Charles University in Prague
Faculty of Mathematics and Physics
Department of Software Engineering

Abstract of Doctoral Thesis

Query by Pictorial Example

Mgr. Pavel Vácha

Supervisor: Prof. Ing. Michal Haindl, DrSc.

Prague, November 2010

This doctoral thesis was elaborated during a doctoral study at the Faculty of Mathematics and Physics of the Charles University in Prague from 2003 until 2010.

Candidate

Mgr. Pavel Vácha

Supervisor

Prof. Ing. Michal Haindl, DrSc.

Department

Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
Pod Vodárenskou věží 4
182 00, Prague 8

Opponents

Dr. Jan-Mark Geusebroek
Informatics Institute
University of Amsterdam
Amsterdam, The Netherlands

Ing. Ondřej Drbohlav, PhD.
Faculty of Electrical Engineering
Czech Technical University in Prague
Prague, Czech Republic

I–2 board chairman

Prof. Ing. František Plášil, DrSc.

This report has been disseminated on ..... .

The thesis defence will be held on ..... at ..... at MFF UK, Malostranské nám. 25, 118 00, Praha 1, room ..... .

The thesis is available at the Department of doctoral study, MFF UK, Ke Karlovu 3, Praha 2.

Univerzita Karlova v Praze
Matematicko-fyzikální fakulta
Katedra softwarového inženýrství

Autoreferát dizertační práce

Indexace obrazové databáze

Mgr. Pavel Vácha

Školitel: Prof. Ing. Michal Haindl, DrSc.

Praha, listopad 2010

Tato dizertační práce byla vypracována v rámci doktorského studia uchazeče na Matematicko-fyzikální fakultě University Karlovy v Praze v letech 2003–2010.

Uchazeč

Mgr. Pavel Vácha

Školitel

Prof. Ing. Michal Haindl, DrSc.

Školicí pracoviště

Ústav teorie informace a automatizace
Akademie věd České republiky
Pod Vodárenskou věží 4
182 00, Praha 8

Oponenti

Dr. Jan-Mark Geusebroek
Informatics Institute
University of Amsterdam
Amsterdam, Nizozemí

Ing. Ondřej Drbohlav, PhD.
Fakulta elektrotechnická
České vysoké učení technické v Praze
Praha, Česká republika

Předseda oborové rady I–2

Prof. Ing. František Plášil, DrSc.

Autoreferát byl rozeslán dne ..... .

Obhajoba dizertační práce se koná dne ..... v ..... hodin v budově MFF UK, Malostranské nám. 25, 118 00, Praha 1, místnost ..... .

S dizertační prací je možno seznámit se na studijním oddělení doktorského studia MFF UK, Ke Karlovu 3, Praha 2.

Contents

1 Introduction

Content-based image retrieval (CBIR) systems are search engines for image databases, which index images according to their content. A typical task solved by CBIR systems is that a user submits a query image or series of images and the systems is required to retrieve images from the database as similar as possible. Another task is a support for browsing through large image databases, where the images are supposed to be grouped or organised in accordance with similar properties.

Although the image retrieval has been an active research area for many years (see surveys [32, 8]) this difficult problem is still far from being solved. There are two main reasons, the first is so called *semantic gap*, which is the difference between the information that can be extracted from the visual data and the interpretation that the same data have for a user in a given situation. The other reason is called *sensory gap*, which is the difference between a real object and its computational representation derived from sensors, which measurements are significantly influenced by the acquisition conditions.

The *semantic gap* is usually approached by learning of concepts or ontologies and subsequent attempts to recognise them. A system can also learn from the interaction with a user or try to employ combination of multimedia information. However, these topics are beyond the scope of this work and we refer to reviews [32, 21] for further information.

This work concerns with the second problem — finding a reliable image representation, which is not influenced by image acquisition conditions. For example, scene or object can be photographed from different positions and the illumination can vary significantly during a day or be artificial, which causes significant changes in appearance. More specifically, we focus on a reliable and robust representation of homogeneous images (textures), which do not comprise the *semantic gap*.

Invariance

A representation is referred as invariant to a given set of acquisition conditions if it does not change with a variation of these conditions. This invariance property allows to recognise objects or textures in the real world, where the conditions during an image acquisition are usually variable and unknown. Therefore a construction of invariant texture representation is advantageous, because the appearance of many natural or artificial textures is highly dependent on illumination colour, illumination and view point direction, as demonstrated in Figs. 4, 7.

It is necessary to keep in mind that an undesired invariance to a broad range of conditions inevitably reduces the discriminability and aggravates the recognition. (An absurd example is the representation by a constant; it is invariant to all possible circumstances, but it has no use.) Consequently, the optimal texture representation should be invariant to all expected variations of acquisition conditions and still it is required to remain highly discriminative, which are often contrary requirements.

1.1 Thesis contribution

This work is focused on a query by and retrieval of homogeneous images (textures) and on robustness against image acquisition conditions, namely illumination and rotation of texture. It is believed that this thesis contributes to the field of pattern recognition with the following original work:

1. The main contribution is a set of novel illumination invariant features, which are derived from an efficient Markovian textural representation based on modelling by either Causal Autoregressive Random models (2D CAR, 3D CAR) or a Gaussian Markov Random Field (GMRF) model. These new features are proved to be invariant to illumination intensity and spectrum changes and also approximately invariant to local intensity changes (e.g. cast shadows). The invariants are efficiently implemented using parameter estimates and other statistics of CAR and GMRF models.

2. The illumination invariants are extended to be simultaneously rotation invariant. The rotation invariance is achieved either by moment invariants or by combination with circularly symmetric texture model.

Although the proposed invariant features are derived with the assumption of fixed viewpoint and illumination positions, our features exhibit significant robustness to illumination direction variation. This is confirmed in thorough experiments with measurements of Bidirectional Texture Function (BTF) [7], which is currently the most advanced representation of realistic material appearance. Moreover, no knowledge of illumination conditions is required and our methods work even with a single training image per texture. The proposed methods are also robust to image degradation with an additive Gaussian noise.

The proposed invariant representation of textures is tested in the task of texture retrieval and recognition under variation of acquisition conditions, including illumination changes and texture rotation. The experiments are performed on five different textural databases and the results are favourably compared with other state of the art illumination invariant representations. Psychophysical tests with our textural representation indicate its relation to the human perception of textures.

We utilise our features in construction of a system for retrieval of similar tiles, which can be used in decoration industry and we show a feasible application in optimisation of parameters in texture compression, used in computer graphics. Finally, our illumination invariants are integrated into a texture segmentation algorithm and our textural features are applied in the recognition of glaucomatous tissue in retina images.

2 State of the Art

Human perception of textures was studied by Julesz [19], for more than thirty years, and his findings were highly influential for construction of texture discrimination algorithms. The following list briefly reviews currently the most popular textural features.

Histograms of colours or intensity values are the simplest features, but they are not able to describe spatial relations in texture. Their advantage is robustness to various geometrical transformations and easy implementation. A multiresolution histogram [12] includes some spatial relations, as well as co-occurrence matrices [17].

Spatial model parameters characterise a texture by parameters of a chosen model, whose parameters are estimated from a texture image. Models are: MultiResolution Simultaneous AutoRegressive (MR-SAR) model [24], Rotation Invariant Simultaneous AutoRegressive (RI-SAR) model [20], Anisotropic Circular Gaussian Markov Random Field (ACGMRF) [9], the last two are invariant to image rotation.

Gabor features are based on an image decomposition by a set of Gabor filters, which are tunable detectors of edges and lines at different orientations and scales. Subsequently, Gabor features [23] are statistics of the Gabor filter responses. The opponent Gabor features [18] analyses also relations of spectral planes.

Steerable pyramid features [30] employ an image decomposition into scales and orientations, similarly to Gabor filters and other wavelets. The features additionally include correlations among adjacent orientations and scales.

LBP (Local Binary Patterns) [27] are histograms of binarised micro-patterns. These features are illumination invariant and their modifications are additionally rotation invariant: LBP^{riu2} [29], LBP-HF[1].

Textons representation characterizes a texture by a histogram of textons, which are texture primitives learned from a set of images as clusters of filter responses. Rotation invariant MR-8 textons [33] were extended to be simultaneously colour invariant [3].

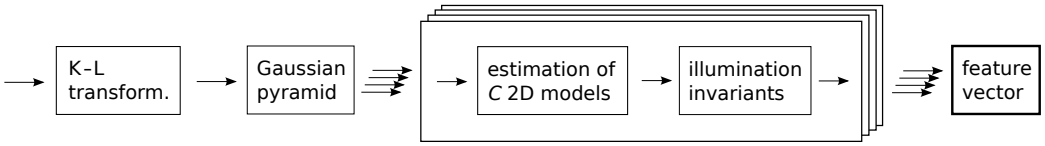

Figure 1: Texture analysis algorithm by means of a set of 2D models and with computation of illumination invariants.

