



Journal of Information and Telecommunication

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tjit20

Texture recognition under scale and illumination variations

Pavel Vácha & Michal Haindl

To cite this article: Pavel Vácha & Michal Haindl (2024) Texture recognition under scale and illumination variations, Journal of Information and Telecommunication, 8:1, 130-148, DOI: 10.1080/24751839.2023.2265190

To link to this article: https://doi.org/10.1080/24751839.2023.2265190

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



0

Published online: 07 Oct 2023.

ſ	Ø,
~	

Submit your article to this journal 🖸

Article views: 172



View related articles 🗹

View Crossmark data 🗹



a OPEN ACCESS

Check for updates

Texture recognition under scale and illumination variations

Pavel Vácha and Michal Haindl

The Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czechia

ABSTRACT

Visual scene recognition is predominantly based on visual textures representing an object's material properties. However, the single material texture varies in scale and illumination angles due to mapping an object's shape. We present a comparative study of the colour histogram, Gabor, opponent Gabor, Local Binary Pattern (LBP), and wide-sense Markovian textural features concerning their sensitivity to simultaneous scale and illumination variations. Due to their application dominance, these textural features are selected from more than 50 published textural features. Markovian features are information preserving, and we demonstrate their superior performance for scale and illumination variable observation conditions over the standard alternative textural features. We bound the scale variation by double size, and illumination variation includes illumination spectra, acquisition devices, and 35 illumination directions spanned above a sample hemisphere. Recognition accuracy is tested on textile patterns from the University of East Anglia and wood veneers from UTIA BTF databases.

ARTICLE HISTORY

Received 19 December 2022 Accepted 26 September 2023

KEYWORDS

Markovian textural features; LBP; Gabor features; scale sensitivity; illumination sensitivity

1. Introduction

A human observer recognizes a visual scene using shape and material attributes. Unfortunately, the surface material's appearance vastly changes under variable observation conditions, negatively affecting its automatic and reliable recognition in numerous artificial intelligence applications. As a consequence, most material recognition attempts apply unnaturally restricted observation conditions, which is shown in Varma and Zisserman (2009), Bell et al. (2015), Gibert et al. (2015).

Scale Invariant Feature Transform (SIFT) features modelled using the Johnson distribution, which allows features to be invariant in rotation, scale, and illumination, are introduced in Hlaing and Zaw (2018). Authors Roy et al. (2018) proposed fractal dimension calculated in the Gaussian scale-space texture representation. Fractal images are combined with LBP images using an indexing function to obtain scale-invariant features. Galois field-based features in Shivashankar et al. (2018) were used for rotation and

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http:// creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

CONTACT Michal Haindl 🖾 haindl@utia.cas.cz 💼 The Institute of Information Theory and Automation, Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 00, Prague, Czechia

scale invariant texture classification. Rotation, scale, and illumination invariant features by Yang et al. (2018) use LBP and log-polar energy-based descriptors in the dual-tree complex wavelet transform domain. Another rotation, illumination, and scale invariance variant of LBP (IRSLBP) was published by Veerashetty and Patil (2020), where partial scale invariance was achieved using three different neighbourhood radii and a scaleselective and noise-robust extended LBP (SNELBP) was proposed in Luo et al. (2022). An LBP modification with some extent of rotation, illumination, and scale invariance was described by the method (Shu et al., 2022). Rotation and scale invariant features based on the LBP and Gabor filter combination were presented in Muzaffar et al. (2023). Although over 50 textural features were compared in Liu et al. (2018), Simon and Uma (2018), we restricted our comparison to the most effective and, thus, dominant textural features.

An ideal model for representing and classifying materials should be capable of capturing fundamental perceptual materials properties. A multi-dimensional visual texture is an appropriate surface reflectance function model paradigm. The sevendimensional Bidirectional Texture Function (BTF) is the best measurable representation, as shown in Haindl and Filip (2013). BTF can be measured simultaneously, even if it is not a trivial task, using state-of-the-art measurement devices, computers, and the most advanced visual data mathematical models, see Haindl (2023). Features derived from such multi-dimensional data models preserve information because they can synthesize data spaces resembling the original measurement data space. The authors have introduced a family of fast multi-resolution Markov random fieldbased models, and in Haindl and Vacha (2015), these models are shown to be robust to illumination conditions.

This paper is an extended version of our ICCCI 2022 paper (Vácha & Haindl, 2022) with the additional comparison of textural features, detailed analysis of recognition with different scale factors, and reciprocity of training and test conditions. This paper's contribution is a joint test of scale and illumination variations to simulate realistic visual scene recognition conditions and we present a comparative analysis with several of the most common alternative textural features representing four alternatives' most commonly used textural features. For this analysis, we use the unique UTIA BTF visual material measurements introduced by Haindl et al. (2015).

