Image Retrieval Measures Based on Illumination Invariant Textural MRF Features

Pavel Vacha
Institute of Information Theory and Automation
Academy of Sciences CR
Prague, CZ182 08, Czech Republic
vacha@utia.cas.cz

Michal Haindl*
Institute of Information Theory and Automation
Academy of Sciences CR
Prague, CZ182 08, Czech Republic
haindl@utia.cas.cz

ABSTRACT

Content-based image retrieval (CBIR) systems, target database images using feature similarities with respect to the query. We introduce fast and robust image retrieval measures that utilise novel illumination invariant features extracted from three different Markov random field (MRF) based texture representations. These measures allow retrieving images with similar scenes comprising colour textured objects viewed with different illumination brightness or spectrum.

The proposed illumination insensitive measures are compared favourably with the most frequently used features like the Local Binary Patterns, steerable pyramid and Gabor textural features, respectively. The superiority of these new illumination invariant measures and their robustness to added noise are empirically demonstrated in the illumination invariant recognition of textures from the Outex database.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Retrieval models*; I.4.7 [**Image Processing and Computer Vision**]: Feature Measurement—*Texture,Invariants*; I.5.1 [**Pattern Recognition**]: Models—*Statistical*; G.3 [**Probability and Statistics**]: Markov processes

General Terms

Algorithms, Theory, Measurement

*Corresponding author.
Tel:+420-266052350
http://ro.utia.cz

CIVR'07, July 9–11, 2007, Amsterdam, The Netherlands

Keywords

Content-based Image retrieval (CBIR), illumination invariants, Markov random fields (MRF)

1. INTRODUCTION

Content-based image retrieval systems typically query large image databases based on some automatically generated colour and textural features. Optimal robust features should be geometry and illumination invariant. Although image retrieval has been an active research area for many years [24] this difficult problem is still far from being solved and proposed solutions are still very immature. Simpler methods based only on colour features achieve illumination invariance by normalising colour bands or using the colour ratio histogram [4], [9]. However, colour based methods rarely perform sufficiently well in natural visual scenes because they cannot detect similar objects in different location, backgrounds or illumination.

Textures are important clues to specify objects present in a visual scene. Unfortunately, the appearance of natural textures is highly illumination 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 [25], [14]. Such learning is obviously clumsy and very often even impossible if required measurements are not available. Authors [3] allow a single training image per class, but they require surfaces of uniform albedo, smooth and shallow relief and most seriously the knowledge of illumination direction for all involved textures. It was demonstrated [15],[2] that for grey image of an object with Lambertian reflectance there are no discriminative functions that are invariant to change of illumination direction.

Colour constancy algorithms, represented by [5], 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 [6] is able to cancel changes of illuminated 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 [26] releases the condition of full invariance and therefore it is less sensitive to noise. Parameters of Weibull-distribution of image edges [8] are proposed as insensitive to illumination brightness. Authors [12], [27] employ properties of correlation functions between different spectral channels to achieve invariance to illumination spectrum changes. Authors [13], [7] introduced a method based on the logarithm of Gabor filter responses together with 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.

We introduce three methods based on simple parametric measures, which are invariant to illumination brightness and spectrum changes and which do not require any knowledge of illumination spectrum. They can be applied for textured object retrieval if only single illumination training image is available for each class.

The paper is organised as follows: Assumptions about illumination model are reviewed in Section 2. Texture model description and derivation of illumination invariant features follow in Section 3. Experimental results are presented in Section 4 and Section 5 concludes the paper.

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 fixed positions of viewpoint and illumination sources. Moreover, the illumination sources are supposed to be far from the texture surface to produce uniform illumination. Furthermore, we assume planar textured Lambertian surfaces with varying albedo and surface texture normal.

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 [5], we approximate the surface reflectance $S(r,\lambda)$ by a linear combination of a fixed basis $S(r,\lambda) = \sum_{c=1}^{C} d_c\, s_c(\lambda)$. 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 illumination, 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.

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 method requires this type of normalisation, it is applied to Gabor features, which are used for comparison purposes, only.