3 Invariant Textural Features

We accept the mathematical definition of texture as a kind of random field, and the texture image is the realisation of this random field. In our textural representation, a texture is locally modelled by a Markov Random Field (MRF) model and the model parameters characterise the texture [I, II]. Subsequently, illumination/colour or rotation invariants are computed from the estimated parameters. We take advantage of a special wide sense Markov model, which enables a fast analytical estimate of its parameters and thus to avoid time-consuming Monte Carlo minimisation prevailing in most of alternative MRF models [22].

Let us assume that a texture is defined on a rectangular lattice I and it is composed of C spectral planes measured by the corresponding sensors (usually {Red, Green, Blue}). The multispectral pixels are $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$, where pixel location $r = [r_1, r_2]$ is a multiindex composed of r_1 row and r_2 column index, respectively.

Algorithm. The texture analysis starts with a spatial factorisation of the texture into K levels of the Gaussian down-sampled pyramid [4]. All spectral planes are factorised using the same pyramid and each pyramid level is either modelled by a 3-dimensional MRF model or a set of C 2-dimensional MRF models. In case of 2D models the image spectral planes are mutually decorrelated by Karhunen-Loève transformation (Principal Component Analysis – PCA) prior to the spatial factorisation by the pyramid. For 3D models the decorrelation is not required. The MRF model parameters are estimated, illumination/colour or rotation invariants are computed, and textural features are formed from them. Finally, the textural features from all the pyramid levels are concatenated into a common feature vector. The overview of the texture analysis algorithm with a set of 2D models is displayed in Fig. 1.

3.1 Markov random field

The each level of Gaussian pyramid level is modelled separately and in the same way. Therefore we omit the level index k and we work generally with multispectral texture pixels Y_r . We use three different models from Markovian family: 3D CAR, 2D CAR, and GMRF. In contrast to MR-SAR model [24], we use restricted shape of neighbourhood, which allows efficient and analytical parameter estimation and we model spatial interaction of spectral planes (colours).

3.1.1 Causal autoregressive random field

The 3D CAR representation assumes that the multispectral texture pixel Y_r can be locally modelled by a 3D CAR model [16] as a linear combination of neighbouring pixels. The shape of contextual neighbourhood is limited to causal or unilateral neighbourhood, see examples in Fig. 2. We denote I_r a selected contextual causal or unilateral neighbour index shift set and its cardinality $\eta = |I_r|$. Let Z_r is a $C\eta \times 1$ data vector, which consists of neighbour pixel values for a given pixel position r . The matrix form of the 3D CAR model is:

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

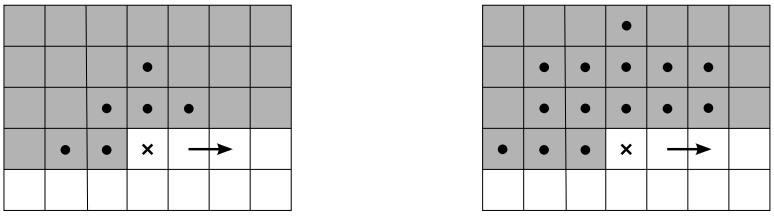

Figure 2: Examples of contextual neighbourhood I_r. From the left, it is the unilateral hierarchical neighbourhood of third and sixth order. X marks the current pixel, the bullets are pixels in the neighbourhood, the arrow shows movement direction, and the grey area indicate permitted pixels.

where $\gamma = [A_s : s \in I_r]$ is the $C \times C\eta$ unknown parameter matrix with square submatrices A_s and r, s are multiindices. 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.

For the 2D CAR models, a set of C models is stacked into formula (1), with diagonal parameters matrices A_s and diagonal noise covariance matrix Σ. Because 2D models are not able to model interspectral relations, the image spectral planes are decorrelated by means of K-L transformation before the estimation of model parameters.

Parameter estimation. The texture is analysed in a chosen direction, where multiindex t changes according to the movement on the image lattice e.g. $t-1 = (t_1, t_2-1)$, $t-2 = (t_1, t_2-2), \ldots$. The task consists in finding the parameter conditional density $p(\gamma \,|\, Y^{(t-1)})$ given the known process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$ and taking its conditional mean as model parameter estimation. Assuming normality of the white noise component ϵ_t, conditional independence between pixels and the normal-Wishart parameter prior, the parameters can be estimated:

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

$$V_{t-1} = \begin{pmatrix} \sum_{r=1}^{t-1} Y_r Y_r^T & \sum_{r=1}^{t-1} Y_r Z_r^T \\ \sum_{r=1}^{t-1} Z_r Y_r^T & \sum_{r=1}^{t-1} Z_r Z_r^T \end{pmatrix} + V_0 = \begin{pmatrix} V_{yy(t-1)} & V_{zy(t-1)}^T \\ V_{zy(t-1)} & V_{zz(t-1)} \end{pmatrix} \ , \tag{3}$$

where V_0 is a positive definite matrix representing prior knowledge, e.g. identity matrix $V_0 = 1_{C\eta+C}$ for uniform prior. Noise covariance matrix Σ is estimated as

$$\begin{aligned} \hat{\Sigma}_{t-1} &= \frac{\lambda_{t-1}}{\psi(t)} \ , \\ \lambda_{t-1} &= V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1} V_{zy(t-1)} \ , \\ \psi(t) &= \psi(t-1) + 1 \ , \ \psi(0) > 1 \ . \end{aligned} \tag{4}$$

The parameter estimation $\hat{\gamma}_t$ can be accomplished using fast and numerically robust recursive statistics. The numerical realisation of the model statistics (2) – (4) is discussed in [16].

Alternatively, the model parameters can be estimated by means of Least Squares (LS) estimation, which leads to the formally same equations as (2) – (4) with zero matrix $V_0 = 0_{C\eta+C}$.

After the estimation of model parameters, the pixel prediction probability $p\left(Y_t | Y^{(t-1)}\right)$ can be computed. Moreover, the optimal contextual neighbourhood I_r can be found analytically. The most probable model M_ℓ with contextual neighbourhood I_r^ℓ can be selected using the Bayesian formula without computation of normalisation constant. Therefore $p(Y^{(t-1)}|M_\ell)$ or its logarithm is maximised, see [16] for details.

3.1.2 2D Gaussian Markov random field

This model is obtained if the local conditional density of the MRF model is Gaussian. 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 of the 3D CAR model (1), but the driving noise ϵ_r and its correlation structure is now more complex:

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

where σ_j, $a_{s,j}$ $\forall s \in I_r$ are unknown parameters. Also topology of the contextual neighbourhood I_r is different, because GMRF model requires a symmetrical neighbourhood.

Parameter estimation. The parameter estimation of the GMRF model is more complicated, because either Bayesian or Maximum Likelihood (ML) estimate requires an iterative minimisation of a nonlinear function. Therefore we use an approximation by the pseudo-likelihood estimator, which is computationally simple although not efficient, see [13] for details.

3.2 Illumination invariance

Illumination conditions of an image acquisition can change due to various reasons. In our approach, we allow changes of brightness and spectrum of illumination sources, and we derived illumination invariants based on parameters and statistics of the previous models. It enabled us to create textural representation, which is invariant to illumination spectrum (colour) and brightness [III, V, XII].

Assumptions. We assume that a planar textured surface is uniformly illuminated and the positions of viewpoint and illumination source remain unchanged. For Lambertian (ideally diffuse) surface reflectance, the two images $\tilde{Y}$, Y acquired with different illumination brightness or spectrum can be transformed to each other by the linear transformation:

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

where B is a $C \times C$ matrix same for all the pixels. We derive that this equation holds even for natural reflectance model of BTF [7] if the surface colour do not depend on camera or light positions. If sensor response functions are changed instead of illumination spectrum, the formula (5) holds again.

Although, the assumption of fixed illumination positions might sound limiting, our experiments with natural and artificial surface materials show that the derived features are very robust even if the illumination positions changes dramatically.