2. Markovian textural features

The texture is factorized into *K* levels of the Gaussian down-sampled pyramid and subsequently each pyramid level is modelled by a wide-sense Markovian type of model – the Causal Auto-regressive Random field (CAR) model. Let us assume that each multispectral (colour) texture is composed of *C* spectral planes (usually C=3), $Y_r = [Y_{r,1}, \ldots, Y_{r,C}]^T$ is the multispectral pixel at location *r*. The multiindex $r = (r_1, r_2)$ is composed of row index r_1 and column index r_2 . The spectral planes are mutually decorrelated by the Karhunen–Loéve transformation. The two-dimensional models assume that the *j*-th spectral plane of the pixel at position *r* can be modelled as:

$$Y_{r,j} = \gamma_j Z_{r,j} + \epsilon_r, \tag{1}$$

132 👄 P. VÁCHA AND M. HAINDL

where $Z_{r,i} = [Y_{r-s,i}: \forall s \in I_r]^T$ is the $\eta \times 1$ data vector, ϵ_r is Gaussian white noise with constant but unknown variance, $\gamma_i = [a_{1,i}, \ldots, a_{n,i}]$ is the $1 \times \eta$ unknown parameter vector. Some selected contextual causal or unilateral neighbour index shift set is denoted I_r and $\eta = cardinality(I_r)$, see example in Figure 1. The texture is analysed in a chosen direction, where multi-index t changes according to the movement on lattice the image Ι. Given the known CAR process history $Y^{(t-1),j} = \{Y_{t-1,j}, Y_{t-2,j}, \dots, Y_{1,j}, Z_{t,j}, Z_{t-1,j}, \dots, Z_{1,j}\}, \hat{\gamma}_j$ can be estimated using fast, numerically robust recursive statistics (Haindl, 2012):

$$V_{t-1,j} = \begin{pmatrix} \sum_{u=1}^{t-1} Y_{u,j} Y_{u,j}^T & \sum_{u=1}^{t-1} Y_{u,j} Z_{u,j}^T \\ \sum_{u=1}^{t-1} Z_{u,j} Y_{u,j}^T & \sum_{u=1}^{t-1} Z_{u,j} Z_{u,j}^T \end{pmatrix} + V_0$$

$$= \begin{pmatrix} V_{y,j(t-1)} & V_{zy,j(t-1)} \\ V_{zy,j(t-1)} & V_{z,j(t-1)} \end{pmatrix},$$
(2)

$$\hat{\gamma}_{t-1,j}^{\prime} = V_{z,j(t-1)}^{-1} V_{zy,j(t-1)}, \tag{3}$$

$$\lambda_{t-1,j} = V_{y,j(t-1)} - V_{zy,j(t-1)}^{T} V_{z,j(t-1)}^{-1} V_{zy,j(t-1)},$$
(4)

where the positive definite matrix V_0 represents prior knowledge.

2.1. Colour invariants

Our textural features are $a_{s,j} \forall s \in I_r$, j = 1, ..., C which are colour invariants and additional colour invariant features derived from this model in Haindl and Vácha (2017). The spectral index j is excluded for simplification in invariants (6)–(20).

$$\alpha_1 = \sum_{j=1}^{C} a_{ij} \quad \forall i , \qquad (5)$$

$$\alpha_2 = 1 + Z_t^T \, V_{z(t)}^{-1} \, Z_t \quad , \tag{6}$$

$$\alpha_{3} = \sqrt{\sum_{\forall r \in I} (Y_{r} - \hat{\gamma}_{t} Z_{r})^{T} \lambda_{t}^{-1} (Y_{r} - \hat{\gamma}_{t} Z_{r})} \quad , \tag{7}$$

$$\alpha_4 = \sqrt{\sum_{\forall r \in I} (Y_r - \mu)^T \lambda_t^{-1} (Y_r - \mu)} \} , \qquad (8)$$

			•			
	•	•	•	•	•	
	•	•	•	•	•	
•	•	•	×	-	→	

Figure 1. Unilateral contextual neighbourhood I_r of sixth-order used for CAR model. X marks the current pixel, the bullets are pixels in the neighbourhood, the arrow shows movement direction, and the grey area indicates acceptable neighbourhood pixels.

$$\beta_1 = \ln\left(\frac{\psi(r)}{\psi(t)}|\lambda_t||\lambda_r|^{-1}\right),\tag{9}$$

$$\beta_{2} = \ln\left(\frac{\psi(r)}{\psi(t)} |V_{z(t)}||^{-1}\right)$$
(10)

$$\beta_3 = \ln(|V_{z(t)}||\lambda_t|^{-\eta}), \qquad (11)$$

$$\beta_4 = \ln(|V_{z(t)}||V_{y(t)}|^{-\eta}), \qquad (12)$$

$$\beta_5 = \operatorname{tr}\{V_{y(t)} \ \lambda_t^{-1}\},\tag{13}$$

$$\beta_{6} = \ln\left(\sum_{\forall r \in I} \frac{1}{|I|} p(Y_{r}|Y^{(r-1)}) |V_{y(t)}|^{\frac{1}{2}}\right), \tag{14}$$