3. TEXTURE REPRESENTATION

Let us assume each texture to be composed of C spectral planes measured by corresponding sensors. The spectral planes are 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. Each spectral plane is factorised into K levels of the Gaussian pyramid and subsequently each pyramid level is modelled by a dedicated MRF model. All C spectral planes are factorised using the same pyramid thus the corresponding multispectral pixels for every pyramid level have C components $Y_r = [Y_{r,1},\ldots,Y_{r,C}]^T$.

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

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

with a multiindex $r = (r_1, r_2)$ and similarly the multiindices s, t. 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 Ω. 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 $p(\gamma \,|\, Y^{(t-1)})$ given the known process history $Y^{(t-1)}$, $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ 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 [11] that the conditional mean value is:

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

The following notation is used in (5):

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

and V_0 is a positive definite matrix. It is easy to check (see [11]) 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 [11]. The optimal contextual neighbourhood I_r can be found analytically by maximising the corresponding posterior probability [11].

Textural features for each texture resolution 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.

Because the CAR models analyse a texture in some fixed movement direction, we have experimented with additional directions to capture supplementary texture properties. In that case, the texture is optionally analysed in two orthogonal directions: row-wise and column-wise. Subsequently, the estimated features for both directions are concatenated into a common feature vector.

3.2 GMRF Factor Model

The alternative representation assumes that spectral planes of the $k-$th resolution 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 centered 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. An optimal neighbourhood is detected using the correlation method [10] favouring neighbours locations corresponding to large correlations over those with small correlations. The GMRF model for centered values $Y_{r,j}$ can be expressed also in the matrix form (4), but the driving noise correlation structure (diagonal Σ) is now more complex:

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

$\sigma_j, a_{s,j} \ \forall s \in I_r$ are unknown parameters. The parameter estimation of the GMRF model is complicated because either Bayesian or ML 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 evaluated for an image index lattice I has the form

$$\begin{aligned}
\hat{\gamma}_j^T &= [a_{s,j} \quad \forall s \in I_r]^T \\
&= \left[\sum_{\forall s \in I} Z_{s,j} Z_{s,j}^T \right]^{-1} \sum_{\forall s \in I} Z_{s,j} Y_{s,j} \ ,
\end{aligned} \tag{11}$$

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

where $j = 1 \ldots C$. 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

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

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 centered $Y_{r,j}$:

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. As a texture retrieval measure, the feature vectors distance between two textures T, S is computed using the L_1 norm or alternatively using $L_{1\sigma}$ norm:

$$L_1(T,S) = \sum_{i=0}^{p} \left| f_i^{(T)} - f_i^{(S)} \right| \ , \tag{16}$$

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

where $\sigma(f_i)$ is the standard deviations of a feature f_i computed over all database, and p is the size of the feature vector.

3.4 Alternative Features

Our proposed measures 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).

The Gabor filters [1], [22] can be considered as orientation and scale tunable edge and line detectors and statistics of Gabor filter responses in a given region are used to characterise the underlying texture information. A two dimensional Gabor function $g(r) : \mathbb{R}^2 \to \mathbb{C}$ can be specified as

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

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 [18] for details. The Gabor features [18] are defined as the mean μ_j and the standard deviation σ_j of the magnitude of filter responses computed separately for each spectral plane and concatenated into the feature vector. These feature vectors are compared in the $L_{1\sigma}$ norm (17).

The Opponent Gabor features [16] are 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 distance between textures T, S using the Opponent Gabor features is measured as the sum

$$L_{2\sigma}(T,S) = \sqrt{\sum_{i=0}^{p} \left(\frac{f_i^{(T)} - f_i^{(S)}}{\sigma(f_i)}\right)^2} \ , \tag{18}$$

where $\sigma(f_i)$ is again the standard deviations of feature f_i computed over all database, and p is the size of feature vector. 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.

The steerable pyramid [23] is an over complete wavelet decomposition similar to the Gabor decomposition. The pyramid is built up of responses to steerable filters, which are oriented complex analytic filters that are polar separable in the Fourier domain. Each level of pyramid extracts certain frequency range and all levels (except the highest and the lowest one) are further decomposed to different orientations. Our steerable pyramids are composed of 4 orientation bands and 4 pyramid levels, which is in accordance with [23]. We use steerable pyramid features 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 [23]. The feature vectors are compared using the $L_{1\sigma}$ norm (17). Again, we have tested steerable pyramid features with and without the channel normalisation.