3.2.1 Causal autoregressive random field

Let us assume that two images Y, $\tilde{Y}$ with different illumination are related by equation (5). Consequently, the model data vectors of 3D CAR model (1) are also related by the linear transformation $\tilde{Z}_r = \Delta\, Z_r$, $\forall r$, where Δ is the $C\eta \times C\eta$ block diagonal matrix with blocks B on the diagonal. By substitution of $\tilde{Y}_r$, $\tilde{Z}_r$ into the parameter estimate of 3D CAR model we derived the following relations of model parameters, where accent $(\tilde{\cdot})$ denotes different illumination:

$$\tilde{A}_s = B\, A_s\, B^{-1} \ , \qquad \tilde{\lambda}_t = B\, \lambda_t\, B^T \ , \qquad \forall s \in I_r \ , \ \forall t \in I \ . \tag{6}$$

As a direct consequence the following features are proved to be colour invariant [III]:

1. trace: $\operatorname{tr} A_s$, $\forall s \in I_r$,
2. eigenvalues: $\nu_{s,j}$ of A_s , $\quad \forall s \in I_r$, $j = 1, \ldots, C$,

3. $\alpha_1 = 1 + Z_t^T \, V_{zz(t)}^{-1} \, Z_t$,

4. $\alpha_2 = \sqrt{\sum\limits_{\forall r \in I} \left(Y_r - \hat{\gamma}_t Z_r\right)^T \lambda_t^{-1} \left(Y_r - \hat{\gamma}_t Z_r\right)}$,

5. $\alpha_3 = \sqrt{\sum\limits_{\forall r \in I} \left(Y_r - \mu\right)^T \lambda_t^{-1} \left(Y_r - \mu\right)}$, μ is the mean value vector of Y_r ,

These colour invariants utilise the linear relation (5), which could be considered too general for some applications, because it allows mutual swaps of sensors or spectral planes. In that case, matrix B can be restricted to a diagonal matrix, which models illumination change as multiplication of each spectral plane. For the diagonal B, the formula BA_sB^{-1} do not change the diagonal elements of A_s . Therefore we can alternatively redefine invariants $\nu_{s,j}$:

2′. diagonals: $\nu_s = \operatorname{diag} A_s \,, \quad \forall s \in I_r$.

Determinant based invariants. Additional colour invariants are derived [XII] from determinants $|V_{yy(t-1)}|$, $|V_{zz(t-1)}|$, $|\lambda_{t-1}|$, and probabilities $p\left(Y_t|Y^{(t-1)}\right)$, $\ln p\left(Y^{(t-1)}|M_\ell\right)$:

$$
\begin{aligned}
\beta_1 &= \ln\left(\tfrac{\psi(r)^C}{\psi(t)^C} \, |\lambda_t| \, |\lambda_r|^{-1}\right) \;, & \beta_8 &= \left(\tfrac{\psi(r)^C}{\psi(t)^C} \, |\lambda_t| \, |\lambda_r|^{-1}\right)^{\frac{1}{2C}} \;, \\
\beta_2 &= \ln\left(\tfrac{\psi(r)^C}{\psi(t)^C} \, \left|V_{zz(t)}\right| \left|V_{zz(r)}\right|^{-1}\right) \;, & \beta_9 &= \left(\tfrac{\psi(r)^C}{\psi(t)^C} \, \left|V_{zz(t)}\right| \left|V_{zz(r)}\right|^{-1}\right)^{\frac{1}{2C\eta}} \;, \\
\beta_3 &= \ln\left(\left|V_{zz(t)}\right| |\lambda_t|^{-\eta}\right) \;, & \beta_{10} &= \left(\left|V_{zz(t)}\right| |\lambda_t|^{-\eta}\right)^{\frac{1}{2C}} \;, \\
\beta_4 &= \ln\left(\left|V_{zz(t)}\right| \left|V_{yy(t)}\right|^{-\eta}\right) \;, & \beta_{11} &= \left(\left|V_{zz(t)}\right| \left|V_{yy(t)}\right|^{-\eta}\right)^{\frac{1}{2C}} \;, \\
\beta_5 &= \operatorname{tr}\left\{V_{yy(t)} \; \lambda_t^{-1}\right\} \;, & \beta_{12} &= \sqrt{\left|V_{yy(t)}\right| |\lambda_t|^{-1}} \;,
\end{aligned}
$$

$$
\begin{aligned}
\beta_6 &= \ln\left(\sum_{\forall r \in I} \frac{1}{|I|} p\left(Y_r|Y^{(r-1)}\right) \; \left|V_{yy(t)}\right|^{\frac{1}{2}}\right) \;, \\
\beta_7 &= \ln\Big(\ln p\left(Y^{(t)}|M_\ell\right) \; + \; (\psi(t+1) + C + 1) \ln\left|V_{yy(t)}\right|\Big) \;.
\end{aligned}
$$

Invariants $\alpha_1 - \alpha_3$, $\alpha_{1'}$, $\beta_3 - \beta_7$, $\beta_{10} - \beta_{12}$ are computed with $V_{zz(t)}$, $V_{yy(t)}$, λ_t estimates from all the image pixels, it means t equal to the last pixel position. However, they can be computed from actual estimates at each pixel position as well, which is useful in texture segmentation. Invariants β_1 , β_2 , β_8 , and β_9 are computed from $V_{zz(r)}$, λ_r estimates at different positions r, t in the texture, e.g. first and last pixel position.

For the 2D CAR model, the invariants $\alpha_1 - \alpha_3$, $\beta_1 - \beta_{12}$ are computed for each spectral plane separately with $C = 1$.

Interpretation of invariants. For ideally homogeneous images, the invariants β_1 , β_2 , β_8 , and β_9 are necessary constant. Therefore, these invariants can be regarded as condensed indicators of texture homogeneity. An intuitive interpretation of the other invariants is quite difficult. The invariants α_2 , α_3 are based the statistic λ which is made illumination invariant. The statistic λ is used in the estimation of noise covariance and actually it expresses the model ability to explain the data. Furthermore, the invariants β_4 , β_{11} are the ratios of correlations in the data vectors to correlations in the pixel vectors, which we consider to be a measure of dependency in the contextual neighbourhood.

3.2.2 2D Gaussian Markov random field

The colour invariants for GMRF model are derived analogically to the previous invariants for 2D CAR. The invariants $\operatorname{tr} A_s\,,\ \nu_s = \operatorname{diag} A_s\,, \forall s \in I_r$ are the same. The invariants $\alpha_1 - \alpha_3\,,$ $\beta_1 - \beta_{12}$ are similar, with the following main differences:

1. $\hat{\Sigma} \cdot |I|$ is used instead of λ_t, which is not defined in GMRF model,
2. $\operatorname{abs}|V_{zz}|$ have to be used instead of $|V_{zz}|\,,$ because V_{zz} is not always positive definite in the GMRF model,
3. invariants α_1, β_7 and β_8 do not have their GMRF counterparts.

3.2.3 Local intensity changes

All previous colour invariants were derived with the assumption of uniform illumination. However, we derived [XIV] that most of them are also invariant to locally constant intensity changes, which can be caused by cast shadows or objects with more textured planar surfaces.

More precisely, if the texture image is composed of n copies of homogeneously illuminates texture tiles S, where each tile is modified by multiplication with some constant, then the parameters $\hat{\gamma}$ estimated on the whole image are approximately same as they were estimated on tile S. Moreover, the tiles S can be even different if the statistics $\sum_{r\in S} Z_r Z_r^T$ and $\sum_{r\in S} Z_r Y_r^T$ remain the same, natural examples are stochastic textures. As result, the illumination invariants $\operatorname{tr} A_s\,,\ \nu_{s,j}$ are approximately invariant to local intensity changes. This can be proved also for the invariants $\alpha_2\,,$ $\beta_3 - \beta_5\,,$ and $\beta_{10} - \beta_{12}$.

Texture size independence. The independence to size of texture sample (more data, not scale) is a special case of the previous invariance to local intensity changes. Almost all previously derived colour invariants: $\operatorname{tr} A_s\,,\ \nu_{s,j}\,,\ \alpha_{1'} - \alpha_3\,,\ \beta_1 - \beta_5\,,\ \beta_8 - \beta_{12}$ comply with this independence to sample size. Exceptions are β_6 and β_7, which depend on texture sample size, because probabilities $p(\tilde{Y}_r|\tilde{Y}^{(r-1)})$ and $\ln p(\tilde{Y}^{(t)}|M_\ell)$ include non-linear functions of the number of previously analysed data. Of course, a texture sample with sufficient size is required for a reliable estimation of the model parameters and subsequently the invariant features.

3.3 Rotation invariance

Rotation invariants are textural features that do not change with texture rotation. The important property of rotation invariants is how they retain their discriminability, because without sufficient discriminability the features would be useless despite their invariance. We propose two different methods for the rotation invariance of MRF features [XIV]. The first method computes rotation invariant features before the estimation of MRF parameters. While the second method build rotation invariants after the MRF parameter estimation by means of moment invariants. The scheme of rotation invariant texture analysis is depicted in Fig. 3. We also developed a simple method for normalisation of texture rotation based on estimation of dominant texture orientation [VI].

Real rotation of surfaces. The construction of rotation invariants assume that the surface rotation can be modelled as a rotation of its image. Unfortunately, this assumption does not apply for rough surfaces and illumination near surface plane [6]. However, we imagine a rotation of rough materials as a two step process. In the first step, the material sample and the illumination source are rotated around the same axis, as they were firmly tied together. This step can be modelled as a simple image rotation and it is handled by the proposed rotation invariants. The second step consists of illumination rotation into its final position. This situation is supposed to be dealt with the proposed illumination invariants, despite the fact that they were derived with the assumption of fixed illumination position. The reason is that our experiments with natural surfaces show that the derived illumination invariants are robust to change of illumination direction.

