Illumination Invariants Based on Markov Random Fields

Pavel Vácha and Michal Haindl

Institute of Information Theory and Automation of the ASCR

Czech Republic

1. Introduction

Ongoing expansion of digital images requires new methods for sorting, browsing, and searching through huge image databases. Content-based image retrieval (CBIR) systems, which are database search engines for images, typically takes a user selected image or series of images and tries to retrieve similar images from a large image database. Although image retrieval has been an active research area for many years (Smeulders et al., 2000) this difficult problem is still far from being solved. One of the reasons is that common image features do not provide required discriminability and invariance. Optimal robust features should be geometry and illumination invariant and still remain highly discriminative, which are often contradictory requirements.

Simpler CBIR methods are based only on colour features and achieve illumination invariance by normalising colour bands or using the colour ratio histogram (Gevers & Smeulders, 2001). However, colour based methods rarely perform sufficiently well in natural visual scenes because they cannot detect similar objects in different locations, backgrounds or illuminations.

Textures are important clues to specify objects present in a visual scene. Unfortunately, the visual appearance of natural prevailing textures is highly illumination and view direction dependent. As a consequence, most recent natural texture based classification or segmentation methods require multiple training images captured under a full variety of possible illumination and viewing conditions for each class (Suen & Healey, 2000; Varma & Zisserman, 2005). Such learning is obviously clumsy and very often even impossible if required measurements are not available. Drbohlav & Chantler (2005) allow a single training image per class, but they require surfaces of uniform albedo, smooth and shallow relief, the illumination sufficiently far from the texture macro-normal and most seriously the knowledge of illumination direction for all involved (trained as well as tested) textures. It was demonstrated by Chen et al. (2000); Jacobs et al. (1998) that for grey image of an object with Lambertian reflectance there are no discriminative functions that are invariant to change of illumination direction. However, multispectral images can also rely on relations of spectral planes and therefore overcome this theoretical property.

Colour constancy algorithms, represented by Finlayson (1995), attempt to recover the image illuminated by some standard illumination, which is an unnecessarily complex task and it induces additional assumptions on a recognised scene. The normalisation of an image before the recognition proposed in Finlyason & Xu (2002) is able to cancel changes of illumination colour, lighting changes caused by the object geometry and even a power (gamma) function,

which is usually applied to image data during the coding process. However, since the method normalises lighting changes caused by the geometry it completely wipes out the structure of rough textures and therefore it destroys the possibility to recognise such textures. Simultaneously, the invariants to geometry introduced lighting changes tend to be unstable because of nonlinear transformations usually involved. An interesting approach of quasi-invariants (Weijer et al., 2005) relieves the condition of full invariance and therefore it is less sensitive to noise. Parameters of Weibull-distribution of image edges (Geusebroek & Smeulders, 2005) are proposed as insensitive to illumination brightness. Healey & Wang (1995), Yang & Al-Rawi (2005) employ properties of correlation functions between different spectral channels to achieve invariance to illumination spectrum changes. Hoang & Geausebroek (2005), Geusebroek et al. (2003) introduced a method based on the logarithm of Gabor filter responses together with a new Gaussian colour model. However, the Gaussian colour model of RGB texture is implemented as a simple matrix multiplication and invariance to any linear transformation of texture values is inherent part of the illumination invariant features proposed in the rest of this article.

Local Binary Patterns (Ojala, Pietikäinen & Mäenpää, 2002) (LBP) are popular illumination invariant features, which we use for comparison. The texton representation (Varma & Zisserman, 2005) based on MR8 filter responses have been extended to incorporate colour information and to be illumination invariant (Burghouts & Geusebroek, 2009). Another approach (Targhi et al., 2008) generates unseen training images using the photometric stereo approach. Although, it improves classification accuracy, this algorithm has strong requirements of three mutually registered images with different illumination direction for each material.

We present textural features, which are invariant to illumination brightness and spectrum changes and which do not require any knowledge of illumination spectrum. The features are robust to illumination direction changes and do not require knowledge of illumination direction or mutual texture registration. They can be applied for textured object retrieval if only a single illumination training image is available for each class.

These properties are verified on Outex database (Ojala, Mäenpää, Pietikäinen, Viertola, Kyllönen & Huovinen, 2002), where texture images are illuminated under three different spectra, and University of Bonn BTF texture measurements (Meseth et al., 2003), where illumination sources are spanned over 75% of possible illumination direction above material samples. Preliminary experiments were published in Vacha & Haindl (2008), Vacha & Haindl (2007).

The chapter is organised as follows: The illumination model assumptions are reviewed in Section 2. Texture model description and derivation of illumination invariant features follow in Section 3. Experimental results are presented in Section 5 and Section 6 concludes this chapter.

2. Illumination Model

Illumination conditions can change due to various reasons. In our approach we allow changes of brightness and spectrum of illumination sources. We assume that positions of viewpoint and illumination sources remain unchanged and that the illumination sources are far enough to produce uniform illumination. Furthermore, we assume planar textured Lambertian surfaces with varying albedo and surface texture normal.

The previous assumptions are quite strong and they are required to theoretically derive the illumination invariance of features. However, our experiments with natural surfaces show that the derived features are very robust even in the setup, which is in contradiction with previous assumptions (test with eighty different illumination positions).

Let us denote a multiindex $r = (r_1, r_2)$ where r_1 is the row and r_2 the column index, respectively. Value acquired by the j−th sensor at the location r can be expressed as

$$Y_{r,j} = \int_\omega E(\lambda)\, S(r, \lambda)\, R_j(\lambda)\, d\lambda \ ,$$

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

Following the work of Finlayson (1995), we approximate the surface reflectance $S(r, \lambda)$ by a linear combination of a fixed basis $S(r, \lambda) = \sum_{c=1}^{C} d_c\, s_c(\lambda)$, where functions s_c are optimal basis functions that represent the data. The method for finding suitable basis was introduced by Marimont & Wandell (1992), they also conclude that, given the human receptive cones, a 3-dimensional basis set is sufficient to model colour observations. However, finding such basis set is not necessary our method, because the key assumption is its existence.

Provided that $j = 1, \ldots, C$ sensor measurements are available and the illumination and view point positions are the same, the images acquired with different illumination spectra can be transformed to each other by the linear transformation:

$$\tilde{Y}_r = B\, Y_r \qquad \forall r \ , \tag{1}$$

where $\tilde{Y}, Y$ are texture images with different illuminations, and B is a $C \times C$ transformation matrix. The formula (1) 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.

More importantly, it can be proved that formula (1) is valid not only for Lambertian surfaces, but also for surface model with specular reflectance component (e.g. dichromatic reflection model (Shafer, 1985), which comprise also the well-known Phong reflection model).

