF_{8,1+24,3}$, grey	69.5	29.9	340
$LBP\text{-}HF_{8,1+16,2+24,3}$, grey	69.9	29.4	448
Burghouts and Geusebroek (2009a)			
MR8-NC	54	**36**	600
MR8-LINC	**67**	30	600

The recognition accuracy per material is displayed in Fig. 5, where the materials are sorted according to their recognition accuracy. This graph implies that the ALOT dataset includes some very easily recognisable materials as well as extremely difficult ones. It is worth noting that one half of the ALOT test set is acquired with camera 3, which is closer to the material surface and which viewpoint angle is more extreme than cameras used in the training set. (Example images from camera 3 are in two bottom lines in Figs. 1, 2). As result, the classification accuracy for these side viewed images is approximately half of the accuracy for the images from top camera positions, or even worse for LBP features (Fig. 5). The reason is that none of the compared features are invariant to perspective projection.

Moreover, large texture sizes in the ALOT enabled us to experiment with an additional level of the Gaussian pyramid (a level with lower resolution) than in the experiments with the other datasets. This additional level captures larger spatial relations in textures, which is confirmed with a significant performance increase in the ALOT column in Table 2. The CUReT column in Table 2 displays that the additional pyramid levels may decrease the performance when the images do not provide enough data.

The average analysis time for large ALOT images was 20 s for "2D RAR-KL", 11 s for "m_1(2D CAR-KL)", and 10 s for "$LBP^{riu2}_{8,1+24,3}$ RGB", all computed on AMD Opteron 2.1 GHz. The analysis of small CUReT images spent 0.8 s, 0.5 s, and 0.4 s of CPU time per image, respectively.

For comparison reasons, we included the results of the proposed features with a widely used SVM classifier instead of FC_3 dissimilarity and 1-NN. The comparison was performed using LibSVM implementation (Chang and Lin, 2001) and the results in Table 3 are slightly worse than those presented in Table 1. The presented results were achieved with the linear C-SVC classifier, which was preceded with whitening of features – scaling to have zero mean and unit variance.

6.2. Experiment 2

In the second experiment, we demonstrate the performance of the proposed features on the Outex database (Ojala et al., 2002a). This database consists of natural material images acquired under three illuminations with almost same positions but different spectra. The illumination sources were 2856 K incandescent CIE A light source – "inca", 2300 K horizon sunlight – "horizon", and 4000 K fluorescent TL84 – "tl84".

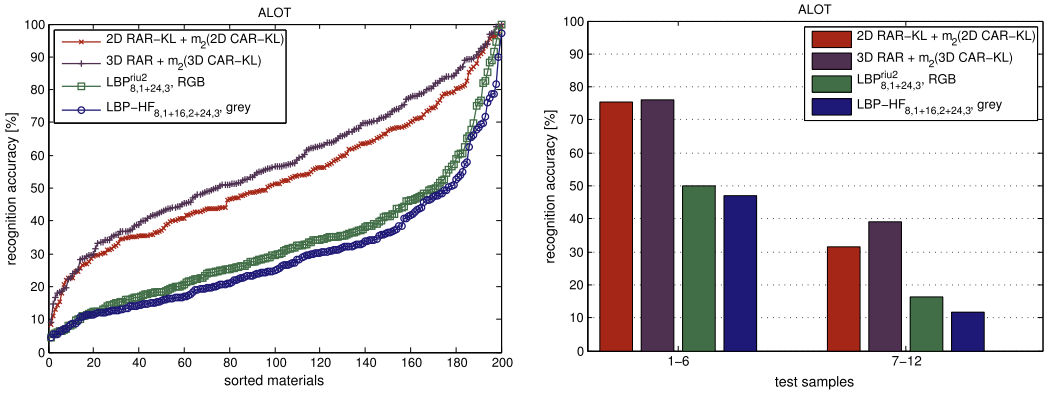

Fig. 5. Accuracy of material recognition [%] for the ALOT dataset, using 4 training samples per materials and 10^3 random selections of training images. On the left, it is displayed the recognition accuracy per material, where the materials are sorted according to their recognition accuracy. On the right, the accuracy for test samples (1–6) acquired from top camera positions is compared to side viewed test samples (7–12).