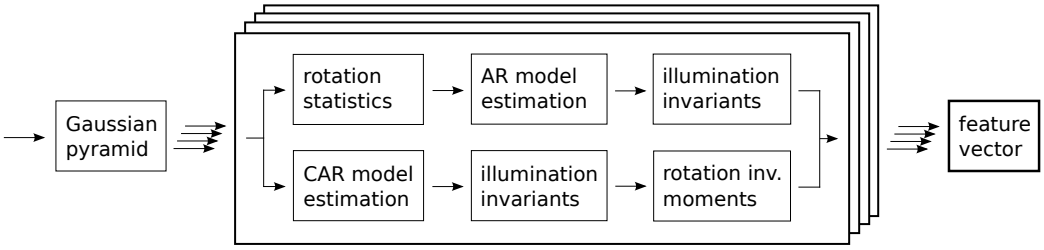

Figure 3: Texture analysis algorithm which combines illumination invariants with two approaches to rotation invariance. It is either an autoregressive model of rotation invariant statistics (RAR) in the upper line, or a causal autoregressive model followed by the computation of rotation moment invariants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

3.3.1 Rotation autoregressive random model

The Rotation Autoregressive Random (RAR) model is inspired by the model [20], which is a regression model of pixel values and averages on concentric circles around these pixels. Although, this model is suitable for modelling of isotropic textures, it has difficulties with anisotropic texture properties. Our model uses multispectal images and extends the regression data with maximum and minimum from circular samples, which enables the model to capture some anisotropic texture properties. The basic modelling equation is similar to (1) with the difference in data vector Z_r, which is composed of mean, maximum, and minimum, all computed separately for each circle in the neighbourhood $I_r^{\odot}$, which is composed of samples on concentric circles. For multispectral images, mean, maximum, and minimum are computed for each spectral plane separately.

The RAR model is used either in 3D or 2D version, which are similar to 3D CAR or 2D CAR models (Section 3.1.1). The differences in the contextual neighbourhood and the datavector Z_r cause that the parameter estimate $\hat{\gamma}_t$ cannot be computed using the analytical Bayesian estimate (2) anymore. Therefore we use the corresponding LS approximation.

Combination with illumination invariants. To achieve simultaneous rotation and colour invariance, the feature vector is composed of the colour invariants derived for CAR models, i.e. $\mathrm{tr}\, A_s$, $\nu_{s,j}$, $\alpha_1 - \alpha_3$, and additionally $\beta_1 - \beta_5$, $\beta_8 - \beta_{12}$ (β_6 and β_7 are not used because they are not valid for the RAR model).

3.3.2 Rotation moment invariants

The rotation moment invariants are used to describe anisotropic texture properties, which are only briefly captured by the RAR model. The CAR model parameters are estimated (Section 3.1.1) and the rotation moment invariants are computed from the illumination invariants $\mathrm{tr}\, A_s$, $\nu_{s,j}$ (Section 3.2.1), according to their position in the unilateral neighbourhood I_r^u. Since the unilateral neighbourhood I_r^u covers only the upper half plane, the values are duplicated in the central symmetry to cover the entire plane, which is advantageous for the rotation invariance of moments.

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

$$c_{pq}^{(f)} = \sum_{r_1} \sum_{r_2} (r_1 + ir_2)^p (r_1 - ir_2)^q f(r_1, r_2) \ , \tag{7}$$

where i is the complex unit. The bilinear interpolation of function $f(r_1, r_2)$ is used to enhance its resolution and precision of computed moments.

The set of invariants should be chosen to be independent, see [11] for more details and additional references. Since our neighbourhood is centrally symmetric, we cannot use any odd-order moment [11]. That is why we use these even-order rotation moment invariants:

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

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

5. second order: $c_{20}^{(\ell)}c_{02}^{(j)}$,

where $\ell = 1$, $j = 2, \ldots, C$ are the individual colour channels. This full set of moment invariants is denoted as $\mathrm{m}_1(model)$. Since the high order moments tend to be numerically unstable, especially for roughly defined f, we also work with the reduced set of invariants denoted as $\mathrm{m}_2(model)$:

1. reduced set of moments: c_{00} , c_{11} , $c_{20}c_{02}$, c_{22} , and $c_{20}^{(1)}c_{02}^{(j)}$.

Combination with illumination invariants. Discrete complex moments c_{pq} are computed from invariants $\operatorname{tr} A_s$ and $\nu_{s,j}$, $j = 1, \ldots, C$ according to their position in the unilateral neighbourhood I_r^u . Each matrix $A_{s=(s_1,s_2)}$ is associated with the position (s_1, s_2) in neighbourhood I_r^u , therefore the input function f is defined from trace of matrices and made centrally symmetric:

$$f_A(r_1, r_2) = \begin{cases} \operatorname{tr} A_{(r_1,r_2)} & (r_1, r_2) \in I_r^u \\ \operatorname{tr} A_{(-r_1,-r_2)} & (-r_1, -r_2) \in I_r^u \\ 0 & \text{otherwise} \end{cases} ,$$

and analogically $f_{\nu,j}(r_1, r_2)$ is constructed from $\nu_{(r_1,r_2),j}$ for each spectral plane j. Subsequently, the previous set of moment invariants (1.– 4.) is computed. The interspectral moment invariant $c_{20}^{(1)}c_{02}^{(j)}$ is computed only from multispectal function $f_{\nu,j}(r_1, r_2)$. Altogether, it makes 34 moment invariants for $C = 3$ and the full set $\mathrm{m}_1(model)$.

The illumination invariants α_1 , α_2 , α_3 , and β_1 – β_{12} are not associated with a position in the contextual neighbourhood, therefore the rotation invariant transformation is not needed, if they are computed with a model with suitable neighbourhood shape. Therefore the illumination invariants α_1 , α_2 , α_3 , and β_1 – β_{12} , computed with hierarchical unilateral neighbourhood, can be added directly into the rotation invariant feature vector.

3.4 Feature vector comparison

All previously described textural representations characterise texture with a feature vector, which is an element of a vector space. Feature vectors are used either directly, i.e. in combination with a classifier to build a class representation, or distance of feature vectors is computed to evaluate similarity of respective textures.

The distance between feature vectors of two textures T, S is computed using Minkowski norms (p-norms) L_1 , $L_{0.2}$, or fuzzy contrast FC_3 proposed in [31]. The norms L_1 , $L_{0.2}$ are preferred to usual L_2 , because they are more robust. Fuzzy contrast, in its symmetric form, is defined as

$$FC_a\,(T, S) = m - \left\{ \sum_{\ell=1}^{m} \min\left\{ \tau(f_\ell^{(T)}), \tau(f_\ell^{(S)}) \right\} - a \sum_{\ell=1}^{m} \left| \tau(f_\ell^{(T)}) - \tau(f_\ell^{(S)}) \right| \right\} , \tag{8}$$

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

where m is the feature vector size, $f_\ell^{(T)}$ and $f_\ell^{(S)}$ are the ℓ-th components of feature vectors of textures T and S, respectively. $\mu(f_\ell)$ and $\sigma(f_\ell)$ are average and standard deviation of the feature f_ℓ computed over all textures in the database.

It is advantageous to compute distance between two feature vectors using the fuzzy contrast, because it normalises different scales of features, which is important for β_ℓ invariants and moment invariants. The drawback is that the fuzzy contrast requires estimate of average and standard deviation of all features.

4 Experiments

The proposed illumination invariant and rotation invariant textural features were tested in the task of natural and artificial material recognition under various circumstances. The experiments were designed to closely resemble real-life conditions and they were conducted on five different textural databases, which differ in the variability of image acquisition conditions and include almost 25 000 of images in total.

At first, we tested the performance of illumination invariant features in texture retrieval and texture classification tasks under various illumination conditions [III, V, IX, XII]. Later, we extended texture recognition tests to illumination variations in combination with different texture rotations and viewpoint positions [VI, XIV]. Such variations of acquisition conditions are usually encountered in an analysis of real-world scenes.

In experiments with texture classification, we used the simple k-Nearest Neighbours (k-NN) classifier, which classifies a texture according to majority of k-nearest training samples. The distance to training samples were computed with L_1 , $L_{0.2}$, or FC_3 dissimilarity measures. In image retrieval applications, we retrieved a given number of images that were nearest according to one of the previous dissimilarities.