If we assume further diagonal transformation (1), then the invariance to illumination colour change can be achieved by the spectral planes normalisation:

$$Y'_{r,j} = \frac{Y_{r,j}}{\sum_s Y_{s,j}} \qquad \forall j = 1, \ldots, C \ . \tag{2}$$

Since neither of our methods requires this type of normalisation, it is applied to Gabor features, which are further used for comparison purposes, only.

No method that guarantee an illumination invariance to surface geometry effects have been used. The primary reason is that the invariance to this geometry effects is not desirable because a surface structure is the inherent part of a surface texture appearance. Moreover, many images are extremely dark and the normalisation of RGB triplets $\frac{z}{r+g+b}, z \in \{r, g, b\}$ produces not only a huge amount of noise, but it is also undefined for black pixels ($r + g + b = 0$). On the other hand, invariants based on pixel hues are ambiguous on the black-white axis and again they are not suitable for rough textures with uniform colours.

3. Texture Representation

Let us assume each texture to be composed of C spectral planes measured by the corresponding sensors. Texture analysis starts with the factorisation of a texture into K lev-

els of the Gaussian pyramid. All C spectral planes are factorised using the same pyramid thus the corresponding multispectral pixels for every pyramid level have C components $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$. Each pyramid level is either modelled by 3-dimensional Markov random field (MRF) model or mutually decorrelated by the Karhunen-Loeve transformation (Principal Component Analysis) and subsequently modelled using a set of C 2-dimensional MRF models. The MRF model parameters are estimated and finally illumination invariants are computed from these parameters.

3.1 CAR Model

The CAR representation assumes that the multispectral texture pixel Y_r at the $k-$th Gaussian pyramid level can be locally modelled by an adaptive simultaneous Causal Autoregressive Random (CAR) field model. We denote the $C\eta \times 1$ data vector

$$Z_r = [Y_{r-s}^T : \forall s \in I_r]^T \tag{3}$$

where $r = (r_1, r_2)$, s, t are multiindices, Z_r consists of neighbour pixel values for given r. The multiindex changes according to the chosen direction of movement on the image plane e.g. $t - 1 = (t_1, t_2 - 1)$, $t - 2 = (t_1, t_2 - 2), \ldots$. Some selected contextual causal or unilateral neighbour index shift set is denoted I_r and $\eta = cardinality(I_r)$. The matrix form of an adaptive CAR model is:

$$Y_r = \gamma Z_r + \epsilon_r \ , \tag{4}$$

where $\gamma = [A_1, \ldots, A_\eta]$ is the $C \times C\eta$ unknown parameter matrix with matrices A_s. In the case of C 2D CAR models stacked into the model equation (4) the parameter matrices A_s are diagonal otherwise they are full matrices for general 3D CAR models. The white noise vector ϵ_r has zero mean and constant but unknown covariance matrix Σ. Moreover, we assume the probability density of ϵ_r to have the normal distribution independent of previous data and being the same for every position r. Additionally for 2D CAR model, we assume uncorrelated noise vector components, i.e.,

$$E\{\epsilon_{r,i}\epsilon_{r,j}\} = 0 \qquad \forall r, i, j, \ i \neq j \ .$$

The task consists in finding the parameter conditional density $p(\gamma \mid Y^{(t-1)})$ given the known process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ and taking its conditional mean as the textural feature representation. Assuming normality of the white noise component ϵ_t, conditional independence between pixels and the normal-Wishart parameter prior, we have shown (Haindl & Šimberová, 1992) that the conditional mean value is:

$$E[\gamma \mid Y^{(t-1)}] = \hat{\gamma}_{t-1} \ . \tag{5}$$

The following notation is used in (5):

$$
\begin{aligned}
\hat{\gamma}_{t-1}^T &= V_{zz(t-1)}^{-1} V_{zy(t-1)} \ , \\
V_{t-1} &= \tilde{V}_{t-1} + V_0 \ , \\
\tilde{V}_{t-1} &= \begin{pmatrix} \sum_{u=1}^{t-1} Y_u Y_u^T & \sum_{u=1}^{t-1} Y_u Z_u^T \\ \sum_{u=1}^{t-1} Z_u Y_u^T & \sum_{u=1}^{t-1} Z_u Z_u^T \end{pmatrix} \\
&= \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}_{zy(t-1)}^T \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{zz(t-1)} \end{pmatrix} \ , \\
\lambda_{t-1} &= V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1} \tag{6}
\end{aligned}
$$

and V_0 is a positive definite matrix. It is easy to check (see Haindl & Šimberová (1992)) also the validity of the following recursive parameter estimator:

$$\hat{\gamma}_t^T = \hat{\gamma}_{t-1}^T + \frac{V_{zz(t-1)}^{-1} Z_t (Y_t - \hat{\gamma}_{t-1} Z_t)^T}{(1 + Z_t^T V_{zz(t-1)}^{-1} Z_t)} \ , \tag{7}$$

and λ_t can be evaluated recursively too. For numerical realisation of the model statistics (5)-(7) see discussion in Haindl & Šimberová (1992).

Textural features for each Gaussian pyramid level k is represented by the parametric matrix $\hat{\gamma}^{(k)}$, $k = 1, \ldots, K$. These parametric estimates are combined into the resulting parametric matrix:

$$\Theta = [\hat{\gamma}^{(k)} \ \forall k] \ . \tag{8}$$

This matrix contains estimations of the multiresolution CAR model (a set of either 2D or 3D CAR models) parameters. Illumination invariants are subsequently derived from these parameters.

3.2 GMRF Factor Model

The alternative representation assumes that spectral planes of the $k-$th pyramid level are locally modelled using a 2D Gaussian Markov Random Field model (GMRF). This model is obtained if the local conditional density of the MRF model is Gaussian:

$$p(Y_{r,j}|Y_{s,j} \ \forall s \in I_r) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp \left\{ -\frac{(Y_{r,j} - \gamma_j Z_{r,j})^2}{2\sigma_j^2} \right\} , \tag{9}$$

where $Y_{r,j}$ are mean centred values and j is the spectral plane index $j = 1 \ldots C$. The data vector is redefined as $Z_{r,j} = [Y_{r+s,j} \quad \forall s \in I_r]^T$ and the parameter vector is $\gamma_j = [a_{s,j} \quad \forall s \in I_r]$. The contextual neighbourhood I_r is non-causal and symmetrical. The GMRF model for centred values $Y_{r,j}$ can be expressed also in the matrix form (4), but the driving noise ϵ_r and its correlation structure is now more complex:

$$E\{\epsilon_{r,i}\epsilon_{r-s,j}\} = \begin{cases} \sigma_j^2 & \text{if } (s) = (0,0) \text{ and } i = j, \\ -\sigma_j^2 a_{s,j} & \text{if } (s) \in I_r \text{ and } i = j, \\ 0 & \text{otherwise,} \end{cases} \tag{10}$$

where $\sigma_j, a_{s,j} \ \forall s \in I_r$ are unknown parameters. The parameter estimation of the GMRF model is complicated because either Bayesian or Maximum likelihood estimate requires an iterative minimisation of a nonlinear function. Therefore we use the pseudo-likelihood estimator which is computationally simple although not efficient. The pseudo-likelihood estimate for $a_{s,j}$ parameters has the form

$$\hat{\gamma}_j^T = [a_{s,j} \quad \forall s \in I_r]^T$$

$$= \left[\sum_{\forall r \in I} Z_{r,j} Z_{r,j}^T \right]^{-1} \sum_{\forall s \in I} Z_{r,j} Y_{r,j} \ , \tag{11}$$

$$\hat{\sigma}_j^2 = \frac{1}{|I|} \sum_{\forall r \in I} (Y_{r,j} - \hat{\gamma}_j Z_{r,j})^2 \ , \tag{12}$$

where $j = 1 \ldots C$, I is an image lattice. Single spectral plane parameters are set up using the direct sum

$$\hat{\gamma}^{(k)} = diag(\hat{\gamma}_1, \ldots, \hat{\gamma}_C) = \oplus_{j=1}^{C} \hat{\gamma}_j \tag{13}$$

and the resulting parametric matrix is again (8).

3.3 MRF Illumination Invariant Features

Illumination invariant feature vectors can be derived from the estimated MRF statistics such as (8), which is composed of the model parameter matrices A_m. On the condition that two images Y, $\tilde{Y}$ under different illumination are related by $\tilde{Y}_r = B\,Y_r$ (see (1)), the model data vectors are also related by the linear transformation $\tilde{Z}_r = \Delta\,Z_r$, where Δ is the $C\eta \times C\eta$ block diagonal matrix with blocks B on the diagonal. By substituting $\tilde{Y}_r, \tilde{Z}_r$ into the parameter estimate of the CAR model (4), (6), (7) we can derive that

$$\tilde{A}_m = B\,A_m B^{-1}, \quad \tilde{\lambda}_r = B\,\lambda_r B^T \ . \tag{14}$$

The matrices $\tilde{A}_m, \tilde{Z}_r, \tilde{\lambda}_r$ are related to the model of the same texture, but with different illumination. The similar substitution into the GMRF parameter estimate (4), (11), (12) produces equations

$$\tilde{A}_m = B\,A_m B^{-1}, \quad \hat{\tilde{\Sigma}} = B\,\hat{\Sigma} B^T \ , \tag{15}$$

where $\hat{\Sigma} = diag(\hat{\sigma}_1, \ldots, \hat{\sigma}_C)$. It is easy to prove that the following features are illumination invariant for both models:

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

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

for each CAR model (for 2D CAR models the invariants $\alpha_1, \alpha_2, \alpha_3$ are computed for each spectral plane separately):

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,

and for each GMRF model with centred $Y_{r,j}$:

6. α_4: $\sqrt{\sum_r \hat{\sigma}_j^{-2} \left(Y_{r,j} - \hat{\gamma}_j Z_{r,j} \right)^2}$,

7. α_5: $\sqrt{\sum_r \hat{\sigma}_j^{-2} \left(Y_{r,j} \right)^2}$.

The feature vector is formed from these illumination invariants. For CAR models we use traces, eigenvalues, α_1, α_2, and α_3 features because they can be easily evaluated during the parameters estimation process. For GMRF models we use trace, eigenvalues, α_4, and α_5 features, respectively.

3.4 Feature Comparison Distances

The distance between illumination invariant feature vectors of two textures T, S is computed using the Minkowski norms L_1, $L_{0.2}$, or alternatively with fuzzy contrast FC_3 proposed by Santini & Jain (1999). The Minkowski norm and its σ normalised variant $L_{p\sigma}$, which is used for comparison of alternative texture features, is defined as follows

$$L_p(T,S) = \left(\sum_{i=0}^{M} \left| f_i^{(T)} - f_i^{(S)} \right|^p \right)^{\frac{1}{p}} , \tag{16}$$

$$L_{p\sigma}(T,S) = \left(\sum_{i=0}^{M} \left| \frac{f_i^{(T)} - f_i^{(S)}}{\sigma(f_i)} \right|^p \right)^{\frac{1}{p}} , \tag{17}$$

$$\tag{18}$$

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 image database, respectively.

Fuzzy contrast $FC_{\alpha\beta}$ models features as predicates in fuzzy logic using sigmoid truth function τ. Subsequently, the feature vector dissimilarity is defined as

$$FC_{\alpha\beta}(T,S) = M - \left\{ \sum_{i=1}^{m} \min \left\{ \tau(f_i^{(T)}), \tau(f_i^{(S)}) \right\} \right. \tag{19}$$

$$- \alpha \sum_{i=1}^{p} \max \left\{ \tau(f_i^{(T)}) - \tau(f_i^{(S)}), 0 \right\}$$

$$\left. - \beta \sum_{i=1}^{p} \max \left\{ \tau(f_i^{(S)}) - \tau(f_i^{(T)}), 0 \right\} \right\} ,$$

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

It is worth to note that $FC_{\alpha\beta}$ is not a metric, because it does not hold $FC_{\alpha\beta}(T,S) = 0$ and it is not necessary symmetrical. However, we use only its symmetrical form FC_3, where $\alpha = \beta = 3$.

4. Alternative Features

Our proposed illumination invariant features are compared with the most frequently used features in image retrieval applications such as the Gabor features, steerable pyramid features and Local Binary Patterns (LBP).