Local Binary Patterns (LBP) [20] are histograms of texture micro patterns. For each pixel, a circular neighbourhood around the pixel is sampled, P is the number of samples and R is the radius of the circle. Sampled points values are thresholded by a central pixel value and the pattern number is formed as follows:

$$LBP_{P,R} = \sum_{s=0}^{P-1} \text{sign}\left(Y_s - Y_c\right) 2^s, \tag{19}$$

where *sign* is the signum function, Y_s 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 LBP^{u2} comprise only a subset of these patterns. Moreover, the features can be also made rotation invariant (see [20]). All LBP histograms were normalised to have a unit L_1 norm. The similarity between texture feature vectors T, S is defined as

$$L_G(T,S) = \sum_{i=1}^{p} f_i^{(T)} \log \frac{f_i^{(T)}}{f_i^{(S)}} \ . \tag{20}$$

We have tested features: $\text{LBP}_{8,1+8,3}$ and $\text{LBP}_{16,2}^{u2}$ which demonstrated the best performance in the test with illumination changes [17], [21] (test set Outex 14). We have also comprised rotation invariant feature $\text{LBP}_{16,2}^{riu2}$. The features were computed either on grey-scale images or on each spectral plane separately and concatenated to form the feature

Table 1: Size of feature vectors. Further, feature vectors with normalisation have the same length and feature vectors of CAR models computed in two directions (denoted as 2x) have double length.

	experiment	
method	1	2
Gabor f.	144	144
Gabor f., grey img.	48	48
Opponent Gabor f.	**252**	**252**
Steerable pyramid	2904	2904
$\mathrm{LBP}_{8,1+8,3}$	1536	1536
$\mathrm{LBP}^{u}_{16,2}$	729	729
$\mathrm{LBP}_{8,1+8,3}$, grey img.	**512**	**512**
$\mathrm{LBP}^{u}_{16,2}$, grey img.	243	243
$\mathrm{LBP}^{u}_{16,2}$, grey img., rotinv	18	18
2CAR-KL, $\alpha_1\alpha_2\alpha_3$	204	54
2CAR-KL, $\alpha_1\alpha_3$	**192**	**48**
2CAR-KL, α_1	180	42
GMRF-KL, $\alpha_4\alpha_5$	192	48
GMRF-KL, α_5	180	42
GMRF-KL	168	36
3CAR, $\alpha_1\alpha_2\alpha_3$	180	48
3CAR, $\alpha_1\alpha_3$	**176**	**44**
3CAR, α_1	172	40

vector. Normalisation to change of illumination brightness or spectrum change is not necessary.

4. EXPERIMENTS

We demonstrate the performance of the proposed illumination invariant MRF features on the Outex database [19], which consists of texture 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. Additionally, all the textures were acquired with a fixed camera position.

We tested three proposed MRF models: 2CAR (2-dimensional), 3CAR (3-dimensional) and GMRF. All models were computed over K levels of the Gaussian pyramids, which were built on C spectral planes or spectral planes decorrelated by the Karhunen-Loeve transformation. The CAR models were tested with and without additional features $\alpha_1, \alpha_2, \alpha_3$ in their feature vectors. Similarly, the GMRF model was tested for feature vectors with and without additional features α_4, α_4 (see Section 3.3). Moreover, due to the directionality of CAR models, we have also tested CAR models estimated in two orthogonal directions (in tables denoted by *2x* suffix).

The proposed measures were compared with the following alternatives: Gabor features, Opponent Gabor features, Steerable pyramid features, all with and without spectral channels normalisations, and also the LBP features (see details in Section 3.4). 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.

4.1 Experiment 1

The first experiment is an illumination invariant texture retrieval based on the Outex texture database. The test set consists of all 318 textures, each with 3 different illuminations, without any rotation and with 100 dpi resolution. All textures were cropped to size 512×512.