4.1 Illumination variation

The performance of the illumination invariant MRF features (see Section 3.2.1) is demonstrated on three image databases, each with different variations of illumination conditions. At first, the Outex texture database [28] was acquired with three illuminations with different spectra and only with slight differences in illumination positions, which complies with our theoretical assumptions. Secondly, the University of Bonn BTF database [25] was acquired with a fixed illumination spectrum and with 91 different illumination directions, which drastically violates our restrictive assumption of fixed illumination position. Finally, the most difficult setup on Amsterdam Library of Textures (ALOT) [3] combined changes in illumination spectrum and direction, and also added slight viewpoint variation. Conditions of these experiments are summarised in Tab. 1.

Proposed features. We tested the proposed illumination invariants based either on 2D CAR, 3D CAR, or GMRF model. The models were usually computed with the 6-th order hierarchical contextual neighbourhood (see Fig. 2), on $K = 4$ levels of the Gaussian pyramids. Optional K-L transformation was indicated with "-KL" suffix. The proposed features were computed according to definitions in Section 3.2.1, the definition of $\nu_{s,j}$ as diagonals of A_s were used together with K-L transformation, otherwise the eigenvalues were employed.

Alternative features. The comparison was performed with the following alternative textural features: Gabor features [23], opponent Gabor features [18], steerable pyramid features [30], and LBP features [29, 27]. The grey value based features as Gabor features and LBP were computed either on grey-scale images or separately on each spectral plane of colour images and concatenated. Moreover, Gabor features and opponent Gabor features were tested with and without normalisation of spectral planes. The mean and standard deviation of features (required by some dissimilarities) were estimated from all images.

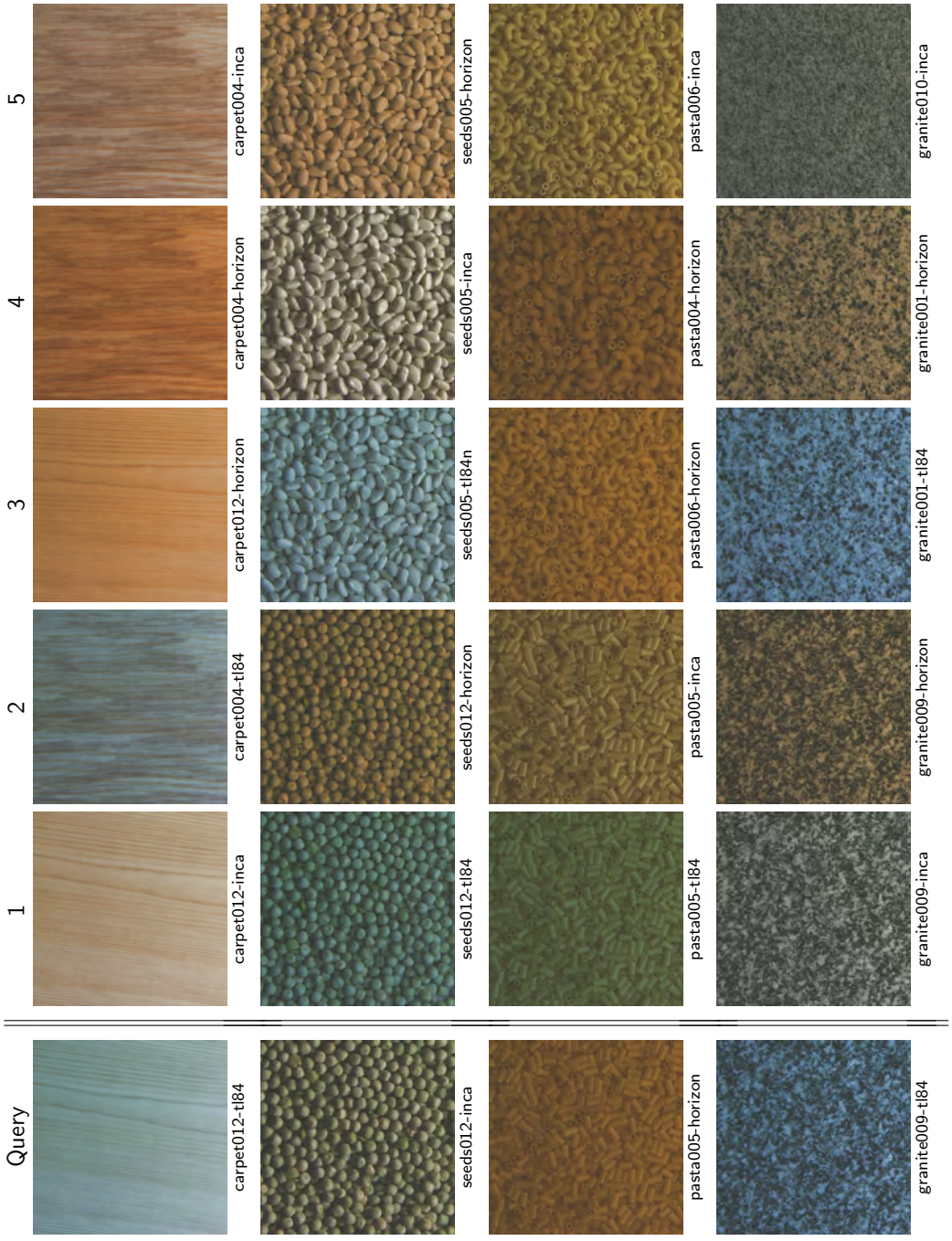

Figure 4: Examples of illumination invariant retrieval from Outex texture database [28], which contains materials illuminated with tree different illumination spectra (inca, horizon, tl84). The query images are followed by retrieved images in order of similarity. The images of query materials acquired under different illumination spectra were successfully retrieved at positions 1 – 3.

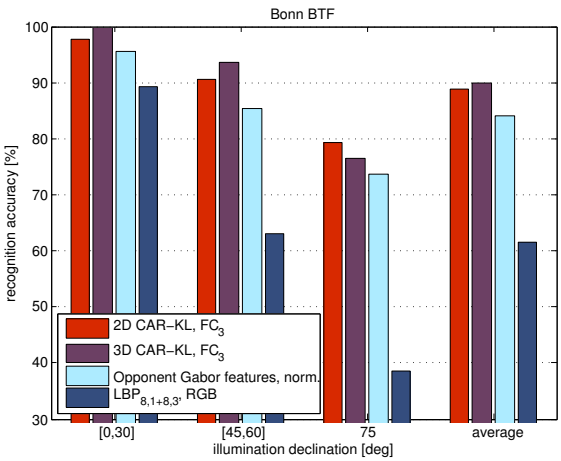

Figure 5: Accuracy of material recognition [%] on the Bonn BTF database [25], which contains materials with 81 illumination directions. A single image per material, with perpendicular illumination, was used for training. The test images, grouped by illumination direction, show performance decrease under more distant lights.

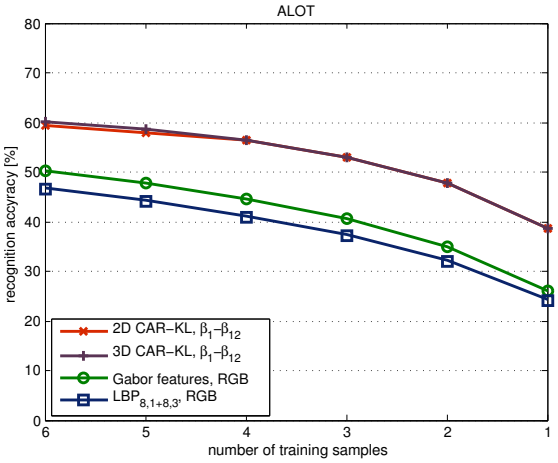

Figure 6: Accuracy of material recognition [%] on the ALOT database [3] for different numbers of training images per material (averaged over 1000 random selections). The superiority of the proposed features (purple and red line) is maintained for all number of training samples.

Figure 7: Examples of materials from the ALOT texture database [3] and their appearance for different camera and light conditions. The two columns on the right are acquired with viewpoint close to the surface (declination angle 60° from the surface macro-normal).

	Experiment					
	i1	i2	i3a	i3b	i4a	i4b
texture database	Outex		Bonn BTF		ALOT	
experiment conditions:						
illumination spectrum	+	+	−	−	+	+
illumination direction	−	−	+	+	+	+
viewpoint azimuth	−	−	−	−	−	−
viewpoint declination	−	−	−	−	+	−
experiment parameters:						
image size (bigger)	512	128	256	256	1536	1536
number of materials	318	68	15	15, 10	200	250

Table 1: Parameters of experiments with illumination invariance and comprised variations of recognition conditions.

Results

The performed experiments confirmed that the proposed illumination invariant features are invariant to changes of illumination spectrum and brightness (see examples in Fig. 4). Moreover, our features also demonstrated considerable robustness to changes of illumination direction and image degradation with an additive Gaussian noise. The reason is that Gaussian noise is an inherent part of the MRF models and the noise is suppressed at higher pyramid levels. This is in contrast with popular LBP features, which exposed their vulnerability to Gaussian noise degradation and illumination direction changes as confirmed in Fig. 5. Most importantly, our illumination invariants retained the discriminability and outperformed the alternative textural features in all but one experiment, see e.g. Figs. 5, 6.