4.1 Gabor Features

The Gabor filters (Bovik, 1991; Randen & Husøy, 1999) can be considered as orientation and scale tunable edge and line detectors. The statistics of Gabor filter responses in a given region are used to characterise the underlying texture information. A two dimensional Gabor

function $g(r) : \mathbb{R}^2 \to \mathbb{C}$ and its Fourier transform can be specified as

$$g(r) = \frac{1}{2\pi\sigma_{r_1}\sigma_{r_2}} \exp\left[-\frac{1}{2}\left(\frac{r_1^2}{\sigma_{r_1}^2} + \frac{r_2^2}{\sigma_{r_2}^2} \right) + 2\pi i V r_1 \right] \ ,$$

$$G(r) = \exp\left\{ -\frac{1}{2}\left[\frac{(r_1 - V)^2}{\sigma_{r_1}^2} + \frac{r_2^2}{\sigma_{r_2}^2} \right] \right\} \ ,$$

where $\sigma_{r_1}, \sigma_{r_2}, V$ are filter parameters. The convolution of the Gabor filter and a texture image extracts edges of given frequency and orientation range. The whole filter set was obtained by four dilatations and six rotations of the function $g(r)$, the filter set is designed so that Fourier transformations of filters cover most of image spectrum, see Manjunath & Ma (1996) for details. The Gabor features (Manjunath & Ma, 1996) are defined as the mean μ_l and the standard deviation σ_l of the magnitude of filter responses computed separately for each spectral plane and concatenated into the feature vector. The suggested distance between feature vectors of textures T, S is $L_{1\sigma}(T, S)$ normalised Minkowski norm (18).

The Opponent Gabor features (Jain & Healey, 1998) are the extension to colour textures, which analyses also relations between spectral channels. The monochrome part of these features is:

$$\eta_{i,m,n} = \sqrt{\sum_r W_{i,m,n}^2(r)} \ ,$$

where $W_{i,m,n}$ is the response to Gabor filter of orientation m and scale n, i is $i-$th spectral band of the colour texture T, while the opponent part of features is:

$$\psi_{i,j,m,m',n} = \sqrt{\sum_r \left(\frac{W_{i,m,n}(r)}{\eta_{i,m,n}} - \frac{W_{j,m',n}(r)}{\eta_{j,m',n}} \right)^2} \ ,$$

for all i, j with $i \neq j$ and $|m - m'| \leq 1$. The previous formula could be also expressed as correlation between spectral planes responses. The distance between textures T, S using the Opponent Gabor features is measured with $L_{2\sigma}(T, S)$ normalised Minkowski norm (18), as suggested by Jain & Healey (1998).

In order to achieve illumination invariance, it is possible to normalise spectral channels using (2) normalisation prior to computation of features. We have tested Gabor features and Opponent Gabor features, the both options with and without the normalisation.

4.2 Steerable Pyramid Features

The steerable pyramid (Portilla & Simoncelli, 2000) is an over complete wavelet decomposition similar to the Gabor decomposition. The pyramid is built up of responses to steerable filters, where level of pyramid extracts certain frequency range. All pyramid levels (except the highest and the lowest one) are further decomposed to different orientations. The transformation is implemented using the set of oriented complex analytic filters B_l that are polar separable in the Fourier domain (see details in Simoncelli & Portilla (1998), Portilla & Simoncelli (2000)):

$$
\begin{aligned}
B_l(R,\theta) &= H(R)G_l(\theta), && l \in [0, L-1], \\[6pt]
H(R) &= \begin{cases} \cos\left(\frac{\pi}{2}\log_2\left(\frac{2R}{\pi}\right)\right), & \frac{\pi}{4} < R < \frac{\pi}{2} \\ 1, & R \geq \frac{\pi}{2} \\ 0, & R \leq \frac{\pi}{4} \end{cases} \\[6pt]
G_l(\theta) &= \begin{cases} \alpha_l\left[\cos\left(\theta - \frac{\pi l}{L}\right)\right]^{L-1}, & \left|\theta - \frac{\pi l}{L}\right| < \frac{\pi}{2}, \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}
$$

where $\alpha_l = 2^{l-1}\dfrac{(L-1)!}{\sqrt{L[2(L-1)!]}}$, R, θ are polar frequency coordinates, $L = 4$ is the number of orientation bands, and $K = 4$ is the number of pyramid levels. The used steerable pyramid features was skewness, kurtosis, mean, variance, minimum and maximum values of image function, and scale-based auto-correlations and subband cross-correlations of filter responses, respectively, which were proposed for texture synthesis in Portilla & Simoncelli (2000). The feature vectors are compared using the $L_{1\sigma}$ norm (18). Again, we have tested steerable pyramid features with and without the channel normalisation.

4.3 Local Binary Patterns

Local Binary Patterns (LBP) (Ojala, Pietikäinen & Mäenpää, 2002) are histograms of texture micro patterns. For each pixel, a circular neighbourhood around the pixel is sampled, then sampled points values are thresholded by the central pixel value and the pattern number is formed as follows:

$$
LBP_{P,R} = \sum_{q=0}^{P-1} \text{sign}\left(Y_q - Y_c\right) 2^q, \quad \text{sign}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0, \end{cases} \tag{20}
$$

where P is the number of samples and R is the radius of the circle, *sign* is the signum function, Y_q is a grey value of the sampled pixel, and Y_c is a grey value of the central pixel. Subsequently, the histogram of patterns is computed. Because of thresholding, the features are invariant to any monotonic grey scale change. The multiresolution analysis is done by growing the circular neighbourhood size. However, complex patterns do not have enough occurrences in a texture, therefore uniform LBP features denoted as LBP^{u2} distinguish only among patterns that include only 2 or less transitions between 0 and 1 at neighbouring bits (20), all other patterns are considered to be the same. Moreover, the features can be also made rotation invariant (Ojala, Pietikäinen & Mäenpää, 2002). All LBP histograms were normalised to have a unit L_1 norm. The similarity between texture feature vectors T, S is defined as

$$
L_G(T,S) = \sum_{i=1}^{M} f_i^{(T)} \log \frac{f_i^{(T)}}{f_i^{(S)}}. \tag{21}
$$

We have tested features: $LBP_{8,1+8,3}$ and $LBP^{u2}_{16,2}$ which demonstrated the best performance in the test with illumination changes (Maenpaa et al., 2002; Pietikainen et al., 2002) (test set Outex 14). We have also comprised rotation invariant feature $LBP^{riu2}_{16,2}$. The features were computed either on grey-scale images or on each spectral plane separately and concatenated to form the feature vector. A normalisation to change of illumination brightness or spectrum change is not necessary.

method	experiment		
	1	2	3
Gabor features	144	144	144
Gabor features, grey img.	48	48	48
Opponent Gabor features	**252**	**252**	**252**
Steerable pyramid	2904	2904	2904
Gabor features, norm.	144	144	144
Gabor features, grey img., norm.	48	48	48
Opponent Gabor features, norm.	**252**	**252**	**252**
$LBP_{8,1+8,3}$	1536	1536	1536
$LBP_{16,2}^{u2}$	729	729	729
$LBP_{16,2}^{riu2}$	54	54	54
$LBP_{8,1+8,3}$, grey img.	**512**	**512**	**512**
$LBP_{16,2}^{u2}$, grey img.	243	243	243
$LBP_{16,2}^{riu2}$, grey img.	18	18	18
2D CAR-KL	**260**	**132**	**325**
2D CAR	**260**	**132**	**325**
GMRF-KL	248	120	310
3D CAR	236	108	295

Table 1. The size of feature vectors.