We have tested texture retrieval using every texture from the test set. The relevant textures were defined as the same texture with the other two illuminations, regardless any texture classes. Therefore there were 2 relevant images present in the test set for each of retrieved textures, a total amount of 3 textures were retrieved. Furthermore, images were degraded with added Gaussian noise to test the robustness of the features. The retrieval performance was measured using recall rate

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

the results are summarised in Tab. 2.

In this test, the proposed illumination invariant MRF features achieved retrieval recall rates over 89%, which clearly present their insensitivity to illumination spectrum variations. The MRF models were computed with the sixth order hierarchical neighbourhood and four levels of the Gaussian pyramid. The LBP features also show their illumination invariance property with 83% recall rate. However, their performance swiftly drops as the standard deviation of additive noise increases. 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. Exceptional results, over 94% recall rate, were achieved by MRF feature compared in $L_{1\sigma}$ norm, however in the following experiment their results are slightly worse than with L_1 norm.

Some examples of retrieved textures are presented in Fig. 1. The results without additive noise are all correct. However, more interesting textures are retrieved at the third and fourth positions. We can observe that features recognise visual similarity of seeds with sand, barley-rice or quartz (the second and the fourth row). In the first and fourth row, there is obvious that the MRF features with pyramids prefer overall structure to micro patterns (such as lines) preferred by the LBP features. This also implies the noise sensitivity of LBP features, as it is shown in the lower half of Fig. 1.

4.2 Experiment 2

Our second experiment was performed on the Outex classification test set number 14 [19]. In this test set, 68 selected textures from the Outex database were treated in the following manner. Twenty subsamples with size 128×128 were extracted from each texture image. The train set consists of 10 samples per texture, all illuminated with the 2586K incandescent CIE A light source. The test set consists of 10 remaining subsamples from each texture, all of them illuminated with 2300K horizon sunlight and 4000K fluorescent TL84. 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 [17], [21].

The highest reported classification accuracy on the test set [17] was 69% for $\mathrm{LBP}^{u2}_{16,2}$ features, which outperformed Gabor features with 66% of accuracy (unfortunately we reached only 54.5%, see the Tab. 3), both features were computed on

Table 2: Illumination invariant texture retrieval from the Outex texture database. Performance is measured as recall rate [%] of 3 retrieved images.

	added noise σ			
method	0	2	4	8
Gabor f.	14.0	13.9	13.6	13.4
Gabor f., grey img.	42.8	42.6	42.3	42.4
Opponent Gabor f.	38.8	37.2	33.8	30.5
Steerable pyramid	19.4	19.1	18.9	18.9
Gabor f., norm.	40.4	38.0	33.1	27.5
Gabor f., grey img, norm.	**53.4**	**58.1**	**58.7**	**56.1**
Opponent Gabor f., norm.	**46.9**	**45.0**	**40.9**	**37.8**
Steerable pyramid, norm.	41.2	41.0	40.5	39.4
$LBP_{8,1+8,3}$	51.5	34.8	25.9	20.0
$LBP^{u}_{16,2}$	47.3	28.4	19.6	11.7
$LBP_{8,1+8,3}$, grey img.	**83.1**	**66.0**	**56.0**	**50.3**
$LBP^{u}_{16,2}$, grey img.	80.6	62.0	49.7	40.8
$LBP^{u}_{16,2}$, grey img., rotinv.	61.5	40.0	29.6	21.3
2CAR-KL, $\alpha_1\alpha_2\alpha_3$	89.7	86.6	81.1	68.6
2CAR-KL, $\alpha_1\alpha_3$	**89.7**	**86.6**	**81.1**	**68.6**
2CAR-KL, α_1	**92.4**	89.2	84.4	75.1
GMRF-KL, $\alpha_4\alpha_5$	82.7	78.2	70.1	56.5
GMRF-KL, α_5	82.7	78.3	70.1	56.4
GMRF-KL	87.7	81.7	75.5	65.3
3CAR, $\alpha_1\alpha_2\alpha_3$	83.9	81.5	76.1	65.6
3CAR, $\alpha_1\alpha_3$	83.9	81.4	76.1	65.6
3CAR, α_1	85.4	81.8	75.7	65.0
2CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$	89.2	86.3	80.5	68.7
2CAR-KL 2x, $\alpha_1\alpha_3$	**89.2**	**86.3**	**80.5**	**68.7**
2CAR-KL 2x, α_1	**91.8**	89.4	83.7	74.7
3CAR 2x, $\alpha_1\alpha_2\alpha_3$	85.1.	82.6	77.5	66.4
3CAR 2x, $\alpha_1\alpha_3$	**85.1**	**82.6**	**77.5**	**66.5**
3CAR 2x, α_1	87.2	83.9	77.5	66.3
2CAR, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	**98.1**	97.9	96.3	92.7
GMRF, α_5, $L_{1\sigma}$	92.4	89.4	86.0	83.8
3CAR, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	87.8	85.7	79.2	67.0
2CAR-KL, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	94.1	91.9	88.6	81.6
2CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	**94.2**	92.9	89.2	81.7
3CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	90.3	88.3	81.8	69.2