Table 2
Accuracy of material recognition [%] for CUReT and ALOT datasets, experiment with one additional Gaussian pyramid level. The setup is same as in Table 1.

Method	CUReT	ALOT	Size
↑ 2D RAR-KL + m_1(2D CAR-KL)	**78.4**	61.6	440
↑ 3D RAR + m_1(3D CAR-KL)	74.9	**65.3**	380

Table 3
Accuracy of material recognition [%] for CUReT and ALOT datasets, using SVM classifier preceded with whitening of features. Otherwise, the setup is same as in Table 1.

Method + SVM	CUReT	ALOT	Size
2D RAR-KL + m_1(2D CAR-KL)	76.1	51.3	440
3D RAR + m_1(3D CAR-KL)	75.1	56.6	380

The experiment was performed on the classification test Outex_TC_0012 (Ojala et al., 2002a). In this set, 24 materials were selected from the Outex database, subsequently 20 subsamples with size 128 × 128 were extracted for each of 9 rotations of each material. The training set consists in 20 subsamples per material, with 0° rotation, illuminated with the "inca" light source. On the other hand, the each of two test sets consists of 20 subsamples per material, with all 9 rotations. The first test set contains subsamples illuminated with "horizon", while the second one contains images with "tl84" light. Consequently, the train set consists of 480 images, while the test sets are composed of 4320 images each. Although this test is focused on colour invariance, all images are available only in the grey-scale. This setup disables an exploitation of interspectral dependences, which are the key properties in illumination spectrum invariance.

We used the nearest neighbour classifier instead of three nearest neighbours as Ojala et al. (2002b), since the performance differences were negligible. The mean and standard deviation of features, used by FC_3 dissimilarity, were estimated on the training set. The averages of correct classification on both test sets are displayed in Table 4. The recently published results of LBP-$HF_{8,1+16,2+24,3}$ are slightly better than the proposed features, however the feature vector of LBP-$HF_{8,1+16,2+24,3}$ is almost five times longer. The proposed features suffered from grey-scale input images, which are, in our opinion, not suitable for testing of colour invariance.

6.3. Experiment 3

The third experiment compares the performance of the proposed features on the KTH-TIPS2 database (Caputo et al., 2005),

Table 4
Outex_TC_0012: averages of material recognition accuracy [%] on both test sets. The last column consists of feature vector sizes.

Method	Total	Size
2D RAR-KL	87.5	48
m_1(2D CAR-KL)	64.6	44
m_2(2D CAR-KL)	68.1	28
2D RAR-KL + m_1(2D CAR-KL)	87.6	92
2D RAR-KL + m_2(2D CAR-KL)	**89.6**	60
$LBP^{riu2}_{8,1+24,3}$, grey	87.6	36
Ojala et al. (2002b)		
$LBP^{riu2}_{8,1+24,3}$, grey	87.2	36
Ahonen et al. (2009)		
$LBP^{riu2}_{8,1+24,3}$, grey	88.3	36
LBP-$HF_{8,1+24,3}$, grey	91.7	340
LBP-$HF_{8,1+16,2+24,3}$, grey	**92.5**	448

Table 5
KTH-TIPS2: accuracy of material recognition [%] averaged over 10^4 random training set selections. The last column consists of feature vector sizes.

Method	Mean	Size
2D RAR-KL	58.6	180
m_1(2D CAR-KL)	59.6	172
m_2(2D CAR-KL)	59.1	108
2D RAR-KL + m_1(2D CAR-KL)	**63.2**	352
2D RAR-KL + m_2(2D CAR-KL)	63.0	288
3D RAR	58.8	156
m_1(3D CAR)	49.6	148
m_1(3D CAR-KL)	58.7	148
m_2(3D CAR-KL)	57.8	84
3D RAR + m_1(3D CAR-KL)	**65.0**	304
3D RAR + m_2(3D CAR-KL)	**65.0**	240
$LBP_{8,1\,+\,8,3}$, RGB	**56.0**	**1536**
$LBP^{riu2}_{8,1+24,3}$, RGB	54.1	108
$LBP^{riu2}_{8,1+24,3}$, grey	49.6	36
Ahonen et al. (2009)		
$LBP^{riu2}_{8,1+24,3}$, grey	50.7	36
LBP-$HF_{8,1+24,3}$, grey	54.2	340
LBP-$HF_{8,1+16,2+24,3}$, grey	**54.6**	448

which includes samples with different scales and rotations. Because the training set includes various scales and rotations, the invariance is not a key issue.