5. Experiments

We demonstrate the performance of the proposed illumination invariant MRF features on two image databases, each with different variations in illumination conditions. The first, BTF database is acquired with a fixed illumination spectrum and with 91 different illumination directions, which drastically violates our restrictive assumption of fixed illumination position. On the other hand, the Outex texture database is acquired with three illuminations with different spectra and only with slight differences in illumination positions, which complies with our assumptions.

We tested three proposed MRF models: 2D CAR (2-dimensional), 3D CAR (3-dimensional) and GMRF. The models were computed over K levels of the Gaussian pyramids, which were built either directly on C spectral planes or on spectral planes decorrelated by the Karhunen-Loeve transformation (indicated with '-KL' suffix).

The proposed features were compared with the following alternatives: Gabor features, Opponent Gabor features, Steerable pyramid features, all with and without spectral channels normalisations (2), and also the LBP features (see details in Section 4). The grey value based features as Gabor features and LBP were computed not only on grey images, but also separately on each spectral plane of colour images and concatenated subsequently. Tab. 1 compares the sizes of feature vectors used in our experiments.

5.1 Experiment 1

We have designed the first experiment to test the feature robustness against illumination direction changes, which is in contradiction with our theoretical assumptions. The experiment is performed on three different sets of BTF texture images. These BTF data are from the University of Bonn database (Meseth et al., 2003) and consists of fifteen BTF colour measurements

Fig. 1. BTF material measurements, row-wise from left to right: ceiling, corduroy, fabric1, fabric2; walk way, foil, floor tile, pink tile; impalla, proposte, pulli, wallpaper; wool, and two lacquered wood textures.
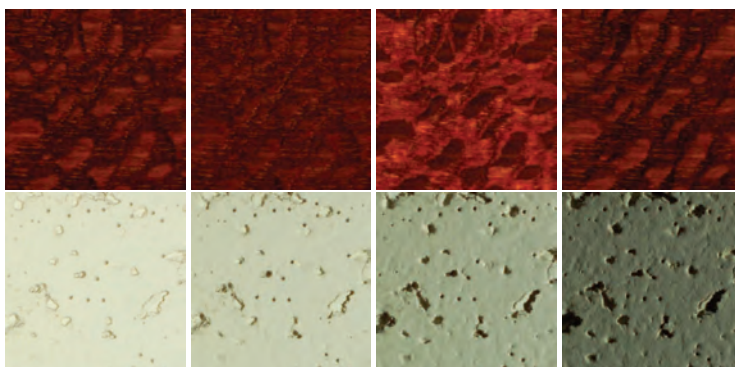

Fig. 2. Illumination variance of BTF materials, the top row is wood with different azimuth of illumination, bottom is ceiling with changing declination of illumination.
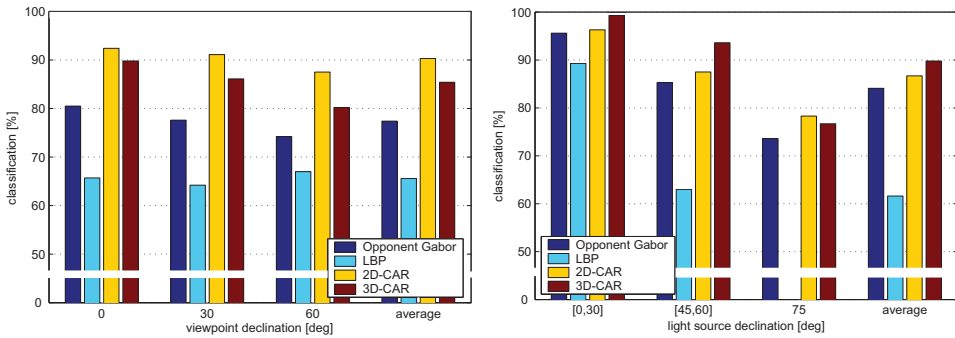
Fig. 3. Correct classification [%], in the left picture for random training samples, in the right picture for training image fixed to top illumination.

(Fig. 1), ten of these measurements are now publicly available[1]. Each BTF material is measured in 81 illumination angles as a RGB image ($C = 3$). Fig. 2 shows examples of material appearance under varying illumination direction. We have three image test sets, which differs in viewpoint direction. The declination angles of viewing direction from surface normal are $0°$, $30°$, and $60°$, in plane rotation is not included. Each test set consists in 1215 images, all cropped to the same size 256×256.

In this experiment, single training image per each material was randomly chosen and the remaining images were classified using the nearest neighbour approach. The MRF models were computed with the sixth order hierarchical neighbourhood and four levels of the Gaussian pyramid, the size of feature vectors is listed in Tab. 1. The results of correct classification are in Tab. 2, all averaged over 10^5 random choices of training images. Bold typeset rows from Tab. 2 are shown in Fig. 3. 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%. Standard deviation is bellow 4% for Gabor features and LBP features, and below 3% for CAR and GMRF models. Although the LBP features 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. For the MRF features, the worst classification result were for ceiling and fabric2 materials. Ceiling material was misclassified as floor tile (for illumination near surface), and fabric2 was sometimes misclassified as fabric1, because they have very similar structures.

We have also explored, how the performance of features depends on light source declination from surface normal. The test set with viewpoint fixed at $0°$ declination was used and single training sample per each material was selected. All selected training samples were illuminated under $0°$ declination angle, other 1200 images were classified. Tab. 3 depicts how correct classification decreases as the illumination position of test sample move away from training sample position. Again, bold typeset rows from Tab. 3 are shown in Fig. 3. The best result were achieved by 3D CAR model and FC_3 dissimilarity with average accuracy 89.8%, similar results 88.7% were achieved by 2D CAR model with FC_3.