Table 3: Results [%] of the Outex classification test set number 14. The classification was performed using three nearest neighbours.

	added noise σ			
method	0	2	4	8
Gabor f.	37.5	37.0	36.2	35.6
Gabor f., grey img .	44.3	43.3	43.2	41.3
Opponent Gabor f.	50.7	49.3	45.3	37.3
Steerable pyramid	37.5	35.9	34.9	32.6
Gabor f., norm.	57.0	59.9	60.3	57.1
Gabor f., grey img, norm.	**54.5**	**61.3**	**63.3**	**62.9**
Opponent Gabor f., norm.	**56.7**	**55.8**	**54.3**	**47.9**
Steerable pyramid, norm.	45.5	45.4	46.8	47.2
$LBP_{8,1+8,3}$	66.8	56.6	48.8	36.7
$LBP^{u}_{16,2}$	62.0	52.9	41.2	28.7
$LBP_{8,1+8,3}$, grey img.	**71.6**	**62.2**	**54.6**	**38.6**
$LBP^{u}_{16,2}$, grey img.	67.6	60.4	49.8	32.9
$LBP^{u}_{16,2}$, grey img., rotinv	56.9	45.2	34.2	19.7
2CAR-KL, $\alpha_1\alpha_2\alpha_3$	66.3	64.1	61.8	55.3
2CAR-KL, $\alpha_1\alpha_3$	**66.3**	**64.1**	**61.8**	**55.3**
2CAR-KL, α_1	**59.8**	59.1	56.5	51.8
GMRF-KL, $\alpha_4\alpha_5$	61.3	60.2	57.1	49.2
GMRF-KL, α_5	61.2	60.3	57.0	49.0
GMRF-KL	55.1	51.8	48.5	45.7
3CAR, $\alpha_1\alpha_2\alpha_3$	58.9	57.5	56.0	49.8
3CAR, $\alpha_1\alpha_3$	58.6	57.5	56.0	49.8
3CAR, α_1	53.7	52.4	49.5	44.6
2CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$	67.5	65.2	61.8	56.4
2CAR-KL 2x, $\alpha_1\alpha_3$	**67.5**	**65.2**	**61.8**	**56.4**
2CAR-KL 2x, α_1	**62.4**	61.1	57.1	54.5
3CAR 2x, $\alpha_1\alpha_2\alpha_3$	61.6	59.1	57.0	50.7
3CAR 2x, $\alpha_1\alpha_3$	**61.5**	**59.0**	**57.0**	**50.7**
3CAR 2x, α_1	57.6	55.4	52.3	46.8
2CAR, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	**63.5**	61.5	59.4	55.1
GMRF, α_5, $L_{1\sigma}$	60.8	60.2	57.2	51.8
3CAR, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	55.6	54.3	52.2	44.0
2CAR-KL, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	62.9	61.1	59.0	55.6
2CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	**64.0**	63.4	63.4	57.9
3CAR-KL 2x, $\alpha_1\alpha_2\alpha_3$, $L_{1\sigma}$	58.5	57.6	52.7	46.3

grey-scale texture images. Moreover, [21] 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 added Gaussian noise to the test set images. Because of small image size, the MRF models have to be restricted to two levels of the Gaussian pyramid and the third order hierarchical neighbourhood. As a consequence the feature vector 2CAR-KL 2x, $\alpha_1\alpha_3$ is about five 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% of classification accuracy. However, the results change dramatically with added noise, the $LBP_{8,1+8,3}$ features drop down to 38.6% showing their vulnerability to added noise. The MRF based features are not so noise sensitive, because Gaussian noise is inherent part of the model. All versions of Gabor features performed better than in the previous experiment, however in most of the cases still worse than the MRF features.