The proposed features excel in recognition of stochastic textures, while the lowest performance was observed with complex regular textures. The most of the discriminative information is concentrated in the invariants ν_s and $\operatorname{tr} A_s$, however, the addition of invariants $\alpha_1 - \alpha_3$, $\beta_1 - \beta_{12}$ still improves the performance. The definition of features ν_s as diagonals of the matrices A_s is preferred to eigenvalues, because it preserves the ordering according to image planes and it should be accompanied with some decorrelation of spectral planes. The size of feature vectors for 2D CAR model were about 260 without β_ℓ invariants and about 400 with all illumination invariants.

Moreover, the fuzzy contrast FC_3 outperformed the other tested dissimilarities of feature vectors. Mean and standard deviation of features, which are required by fuzzy contrast, can be estimated with a sufficiently precision on a small subset of images. If such estimate is not available, we suggest using L_1 norm without β_ℓ invariants.

Additionally, the proposed illumination invariants are also fast to compute and the feature vector has a reasonable low dimension. A disadvantage is that a reliable estimation of the MRF parameters requires a sufficient size of training data. Interactive demonstrations [IV] of the performance of the proposed features are available online[1].

4.2 Rotation and illumination variation

We present performance of the proposed method which combines illumination invariant CAR features with rotation invariance (Section 3.3). The comparison [XIV] was performed on four different texture databases, in three experimental setups, which included almost 300 natural and artificial materials.

The first experiment was performed on Columbia-Utrecht Reflectance and Texture (CUReT) database [7] — dataset from [33], and on ALOT dataset [3]. This experiment is focused on robustness

[1] `http://cbir.utia.cas.cz/outex/`

	Experiment			
	$\varrho 1$		$\varrho 2$	$\varrho 3$
texture database	CUReT	ALOT	Outex	KTH-TIPS2
experiment conditions:				
illumination spectrum	−	+	+	+
illumination direction	+	+	−	+
viewpoint azimuth	+	+	−	−
viewpoint declination	+	+	−	−
experiment parameters:				
image size (bigger)	200	1536	128	200
number of materials	61	200	24	11

Table 2: Parameters of experiments with combined illumination and rotation invariance, including variations of recognition conditions.

of textural features under varying illumination and viewpoint positions, which resembles real-world scenes with natural materials. In the second experiment, we tested features under varying illumination spectrum and texture rotation on Outex classification test [28], which simulates different day light or artificial illuminations. In the third experiment, our results were compared with recently published features on KTH-TIPS2 database [5]. A summary of the experiments and the tested recognition conditions is displayed in Tab. 2.

Proposed features. The proposed illumination and rotation invariant features were again computed on $K = 4$ levels of Gaussian pyramid, with CAR models with the 6-th order hierarchical neighbourhood ($\eta = 14$ neighbours), which corresponds to maximum radius 3 used in the RAR models. The moment based features were composed of either the full set of invariants "$\mathrm{m}_1(model)$" or the reduced set "$\mathrm{m}_2(model)$". Finally, the feature vectors were compared with fuzzy contrast FC_3 (8), since the normalisation of different feature scales is necessary. The feature means and standard deviations, required by fuzzy contrast, were estimated either on a parameter tuning set or on a training set if the tuning set was not available.

Alternative features. The proposed features were compared with the following illumination and rotation invariant features: MR8-NC and MR8-LINC, which were reported with the best performance on ALOT dataset [3]; $\mathrm{LBP}^{riu2}_{P,R}$ [29], and LBP-HF features [1].

Results

The experiments confirmed that our illumination invariants were successfully integrated with two constructions of rotation invariants: either modelling of rotation invariant statistics (RAR model) or moment invariants computed from direction sensitive model parameters (m(CAR) model). Although the construction of invariants assumed image rotation, the proposed features are robust to real rotation of material surface. As the overall best method we suggest the combination "3D RAR + m_2(3D CAR-KL)" or its 2D counterpart if less training data are available. The size of feature vectors was 352 and 304, for 3D and 2D versions, respectively.

In summary, improvements to the best alternative features were 7%, 22%, -3%, 9% for experiments ϱ_1-CUReT, ϱ_1-ALOT, ϱ_2, ϱ_3, respectively. The proposed features performed only slightly worse than LBP features on classification test OUTEX_TC_00014 [28]. However, we argue that although this test is focused on colour invariance, it is not suitable setup, because gray-scale images disable exploitation of interspectral dependences, which are the key properties in illumination spectrum invariance. The performance on CUReT and ALOT datasets can be thoroughly examined in Figs. 8, 9.

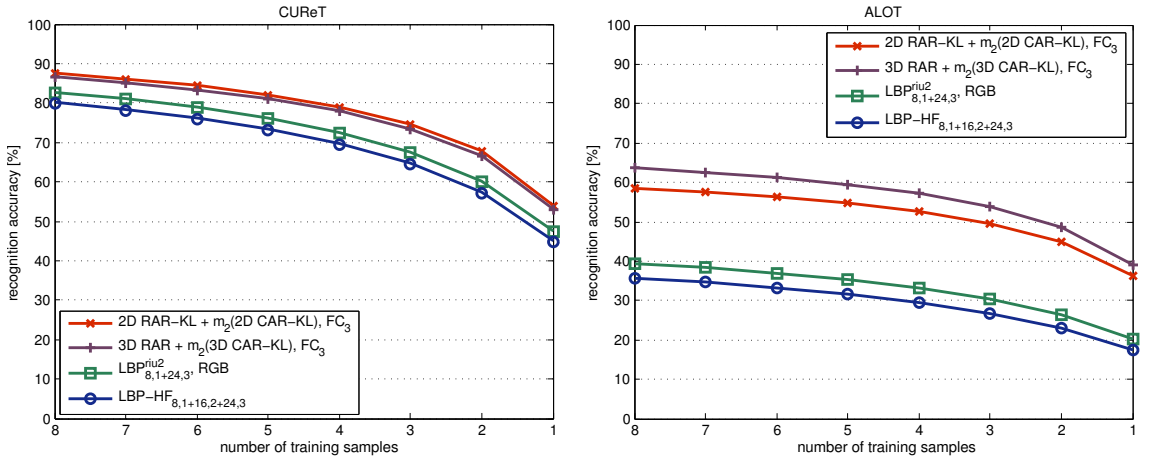

Figure 8: Accuracy of material recognition [%] for CUReT and ALOT datasets, which contains materials under different viewpoint and illumination directions, plus illumination spectrum for ALOT. The training images were randomly selected from the training set and the test images were classified using k-NN classifier, all averaged over 1000 repetitions. The proposed rotation invariant features (purple and red line), which is combination of RAR model parameters and moment invariants of CAR model parameters, outperformed alternative features for all numbers of training images per material. The results are directly comparable with [3], where the best classification accuracy monotonously decreased from 75% to 45% for MR8-LINC features on the CUReT and from 40% to 20% for MR8-NC features on the ALOT dataset.

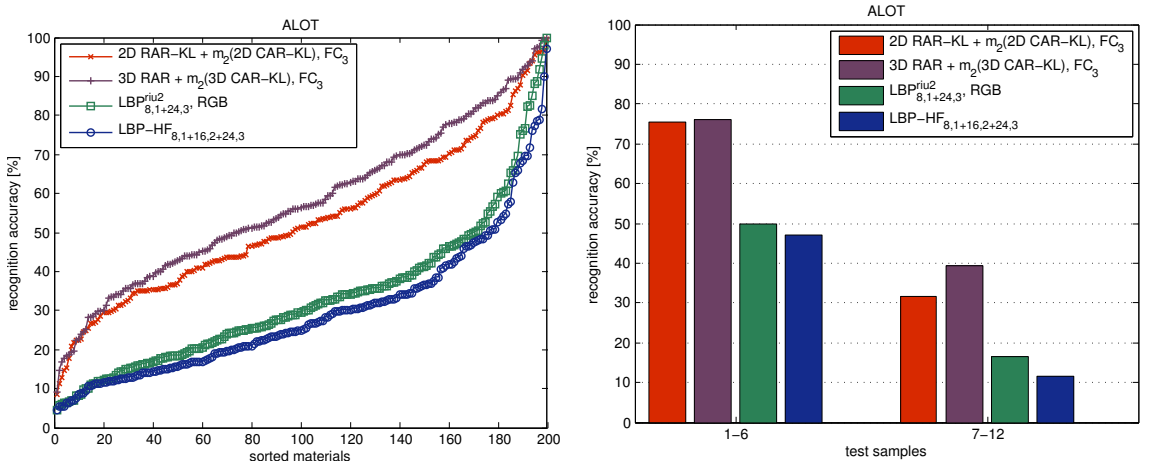

Figure 9: Accuracy of material recognition [%] for the ALOT dataset [3], using 4 training samples per material. On the left, the materials were sorted by their recognition accuracy. This graph implies that the ALOT dataset includes some very easily recognisable materials as well as extremely difficult ones. On the right, the recognition accuracy is grouped by camera position of test samples: top (1-6), from side (7-12), see example images in Fig. 7. The classification accuracy for side viewed images is approximately half of the accuracy for the images from top camera positions, or even worse for LBP features. The reason is that training set do no include such extreme side viewed images and the displayed features are not invariant to perspective projective transformation.