---

[1] http://btf.cs.uni-bonn.de/

| | viewpoint declination angle | | | |
method	0°	30°	60°	avg.
Gabor features	71.7	64.6	60.1	65.5
Gabor features, grey img.	69.8	62.9	55.6	62.8
Opponent Gabor features	**82.5**	77.7	71.7	77.3
Steerable pyramid	72.3	65.5	60.4	63.1
Gabor features, norm.	60.1	58.1	57.9	58.7
Gabor features, grey img., norm.	50.8	50.1	51.3	50.7
Opponent Gabor features, norm.	**80.5**	**77.6**	**74.2**	**77.4**
$LBP_{8,1+8,3}$	**65.7**	**64.2**	**67.0**	**65.6**
$LBP_{16,2}^{u2}$	62.5	61.6	64.6	62.9
$LBP_{16,2}^{riu2}$	**68.4**	60.7	57.4	62.2
$LBP_{8,1+8,3}$, grey img.	61.2	61.1	65.4	62.6
$LBP_{16,2}^{u2}$, grey img.	55.7	56.3	60.7	57.6
$LBP_{16,2}^{riu2}$, grey img.	58.6	52.1	52.5	54.4
2D CAR-KL, L_1	**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
2D CAR, FC_3	**88.7**	**87.3**	**82.9**	**86.3**
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

Table 2. Correct classification [%], using single training image per texture. The results are averages over 10^5 random choices of training images, the last column consists in averages of previous columns.

5.2 Experiment 2

In the second experiment, we demonstrate the performance of the proposed illumination invariant MRF features on the Outex database (Ojala, Mäenpää, Pietikäinen, Viertola, Kyllönen & Huovinen, 2002), which consists of natural material images acquired, under three different illuminations. The illumination sources were 2856K incandescent CIE A light source, 2300K horizon sunlight, and 4000K fluorescent TL84, the illumination positions are very close. All the images were acquired with a fixed camera position.

The experiment was performed on the Outex classification test set number 14 (Ojala, Mäenpää, Pietikäinen, Viertola, Kyllönen & Huovinen, 2002). In this test set, 68 materials selected from the Outex database were treated in the following manner. Twenty subsamples with size 128×128 were extracted from each material image. The train set consists of 10 samples per material, all illuminated with the 2586K incandescent CIE A light source. The test set consists of 10 remaining subsamples for each material, all of them illuminated with other two illuminants. Consequently, the train set consists of 680 images, while the test set is composed of 1360 images. The classification was performed using 3 nearest neighbours as in Maenpaa et al. (2002), Pietikainen et al. (2002).

	light source declination			
method	0:30°	45:65°	75°	avg.
Gabor features	96.3	71.4	28.9	64.2
Gabor features, grey img.	95.2	64.7	34.7	62.6
Opponent Gabor features	95.6	83.9	50.0	76.4
Steerable pyramid	90.7	69.5	36.1	64.3
Gabor features, norm.	81.9	49.1	19.4	47.6
Gabor features, grey img., norm.	81.9	38.8	13.9	41.0
Opponent Gabor features, norm.	**95.6**	**85.3**	**73.6**	**84.1**
$LBP_{8,1+8,3}$	**89.3**	**63.0**	**38.6**	**61.6**
$LBP_{16,2}^{u2}$	84.4	51.4	35.6	54.1
$LBP_{16,2}^{riu2}$	84.4	44.6	31.9	49.8
$LBP_{8,1+8,3}$, grey img.	86.3	57.4	38.3	58.2
$LBP_{16,2}^{u2}$, grey img.	79.3	50.7	34.7	52.3
$LBP_{16,2}^{riu2}$. grey img.	74.1	36.8	16.7	39.2
2D CAR-KL, L_1	**96.3**	**87.5**	**78.3**	**86.7**
3D CAR, L_1	97.8	89.6	75.6	87.2
GMRF-KL	95.9	82.6	65.3	80.4
2D CAR, FC_3	**96.7**	**91.6**	**78.1**	**88.7**
GMRF-KL, FC_3	93.3	86.5	72.2	83.7
GMRF-KL, $L_{0.2}$	94.4	82.3	68.3	80.8
2D CAR-KL, FC_3	97.8	90.5	79.4	88.8
2D CAR-KL, $L_{0.2}$	96.7	85.4	78.3	85.8
3D CAR, FC_3	**99.3**	**93.6**	**76.7**	**89.8**
3D CAR, $L_{0.2}$	97.8	91.2	72.8	87.2

Table 3. Correct classification [%] with traning image fixed to top illumination. The performance is grouped for different intervals of illumination declination angles of test images, the last column is average for all test images. Viewpoint declination angle was 0° all images.

The highest reported classification accuracy on the test set (Maenpaa et al., 2002) was 69% for $LBP_{16,2}^{u2}$ features, which outperformed Gabor features with 66% of accuracy (unfortunately our implementation of Gabor features reached only 54.5% in Tab. 4), both features were computed on grey-scale images. Moreover, Pietikainen et al. (2002) reported 68.4% accuracy for $LBP_{8,1+8,3}$ also on grey-scale images, and 53.3% of accuracy achieved by Opponent Gabor features on colour images preceded by comprehensive colour normalisation.

In addition to the previously described experiment, we have also degraded test set images with an additive Gaussian noise. The experiment was performed directly on noisy images, without any noise removal method. The application of such method might increase classification accuracy, but only on condition that it would not introduce any artificial micro structures into the images.

Because of small image size, the neighbourhood of MRF models have to be restricted to the third order hierarchical neighbourhood. As a consequence the feature vector of 2D CAR, is about four times smaller than the vector of $LBP_{8,1+8,3}$ features. The best result on the original test set was achieved with $LBP_{8,1+8,3}$ on grey images with 71.6% followed by the best of MRF features with 67.5% classification accuracy. However, the results change dramatically

method	added noise σ			
	0	2	4	8
Gabor features	37.5	37.0	36.2	35.6
Gabor features, grey img.	44.3	43.3	43.2	41.3
Opponent Gabor features	50.7	49.3	45.3	37.3
Steerable pyramid	37.5	35.9	34.9	32.6
Gabor features, norm.	57.0	59.9	60.3	57.1
Gabor features, grey img., norm.	**54.5**	**61.3**	**63.3**	**62.9**
Opponent Gabor features, norm.	**56.7**	**55.8**	**54.3**	**47.9**
$LBP_{8,1+8,3}$	66.8	56.6	48.8	36.7
$LBP_{16,2}^{u2}$	62.0	52.9	41.2	28.7
$LBP_{16,2}^{riu2}$	44.6	30.8	22.6	15.3
$LBP_{8,1+8,3}$, grey img.	**71.6**	**62.2**	**54.6**	**38.6**
$LBP_{16,2}^{u2}$, grey img.	67.6	60.4	49.8	33.0
$LBP_{16,2}^{riu2}$, grey img.	56.9	45.2	34.2	19.7
2D CAR-KL, L_1	**67.6**	**60.8**	**55.7**	**52.3**
GMRF-KL, L_1	61.5	57.0	51.1	46.1
3D CAR, L_1	63.6	61.3	60.6	54.9
2D CAR, FC$_3$	**67.5**	**62.2**	**61.0**	**56.6**
2D CAR-KL, FC_3	67.5	63.3	55.8	51.0
2D CAR-KL, $L_{0.2}$	66.3	60.5	55.2	51.0
3D CAR, FC_3	65.3	60.4	58.0	51.3
3D CAR, $L_{0.2}$	63.5	59.7	55.4	47.4

Table 4. Results [%] of the Outex classification test set number 14. The classification was performed using three nearest neighbours.

with added noise, the $LBP_{8,1+8,3}$ features drop down to 38.6% showing their vulnerability to noise degradation. The MRF based features are not so noise sensitive, because Gaussian noise is inherent part of the model and the Gaussian pyramid suppresses noise at its higher levels. The MRF features without KL transform performed better on noisy images than features with KL transform, which was deflected by uncorrelated noise. In this experiment, Gabor features performed better than Opponent Gabor Features, especially on noisy images.