5. CONCLUSIONS

We have proposed new illumination invariant features for content based image retrieval systems. These features are derived from the underlying Markov random field texture representation and they are invariant to variations of brightness and spectrum of illumination sources. We have experimentally verified that proposed features are simultaneously robust to image degradation by the additive Gaussian noise.

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 textures degraded by an additive noise. The best results were achieved with invariants based on the 2-dimensional CAR model with decorrelated spectral planes.

6. ACKNOWLEDGMENTS

This research was supported by the EC project no. FP6-507752 MUSCLE, grants No.A2075302, 1ET400750407 of the Grant Agency of the Academy of Sciences CR and par-

tially by the MŠMT grant 1M0572 DAR and 2C06019.

7. REFERENCES

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

[2] H. F. Chen, P. N. Belhumeur, and D. W. Jacobs. In search of illumination invariants. In *CVPR*, pages I: 254–261, 2000.

[3] O. Drbohlav and M. Chantler. Illumination-invariant texture classification using single training images. In M. Chantler and O. Drbohlav, editors, *Texture 2005.* , pages 31–36, Edinburgh, October 2005. Heriot-Watt University.

[4] G. D. Finlayson. Colour object recognition. Master's thesis, Simon Fraser University, 1992.

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

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

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

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

[9] T. Gevers and A. W. M. Smeulders. Color constant ratio gradients for image segmentation and similarity of texture objects. In *CVPR*, pages 18–25. IEEE Computer Society, 2001.

[10] M. Haindl and V. Havlíček. Prototype Implementation of the Texture Analysis Objects. Technical Report 1939, ÚTIA AV ČR, Praha, 1997.

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

[12] G. Healey and L. Wang. The illumination-invariant recognition of texture in color texture. In *ICCV*, pages 128–133, 1995.

[13] M. A. Hoang and A. W. Geausebroek, Jan-Mark Smeulders. Color texture measurement and segmentation. *Signal Processing*, 85:295–275, 2005.

[14] P. Hsiu Suen and G. Healey. The analysis and reconstruction of real-world textures in three dimensions. *IEEE Trans. on Pattern Anal. and Mach. Intell*, 22(5):491–503, May 2000.

[15] D. Jacobs, P. Belhumeur, and R. Basri. Comparing images under variable illumination. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2000*, volume 1, pages 610 – 617. IEEE, IEEE, June 1998.

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

[17] T. Maenpaa, M. Pietikainen, and J. Viertola. Separating color and pattern information for color texture discrimination. In *ICPR*, pages I: 668–671, 2002.

[18] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. Pattern Anal. and Mach. Intell*, 18(8):837–842, August 1996.

[19] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, and S. Huovinen. Outex- new framework for empirical evaluation of texture analysis algorithms. In *16th ICPR*, pages 701–706, August 2002.

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

[21] M. Pietikainen, T. Maenpaa, and J. Viertola. Color texture classification with color histograms and local binary patterns. In *Workshop on Texture Analysis in Machine Vision*, pages 109–112, 2002.

[22] T. Randen and J. H. Husøy. Filtering for texture classification: A comparative study. *IEEE Trans. Pattern Anal. and Mach. Intell*, 21(4):291–310, April 1999.

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

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

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

[26] J. Weijer, T. Gevers, and J. Geusebroek. Edge and corner detection by photometric quasi-invariants. *IEEE Trans. Pattern Anal. and Mach. Intell*, 27(4):625–630, 2005.