5 Applications

The proposed textural features were applied in various fields, which range from decoration industry to psychophysical studies and a medical application. The second, third, and fourth applications were developed jointly with colleagues from Pattern recognition department and DAR research centre.

5.1 Content-based tile retrieval system

Firstly, we present the content-based tile retrieval system [X], which utilises the proposed colour invariant textural features, supplemented with colour histograms and LBP features. This computer-aided tile consulting system retrieves tiles from digital tile catalogues, so that the retrieved tiles have as similar pattern and/or colours to the query tile as possible. Examples of query and retrieved images are depicted in Fig. 10. The performance of the system was verified on a large commercial tile database in a visual psychophysical experiment.

The system can be exploited in many ways: A user can take a photo of old tile lining and find a suitable replacement of broken tiles from recent production. Or during browsing of digital tile catalogues, the system can offer another tiles that "you may like" based on similar colours or patterns, which could be integrated into an internet tile shop. Or tiles can be clustered according to visual similarity and, consequently, digital catalogues can be browsed through the representatives of visually similar groups. In all cases, the system benefits from its robustness to illumination changes and possible noise degradation.

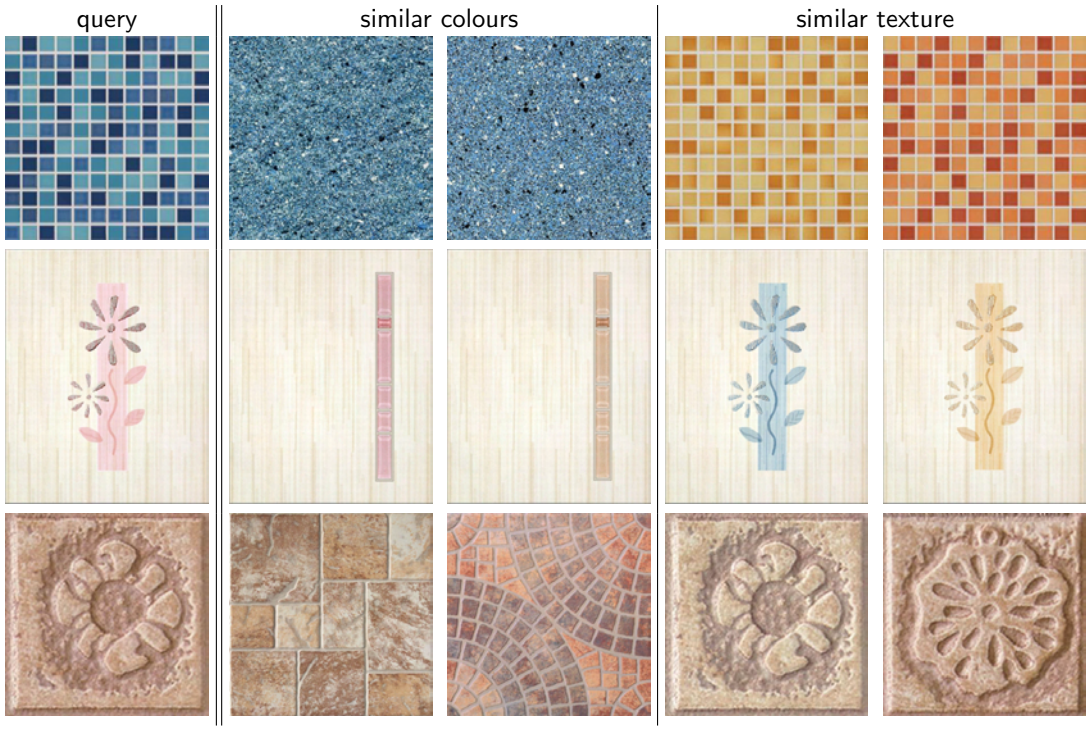

Figure 10: Examples of similar tiles retrieved by our system, which is available online at `http://cbir.utia.cas.cz/tiles/`. The query image, on the left, is followed by two images with similar colours and texture. The images are from the internet tile shop `http://sanita.cz` .

5.2 Illumination invariant unsupervised segmenter

The second application [VII] integrated the proposed colour invariants into the unsupervised texture segmentation method [14, 26], which works with multispectral textures and unknown number of classes. The performance of the presented method was tested on the large illumination invariant benchmark from the Prague Segmentation Benchmark [15] using 21 frequently used segmentation criteria and compared favourably with an alternative segmentation method.

Segmentation is the fundamental process of computer vision and its performance critically determines results of many automated image analysis systems. The segmentation applications [26] include: remote sensing, defect detection, mammography, and cultural heritage applications.

5.3 Psychophysical evaluation of texture degradation descriptors

In the third application [XI], the proposed textural features were successfully used as statistical descriptors of subtle texture degradations. The features were markedly correlated with the psychophysical measurements, employing gaze tracking, and therefore they can be used for automatic detection of subtle texture changes on rendered surfaces in accordance with human vision. Such degradation descriptors are beneficial for compression methods, where the compression parameters have to be set so that the compression is efficient and visual appearance changes remain negligible.

More precisely, the performed experiments were targeted to compression of view- and illumination-dependent textures, which depend on massive measured data of BTF and therefore their compression is inevitable. The descriptors allow automatic tuning of compression parameters to a specific material so that subsequent BTF based rendering methods can deliver realistic appearance of materials [10].

5.4 Texture analysis of the retinal nerve fiber layer in fundus images

Finally, the proposed textural features were applied [VIII] to analysis of images of retinal nerve fibers (RNF) layer, whose texture changes indicate gradual loss of the RNF that it is one of glaucoma symptoms. The early stage detection of RNF losses is desired since the glaucoma is the second most frequent cause of permanent blindness in industrial developed countries.

We have shown that the proposed textural features can be used for discrimination between healthy and glaucomatous tissue, with classification error slightly below 4%. Therefore the features may be used as a part of feature vector in Glaucoma Risk Index, as described in [2] or in a screening program.

6 Conclusions

We proposed several illumination invariant textural representations, which are based on the modelling of local spatial relations. The texture characteristics are modelled by 2D/3D CAR or GMRF models, which are special types from the Markovian model family and which allow a very efficient estimation of their parameters, without the demanding Monte Carlo minimisation. We derived the novel illumination invariants, which are invariant to brightness, illumination colour/spectrum and which are simultaneously approximately invariant to local intensity changes. These illumination invariants were extended to be simultaneously illumination and rotation invariant. On top of that, the experiments with the proposed invariant textural features showed their robustness to illumination direction variations and to the image degradation with an additive Gaussian noise.

The experimental evaluation was performed on five different textural databases: Outex, Bonn BTF, CUReT, ALOT, and KTH-TIPS2, which include images of real-world materials acquired at various conditions. The experiments were designed to closely resemble real-life conditions and the proposed features confirmed their ability to recognise materials in variable illumination conditions and also different viewpoint directions. Our methods do not require any knowledge of acquisition

conditions and the recognition is possible even with a single training image per material, if substantial scale variation or perspective projection is not included. The proposed representation outperformed other state of the art textural representations (among others opponent Gabor features, LBP, LBP-HF, and MR8-LINC), only LBP features performed slightly better in two tests with small texture samples. Although, LBP features are nowadays very popular and effective in many situations, they turned out to be very sensitive to noise degradation and illumination direction variations.

The proposed methods for evaluation of textural similarity are also related to the human perception of textures, according to the performed psychophysical experiments. They were either the low level perception of texture degradations or the subjective ranking of tile similarity.

The presented applications include the content based tile retrieval system, which is able to find tiles with similar textures or colours and, consequently, to ease browsing of digital catalogues. The proposed invariants were also integrated into a segmentation algorithm, in order that computer vision applications can analyse images regardless of illumination conditions. In computer graphics, the features were used for texture degradation description, which opens utilisation in an optimisation of texture compression methods. Last but not least, we applied our textural features in medical imaging and presented their ability to recognise a glaucomatous tissue in retina images.

The results of the invariant texture retrieval or recognition can be reviewed online in our interactive demonstrations[2] so as the presented tile retrieval system[3].