5.3 Experiment 3

The last test is an illumination invariant image retrieval from the Outex texture database. In this task, a CBIR system tries to retrieve images similar with a given query image, it differs from the previous experiment in two major points. At first, all 318 materials were used and at second, the task of image retrieval do not allow more training samples per material (previous experiment used 20 training samples per each of 68 materials).

The test set consists of 3 different illuminations for each material, without any rotation and with 100 dpi resolution, which is 954 images in total. All images were cropped to size 512×512 pixels. We have tested image retrieval using every image from the test set. The relevant images were defined as images of the same material with the other two illuminations. Therefore there were 2 relevant images present in the test set for each query image, a total amount of 3 images were retrieved. Furthermore, images were again corrupted with an additive Gaussian noise to test the noise robustness of the features. The retrieval performance was measured using the

method	added noise σ	
	0	8
Gabor features	14.0	13.4
Gabor features, grey img.	42.8	42.4
Opponent Gabor features	38.8	30.5
Steerable pyramid	19.4	18.9
Gabor features, norm.	40.4	27.5
Gabor features, grey img, norm.	**53.4**	**56.1**
Opponent Gabor features, norm.	**46.9**	**37.8**
$LBP_{8,1+8,3}$	51.5	20.0
$LBP_{16,2}^{u2}$	47.3	11.7
$LBP_{16,2}^{riu2}$	24.3	3.1
$LBP_{8,1+8,3}$, grey img.	**83.1**	**50.3**
$LBP_{16,2}^{u2}$, grey img.	80.6	40.8
$LBP_{16,2}^{riu2}$, grey img.	61.5	21.3
2D CAR-KL L_1	**83.4**	**60.8**
GMRF-KL, L_1	73.4	50.4
3D CAR, L_1	79.6	61.5
2D CAR, FC$_3$	**94.0**	**83.0**
GMRF-KL, FC_3	81.3	55.9
GMRF-KL, $L_{0.2}$	80.7	60.2
2D CAR-KL, FC_3	90.2	68.0
2D CAR-KL, $L_{0.2}$	87.8	69.4
3D CAR, FC_3	82.7	63.0
3D CAR, $L_{0.2}$	79.1	58.8

Table 5. Illumination invariant retrieval from the Outex texture database. The performance is measured as recall rate [%] of 3 retrieved images.

recall rate

$$\text{rr} = \frac{\text{retrieved and relevant}}{\text{all retrieved}} \quad ,$$

the results are summarised in Tab. 5.

In this test, the proposed illumination invariant MRF features achieved retrieval recall rate 94%, which clearly present their insensitivity to illumination spectrum variations. The MRF models were computed with the sixth order hierarchical neighbourhood and five levels of the Gaussian pyramid, since the images are large enough. This experiment also confirms that MRF features without Karhunen-Loeve transform are more robust to noise degradation. The LBP features also show their illumination invariance property with 83% recall rate. However, their performance drops with added Gaussian noise increase. This results also demonstrate that the spectral channel normalisation is essential for Gabor and steerable pyramid features, nevertheless, any variant of Gabor or steerable pyramid features did not performed satisfactory in this test.

Some examples of retrieved textures are presented in Fig. 4. The first two retrieved images are both correct, however, more interesting are images retrieved at the further positions. We can observe that the proposed features managed to recognise visual similarity of various canvas materials, and at the bottom, it is obvious that MRF features with pyramids prefer over-

Query image:

canvas034-inca

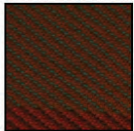**Retrieval result (Outex - proposed 2DCAR):**
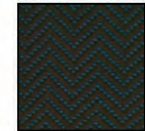
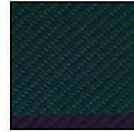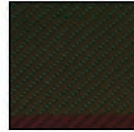| canvas034-horizon | canvas034-tl84 | canvas035-horizon |
| canvas035-inca | canvas037-horizon | canvas035-tl84 |

Retrieval result (Outex - LBP):

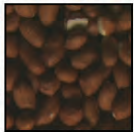| canvas034-tl84 | canvas034-horizon | gravel005-tl84 |
| gravel005-inca | gravel005-horizon | flour013-inca |

Query image:

seeds013-horizon

Retrieval result (Outex - proposed 2DCAR):

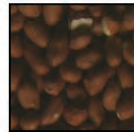| seeds013-tl84 | seeds013-inca | plastic031-horizon |
| plastic029-horizon | plastic030-horizon | plastic031-tl84 |

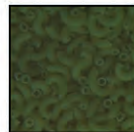**Retrieval result (Outex - LBP):**

| seeds013-tl84 | seeds013-inca | pasta006-horizon |
| pasta004-inca | pasta006-tl84 | pasta004-horizon |

Fig. 4. Illumination invariant image retrieval from Outex database for materials canvas034 (top), seed013 (bottom). Screenshots are from our online demonstration at `http://cbir.utia.cas.cz` or `http://ro.utia.cz/dem.html`.

all structure contrary to micro patterns (such as lines) preferred by the LBP features. These screenshots were taken from our online demonstration,[2] where it is possible to try performance of the proposed features on images from Outex database.

6. Conclusions

We have proposed new illumination invariant features, which are suitable for content based image retrieval systems and other analysis of natural scene images. These features are derived from the underlying Markov random field texture representation and they are invariant to brightness and illumination spectrum variations. We have experimentally verified that introduced MRF features are considerably robust to illumination direction changes and they are simultaneously robust to image degradation by an additive Gaussian noise. Moreover, the MRF features are also fast to compute and the feature vector has low dimension. The disadvantage is that reliable estimation of MRF parameters requires sufficient size of training data.

The proposed methods were compared with Gabor features, Opponent Gabor features, steerable pyramid and LBP features, respectively. Although the LBP features confirmed their illumination spectrum invariance, they had significant difficulties with natural materials illuminated with different directions and images degraded by an additive noise. The best results were achieved with invariants based on the 2-dimensional CAR model with Fuzzy contrast dissimilarity.