The KTH-TIPS database contains 4 samples of 11 materials categories, each sample consists of images with 4 different illuminations, 3 in-plane rotations and 9 scales. The illumination conditions consist in 3 different directions plus 1 image with different spectrum. There are 4572 images in total and their resolution is varying around 200 × 200 pixels. We follow the experimental setup of Ahonen et al. (2009), where the nearest neighbour classifier was trained with one random sample (4 × 3 × 9 images) per material category. The remaining images (3 × 108 per category) were used for testing. This was repeated for 10^4 random partitioning to training and test sets. Since the setup do not define a parameter tuning set, we defined it as the subset of training set which contains the first sample of each material category.

Table 5 compares the average classification accuracy, where standard deviation is 2% or below. Although, a large variety of training image conditions allowed non-invariant features to perform comparably, still the proposed features took advantage of their invariance and outperformed alternatives by more than 10%.

7. Conclusion

We have proposed rotation and colour invariant features, which are simultaneously robust to illumination direction and local intensity changes. The features advantageously combine illumination invariants with two constructions of rotation invariants: either modelling of rotation invariant statistics or moment invariants computed from direction sensitive model parameters. The experiments were designed to closely resemble real life conditions of natural material recognition. The tests were performed on 4 different texture databases, which included almost 300 natural materials in total and the materials were acquired with different viewpoint, illumination colour and direction. The proposed features outperformed leading alternative features as MR8-∗, LBP^{riu2} and LBP-HF. As overall best method we suggest "3D RAR + m_2(3D CAR-KL)" or its 2D counterpart if less training data is available.[1]

In future research we are going to integrate a more elaborate classifier for the combination of features.

[1] Demonstration is available at http://cbir.utia.cas.cz/rotinv/.

Acknowledgements

The authors thank Jan-Mark Geusebroek from University of Amsterdam and Gertjan J. Burghouts from TNO Observation Systems for the textures and details for the first experiment.

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

References

Ahonen, T., Matas, J., He, C., Pietikäinen, M., 2009. Rotation invariant image description with local binary pattern histogram Fourier features. In: Salberg, A.-B., Hardeberg, J.Y., Jenssen, R. (Eds.), Proc. 16th Scandinavian Conf. on Image Analysis, SCIA 2009, Lecture Notes in Computer Science, vol. 5575. Springer-Verlag, pp. 61–70.

Burghouts, G.J., Geusebroek, J.-M., 2009a. Material-specific adaptation of color invariant features. Pattern Recognition Lett. 30, 306–313.

Burghouts, G.J., Geusebroek, J.-M., 2009b. Performance evaluation of local colour invariants. Computer Vision and Image Understanding 113 (1), 48–62.

Caputo, B., Hayman, E., Mallikarjuna, P., 17–21 October 2005. Class-specific material categorisation. In: Proc. 10th IEEE Internat. Conf. on Computer Vision, ICCV 2005, IEEE, pp. 1597–1604.

Chang, C.-C., Lin, C.-J., 2001. LIBSVM: A Library for Support Vector Machines. Available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chantler, M., 1995. Why illuminant direction is fundamental to texture analysis. IEE Proc. – Vision Image Signal Process. 142 (4), 199–206.

Dana, K., Van-Ginneken, B., Nayar, S., Koenderink, J., 1999. Reflectance and texture of real world surfaces. ACM Trans. Graph. 18 (1), 1–34.

Deng, H., Clausi, D.A., 2004. Gaussian MRF rotation-invariant features for image classification. IEEE Trans. Pattern Anal. Machine Intell. 26 (7), 951–955.

Diplaros, A., Gevers, T., Patras, I., 2006. Combining color and shape information for illumination-viewpoint invariant object recognition. IEEE Trans. Image Process. 15 (1), 1–11.

Filip, J., Haindl, M., 2009. Bidirectional texture function modeling: A state of the art survey. IEEE Trans. Pattern Anal. Machine Intell. 31 (11), 1921–1940.