Future research and applications

In the future, we are going to create a texture-based image representation, which will characterise an image by the proposed invariant textural features. The features will be computed from homogeneous regions, which will be extracted by the introduced illumination invariant segmenter. This would be an advantageous extension for current CBIR systems based on colours and SIFT features. Moreover, already reasonable size of our feature vector will be further reduced by feature selection methods [XIII].

We expect that the presented results can be applied in specialised CBIR systems concerning narrow domain images, e.g. medical images or decoration industry, similarly as the presented tile retrieval system. Other possible applications include area of computer vision, since analysis of real scenes inevitably includes description of textures under various light conditions.

A long therm objective is a retrieval from a large medical database, where the texture analysis methods can be successfully exploited. Particularly, we intend to study dermatological images, which would create an online automated dermatology consulting system provided that we will have access to relevant medical images.

[2] `http://cbir.utia.cas.cz`, `http://cbir.utia.cas.cz/rotinv/`
[3] `http://cbir.utia.cas.cz/tiles/`

Selected Bibliography

[1] T. Ahonen, J. Matas, C. He, and M. Pietikäinen. Rotation invariant image description with local binary pattern histogram Fourier features. In *Proc. of the 16th Scandinavian Conference on Image Analysis, SCIA 2009*, *LNCS* 5575, pp. 61–70. Springer-Verlag, 2009.

[2] R. Bock, J. Meier, G. Michelson, L. G. Nyúl, and J. Hornegger. Classifying glaucoma with image-based features from fundus photographs. In *Proc. of the 29th DAGM conference on Pattern recognition*, pp. 355–364. Springer-Verlag, 2007.

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

[4] P. J. Burt. Fast algorithms for estimating local image properties. *Computer Vision, Graphics, and Image Processing*, 21(3):368–382, 1983.

[5] B. Caputo, E. Hayman, and P. Mallikarjuna. Class-specific material categorisation. In *Proc. of the 10th IEEE International Conference on Computer Vision, ICCV 2005*, pp. 1597–1604. IEEE, 17-21 October 2005.

[6] M. Chantler. Why illuminant direction is fundamental to texture analysis. *IEE Proc. - Vision, Image and Signal Processing*, 142(4):199–206, August 1995.

[7] K. Dana, B. Van-Ginneken, S. Nayar, and J. Koenderink. Reflectance and Texture of Real World Surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[8] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Image retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys*, 40(2):1–60, 2008.

[9] H. Deng and D. A. Clausi. Gaussian MRF rotation-invariant features for image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):951–955, July 2004.

[10] J. Filip and M. Haindl. Bidirectional texture function modeling: A state of the art survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):1921–1940, 2009.

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

[12] E. Hadjidemetriou, M. Grossberg, and S. Nayar. Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):831–847, July 2004.

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

[14] M. Haindl and S. Mikeš. Unsupervised texture segmentation using multispectral modelling approach. In Y. Tang, S. Wang, D. Yeung, H. Yan, and G. Lorette, editors, *Proc. of the 18th International Conference on Pattern Recognition, ICPR 2006*, pp. 203–206. IEEE, 20-24 August 2006.

[15] M. Haindl and S. Mikeš. Texture segmentation benchmark. In B. Lovell, D. Laurendeau, and R. Duin, editors, *Proc. of the 19th International Conference on Pattern Recognition, ICPR 2008*, pp. 1–4. IEEE, 8-11 December 2008.

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

[17] R. M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5):786–804, May 1979.

[18] A. Jain and G. Healey. A multiscale representation including opponent colour features for texture recognition. *IEEE Transactions on Image Processing*, 7(1):124–128, January 1998.

[19] B. Julesz. Early vision and focal attention. *Reviews of Modern Physics*, 63(3):735–772, July 1991.

[20] R. L. Kashyap and A. Khotanzad. A model-based method for rotation invariant texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(4):472–481, July 1986.

[21] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2(1):1–19, 2006.

[22] S. Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer Publishing Company, Incorporated, 2009.

[23] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, August 1996.

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

[25] J. Meseth, G. Müller, and R. Klein. Preserving realism in real-time rendering of bidirectional texture functions. In *OpenSG Symposium 2003*, pp. 89–96. Eurographics Association, Switzerland, April 2003.

[26] S. Mikeš. *Image Segmentation*. PhD thesis, Charles University in Prague, Prague, 2010.

[27] T. Mäenpää and M. Pietikäinen. Classification with color and texture: jointly or separately? *Pattern Recognition*, 37(8):1629–1640, 2004.

[28] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, and S. Huovinen. Outex - new framework for empirical evaluation of texture analysis algorithms. In *Proc. of the 16th International Conference on Pattern Recognition, ICPR 2002*, vol. 1, pp. 701–706. IEEE, 11-15 August 2002.

[29] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, Jul 2002.

[30] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1):49–70, 2000.

[31] S. Santini and R. Jain. Similarity measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9):871–883, September 1999.

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

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

List of Publications

[I] P. Vácha. Texture similarity measure. In J. Safrankova, editor, *WDS'05 Proceedings of Contributed Papers: Part I - Mathematics and Computer Sciences*, pages 47–52, Prague, 2005. Charles University in Prague Faculty of Mathematics and Physics, Matfyzpress.

[II] M. Haindl and P. Vacha. Illumination invariant texture retrieval. In Y. Tang, S. Wang, D. Yeung, H. Yan, and G. Lorette, editors, *Proceeding of the 18th International Conference on Pattern Recognition, ICPR 2006*, volume 3, pages 276–279. IEEE, 20-24 August 2006.

[III] P. Vacha and M. Haindl. Image retrieval measures based on illumination invariant textural MRF features. In N. Sebe and M. Worring, editors, *Proceedings of ACM International Conference on Image and Video Retrieval, CIVR 2007*, pages 448–454. ACM, 9-11 July 2007.

[IV] P. Vacha and M. Haindl. Demonstration of image retrieval based on illumination invariant textural MRF features. In N. Sebe and M. Worring, editors, *Proceedings of ACM International Conference on Image and Video Retrieval, CIVR 2007*, pages 135–137. ACM, 9-11 July 2007.

[V] P. Vacha and M. Haindl. Illumination invariants based on Markov random fields. In B. Lovell, D. Laurendeau, and R. Duin, editors, *Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008*, pages 1–4. IEEE, 8-11 December 2008.

[VI] P. Vacha and M. Haindl. Illumination invariant and rotational insensitive textural representation. In *Proceeding of IEEE International Conference on Image Processing, ICIP 2009*, pages 1333–1336. IEEE, 7-10 November 2009.

[VII] M. Haindl, S. Mikes, and P. Vacha. Illumination invariant unsupervised segmenter. In *Proceeding of IEEE International Conference on Image Processing, ICIP 2009*, pages 4025–4028. IEEE, 7-10 November 2009.

[VIII] R. Kolář and P. Vacha. Texture analysis of the retinal nerve fiber layer in fundus images via Markov random fields. In O. Dössel and W. C. Schlegel, editors, *World Congress on Medical Physics and Biomedical Engineering*, volume 25/11 of *IFMBE Proceedings*, pages 247–250. Springer-Verlag, 7-12 September 2009.

[IX] P. Vacha and M. Haindl. *Pattern Recognition, Recent Advances*, chapter Illumination Invariants Based on Markov Random Fields, pages 253–272. In-Teh, Vukovar, Croatia, 2010.

[X] P. Vacha and M. Haindl. Content-based tile retrieval system. In E. R. Hancock, R. C. Wilson, T. Windeatt, I. Ulusoy, and F. Escolano, editors, *Structural, Syntactic, and Statistical Pattern Recognition*, volume 6218 of *Lecture Notes in Computer Science*, pages 434–443. Springer-Verlag, 2010.

[XI] J. Filip, P. Vacha, M. Haindl, and P. R. Green. A psychophysical evaluation of texture degradation descriptors. In E. R. Hancock, R. C. Wilson, T. Windeatt, I. Ulusoy, and F. Escolano, editors, *Structural, Syntactic, and Statistical Pattern Recognition*, volume 6218 of *Lecture Notes in Computer Science*, pages 423–433. Springer-Verlag, 2010.

[XII] P. Vacha and M. Haindl. Natural material recognition with illumination invariant textural features. In *Proceedings of the 20th International Conference on Pattern Recognition, ICPR 2010*, pages 858–861. IEEE, 23-26 August 2010.

[XIII] P. Somol, P. Vácha, S. Mikeš, J. Hora, P. Pudil, and P. Žid. Introduction to Feature Selection Toolbox 3 - the C++ library for subset search, data modeling and classification. Technical Report UTIA TR No. 2287, Czech Academy of Sciences, 2010.

[XIV] P. Vacha, M. Haindl, and T. Suk. Colour and rotation invariant textural features based on Markov random fields. *Pattern Recognition Letters*, submitted.