7. Acknowledgements

8. References

Bovik, A. (1991). Analysis of multichannel narrow-band filters for image texture segmentation, *IEEE Trans. on Signal Processing* **39**(9): 2025–2043.

Burghouts, G. J. & Geusebroek, J. M. (2009). Material-specific adaptation of color invariant features, *Pattern Recognition Letters* **30**: 306–313. `www.science.uva.nl/research/publications/2009/BurghoutsPRL2009`.

Chen, H. F., Belhumeur, P. N. & Jacobs, D. W. (2000). In search of illumination invariants, *IEEE Computer Vision and Pattern Recognition or CVPR*, pp. I: 254–261.

Drbohlav, O. & Chantler, M. (2005). Illumination-invariant texture classification using single training images, *Texture 2005: Proceedings of the 4th international workshop on texture analysis and synthesis*, pp. 31–36.

Finlayson, G. D. (1995). *Coefficient color constancy*, PhD thesis, Simon Fraser University.

Finlyason, G. & Xu, R. (2002). Illuminant and gamma comprehensive normalisation in logrgb space, *Patterm Recognition Letters* **24**: 1679–1690.

Geusebroek, J.-M., Boomgaard, R. v. d., Smeulders, A. W. & Gevers, T. (2003). Colour constancy from physical principes, *Pattern Recognition Letters* **24**: 1653–1662.

Geusebroek, J.-M. & Smeulders, A. W. (2005). A six-stimulus theory for stochastic texture, *International Journal of Computer Cision* **62**: 7–16.

---

[2] `http://cbir.utia.cas.cz` or `http://ro.utia.cz/dem.html`

Gevers, T. & Smeulders, A. W. M. (2001). Color constant ratio gradients for image segmenta-
    tion and similarity of texture objects, *CVPR*, IEEE Computer Society, pp. 18–25.
Haindl, M. & Šimberová, S. (1992). *Theory & Applications of Image Analysis*, World Scien-
    tific Publishing Co., Singapore, chapter A Multispectral Image Line Reconstruction
    Method, pp. 306–315.
Healey, G. & Wang, L. (1995). The illumination-invariant recognition of texture in
    color texture, *ICCV*, pp. 128–133. `computer.org/proceedings/iccv/7042/`
    `70420128abs.htm`.
Hoang, M. A. & Geausebroek, Jan-Mark Smeulders, A. W. (2005). Color texture measurement
    and segmentation, *Signal Processing* **85**: 295–275.
Jacobs, D., Belhumeur, P. & Basri, R. (1998). Comparing images under variable illumination,
    Proceedings IEEE Conference on Computer Vision and Pattern Re cognition, 2000, Vol. 1,
    IEEE, IEEE, pp. 610 – 617.
Jain, A. & Healey, G. (1998). A multiscale representation including opponent colour features
    for texture recognition, *IEEE Transactions on Image Processing* **7**(1): 125–128.
Maenpaa, T., Pietikainen, M. & Viertola, J. (2002). Separating color and pattern information
    for color texture discrimination, *International Conference on Pattern Recognition*, pp. I:
    668–671.
Manjunath, B. S. & Ma, W. Y. (1996). Texture features for browsing and retrieval of image data,
    IEEE Transactions on Pattern Analysis and Machine Intelligence **18**(8): 837–842.
Marimont, D. H. & Wandell, B. A. (1992). Linear models of surface and illuminant spectra,
    Journal of the Optical Society of America **9**: 1905–1913.
Meseth, J., Müller, G. & Klein, R. (2003). Preserving realism in real-time rendering of bidirec-
    tional texture functions, *OpenSG Symposium 2003*, Eurographics Association, Switzer-
    land, pp. 89–96.
Ojala, T., Mäenpää, T., Pietikäinen, M., Viertola, J., Kyllönen, J. & Huovinen, S. (2002). Ou-
    tex - new framework for empirical evaluation of texture analysis algorithms, *16th
    International Conference on Pattern Recognition*, pp. 701–706.
Ojala, T., Pietikäinen, M. & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invari-
    ant texture classification with local binary patterns, *IEEE Trans. Pattern Anal. Mach.
    Intell* **24**(7): 971–987.
Pietikainen, M., Maenpaa, T. & Viertola, J. (2002). Color texture classification with color his-
    tograms and local binary patterns, *Workshop on Texture Analysis in Machine Vision*,
    pp. 109–112.
Portilla, J. & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of
    complex wavelet coefficients, *International Journal of Computer Vision* **40**(1): 49–71.
Randen, T. & Husøy, J. H. (1999). Filtering for texture classification: A comparative study,
    IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(4): 291–310.
Santini, S. & Jain, R. (1999). Similarity measures, *IEEE Trans. Pattern Anal. Mach. Intell*
    21(9): 871–883. `www.computer.org/tpami/tp1999/i0871abs.htm`.
Shafer, S. A. (1985). Using color to seperate reflection components, *COLOR research and appli-
    cation* **10**(4): 210–218.
Simoncelli, E. P. & Portilla, J. (1998). Texture characterization via joint statistics of wavelet
    coefficient magnitudes, *Proceedings of Fifth International Conference on Image Processing*,
    IEEE Computer Society, pp. I: 62–66.

Smeulders, A. W., Worring, M., Santini, S., Gupta, A. & Jain, R. (2000). Content-based image retrieval at the end of the early years, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(12): 1349–1380.

Suen, P. H. & Healey, G. (2000). The analysis and reconstruction of real-world textures in three dimensions, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(5): 491–503.

Targhi, A. T., Geusebroek, J.-M. & Zisserman, A. (2008). Texture classification with minimal training images, *Proceedings of the 19th International Conference on Pattern Recognition*.

Vacha, P. & Haindl, M. (2007). Image retrieval measures based on illumination invariant textural MRF features, *in* N. Sebe & M. Worring (eds), *CIVR*, ACM, pp. 448–454. `doi.acm.org/10.1145/1282280.1282346`.

Vacha, P. & Haindl, M. (2008). Illumination invariants based on markov random fields, *Proceedings of the 19th International Conference on Pattern Recognition*.

Varma, M. & Zisserman, A. (2005). A statistical approach to texture classification from single images, *International Journal of Computer Vision* **62**(1-2): 61–81. `dx.doi.org/10.1007/s11263-005-4635-4`.

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

Yang, J. & Al-Rawi, M. (2005). Illumination invariant recognition of three-dimensional texture in color images, *J. Comput. Sci. Technol* **20**(3): 378–388.