Finlayson, G.D., 1995. Coefficient Color Constancy. Ph.D. Thesis, Simon Fraser University.

Flusser, J., Suk, T., 2006. Rotation moment invariants for recognition of symmetric objects. IEEE Trans. Image Process. 15 (12), 3784–3790.

Flusser, J., Suk, T., Zitová, B., 2009. Moments and Moment Invariants in Pattern Recognition. Wiley, Chichester.

Geusebroek, J.-M., van den Boomgaard, R., Smeulders, A.W.M., Gevers, T., 2003. Color constancy from physical principles. Pattern Recognition Lett. 24 (11), 1653–1662.

Haindl, M., Mikes, S., Vacha, P., 7–10 November 2009. Illumination invariant unsupervised segmenter. In: Proc. IEEE Internat. Conf. on Image Processing, ICIP 2009, IEEE, pp. 4025–4028.

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

Haley, G.M., Manjunath, B.S., 1999. Rotation-invariant texture classification using a complete space-frequency model. IEEE Trans. Image Process. 8 (2), 255–269.

Healey, G., Wang, L., 20–23 June 1995. The illumination-invariant recognition of color texture. In: Proc. 5th IEEE Internat. Conf. on Computer Vision, ICCV 1995. IEEE, pp. 128–133.

Jafari-Khouzani, K., Soltanian-Zadeh, H., 2005. Radon transform orientation estimation for rotation invariant texture analysis. IEEE Trans. Pattern Anal. Machine Intell. 27 (6), 1004–1008.

Kashyap, R.L., Khotanzad, A., 1986. A model-based method for rotation invariant texture classification. IEEE Trans. Pattern Anal. Machine Intell. PAMI-8 (4), 472–481.

Liao, S., Law, M.W.K., Chung, A.C.S., 2009. Dominant local binary patterns for texture classification. IEEE Trans. Image Process. 18 (5), 1107–1118.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Internat. J. Comput. Vision 60 (2), 91–110.

Mao, J., Jain, A.K., 1992. Texture classification and segmentation using multiresolution simultaneous autoregressive models. Pattern Recognition 25 (2), 173–188.

Marimont, D.H., Wandell, B.A., 1992. Linear models of surface and illuminant spectra. J. Opt. Soc. Amer. 9, 1905–1913.

Nicodemus, F., Richmond, J.C., Hsia, J., Ginsburg, I., Limperis, T., October 1977. Geometrical considerations and nomenclature for reflectance. NBS Monograph 160, National Bureau of Standards, US Department of Commerce, Washington, DC, pp. 1–52.

Ojala, T., Mäenpää, T., Pietikäinen, M., Viertola, J., Kyllönen, J., Huovinen, S., 11–15 August 2002a. Outex – New framework for empirical evaluation of texture analysis algorithms. In: Proceeding of the 16th International Conference on Pattern Recognition, ICPR 2002, vol. 1. IEEE, pp. 701–706.

Ojala, T., Pietikäinen, M., Mäenpää, T., 2002b. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Machine Intell. 24 (7), 971–987.

Ravichandran, G., Trivedi, M.M., 1995. Circular-Mellin features for texture segmentation. IEEE Trans. Image Process. 4 (12), 1629–1640.

Santini, S., Jain, R., 1999. Similarity measures. IEEE Trans. Pattern Analysis Machine Intell. 21 (9), 871–883.

Shotton, J.D.J., Winn, J., Rother, C., Criminisi, A., 2009. Textonboost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context. Internat. J. Comput. Vision 81 (1), 2–23.

Smeulders, A.W., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Machine Intell. 22 (12), 1349–1380.

Vacha, P., Haindl, M., 9–11 July 2007. Image retrieval measures based on illumination invariant textural MRF features. In: Sebe, N., Worring, M. (Eds.), Proc. ACM Internat. Conf. on Image and Video Retrieval, CIVR 2007. ACM, pp. 448–454.

Vacha, P., Haindl, M., 8–11 December 2008. Illumination invariants based on Markov random fields. In: Proc. 19th Internat. Conf. on Pattern Recognition, ICPR 2008. IEEE, pp. 1–4.

Varma, M., Zisserman, A., 2005. A statistical approach to texture classification from single images. Int. J. Comput. Vision 62 (1-2), 61–81.