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

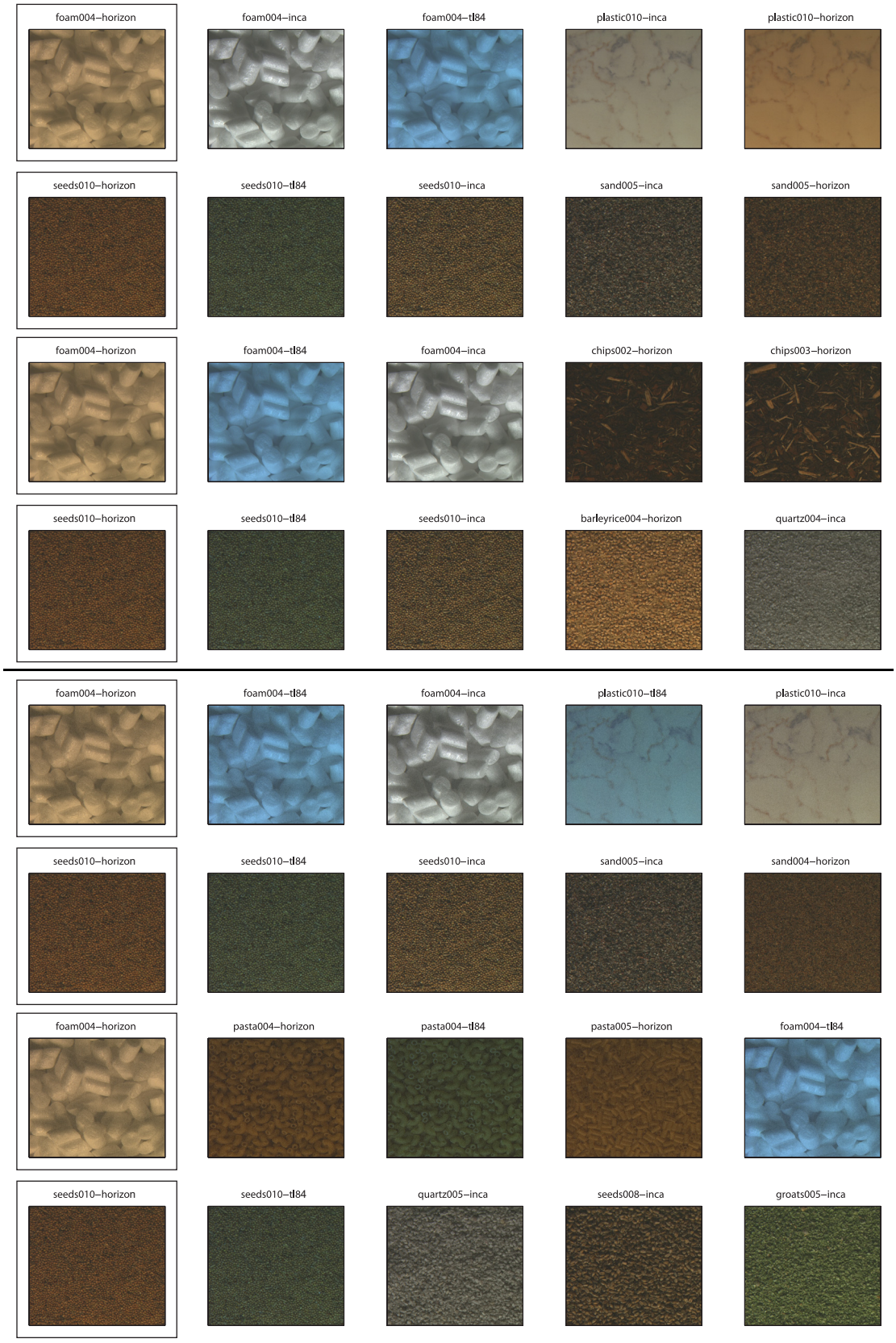

Figure 1: Examples of texture retrieval from the Outex database. The upper half shows retrieval of original textures, the 1st and 2nd row are results of "2CAR-KL, $\alpha_1\alpha_3$", while the 3rd and 4th row are from "LBP$^u_{16,2}$, grey img." method. The lower half shows the same retrievals on the textures with additive noise $\sigma = 8$.