$$\beta_7 = \ln(\ln p(Y^{(t)}|M) + (\psi(t+1)+2)\ln|V_{y(t)}|), \tag{15}$$

$$\beta_8 = \left(\frac{\psi(r)}{\psi(t)}|\lambda_t||\lambda_r|^{-1}\right)^{\frac{1}{2}},\tag{16}$$

$$\beta_{9} = \left(\frac{\psi(r)}{\psi(t)} |V_{z(t)}| |V_{z(t)}|^{-1}\right)^{\frac{1}{2\eta}}$$
(17)

$$\beta_{10} = \left(\left| V_{z(t)} \right| |\lambda_t|^{-\eta} \right)^{\frac{1}{2}},\tag{18}$$

$$\beta_{11} = (|V_{z(t)}||V_{y(t)}|^{-\eta})^{\frac{1}{2}},$$
(19)

$$\beta_{12} = \sqrt{|V_{y(t)}| |\lambda_t|^{-1}},$$
(20)

where μ is the mean value of Y_r and $\psi(t)$ is a number of the pixel from the beginning. $p(Y^{(t)}|M)$ is the posterior probability of the model (1), and $p(Y_r|Y^{(r-1)})$ is prediction probability, both defined in Haindl (2012).

We used neighbourhood l_r of sixth order (see Figure 1), where $\eta = 14$, r = 0 corresponding to prior, and t equals to the last pixel in the image. All invariants (6)–(20) were computed on all spectral planes and concatenated into the feature vector as shown on the diagram in Figure 2. The CAR model and colour invariant feature vector were computed on K = 5 Gaussian pyramid levels and in 3 directions, and the features were again concatenated. Finally, the feature vectors were compared with fuzzy contrast FC_3 introduced by Santini and Jain (1999). Downscaling on the Gaussian pyramid is possible as the image provides sufficient resolution. It may be needed for lower-resolution images to use K = 4 levels of the Gaussian pyramid, which was also tested. When the Karhunen–Loéve transformation preceded CAR features computation, they were denoted by the '-KL' suffix.

3. Frequented alternative features

Hundreds of textural features were published, and testing all these features on the extensive UTIA BTF wood database (426 465 wood images, 260 TB of data) is infeasible. Hence, we compare the CAR features with the following most frequented alternatives, each compared with their author's suggested distance.

3.1. Histogram based features

The most straightforward features used in this study are based on histograms of colours or intensity values. Although, these features cannot be considered proper textural features because they are not able to describe spatial relations, which are the critical texture properties, their advantage is robustness to various geometrical transformations, fast and easy implementation. The cumulative histogram proposed in Stricker and Orengo (1995) is defined as the distribution function of an image histogram. The *i*-th bin H_i is computed as $H_i = \sum_{\ell \leq i} h_{\ell}$, where h_{ℓ} is the ℓ -th bin of ordinary histogram. The distance between two cumulative histograms is computed in the L_1 metric. The cumulative histogram is more robust than the ordinary histogram because a small intensity change characterized by a one-bin shift in the ordinary histogram, have only negligible effect on the cumulative histogram.

3.2. Gabor features

The Gabor filters introduced and used by Bovik (1991), Randen and Husøy (1999), Grigorescu et al. (2002), Han and Ma (2007), Li et al. (2009) can be considered as orientation and scale tuneable edge and line (bar) detectors and statistics of Gabor filter responses in a given region are used to characterize the underlying texture information. A two-dimensional Gabor function $g(r):\mathbb{R}^2 \to \mathbb{C}$ can be specified as

$$g(r) = \frac{1}{2\pi\sigma_{r_1}\sigma_{r_2}} \exp\left[-\frac{1}{2}\left(\frac{r_1^2}{\sigma_{r_1}^2} + \frac{r_2^2}{\sigma_{r_2}^2}\right) + 2\pi i Ur_1\right],$$

where σ_{r_1} , σ_{r_2} , U are filter parameters. The convolution of a texture image and Gabor filter extracts the edges of given frequency and orientation range. The whole filter set is 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 the image spectrum; see Manjunath and Ma (1996) for details. Finally, given a single spectral



Figure 2. The texture analysis algorithm flowchart uses 2D random field models; the K-L transformation step is optional.

image with values $Y_{r,j}$, $r \in I$, j = 1, its Gabor wavelet transform is defined as

$$W_{k\phi j}(r_1, r_2) = \int_{u_1, u_2 \in \mathbb{R}} Y_{rj} g^*_{k\phi}(r_1 - u_1, r_2 - u_2) du_1 du_2,$$

where $(\cdot)^*$ indicates the complex conjugate, ϕ and k are orientation and scale of the filter. The Gabor features are defined as the mean μ_j and the standard deviation σ_j of the magnitude of filter responses W. The Gabor features of colour images have been computed on grey images or each spectral plane separately and concatenated to form a feature vector. The distance between feature vectors is measured by $L_{1\sigma}$ metric, where each feature is normalized by its standard deviation (estimated from all datasets) before L_1 metric is computed.

Another extension of the Gabor filters to colour textures by Jain and Healey (1998) is based on adding a chromatic antagonism, while Gabor filters themselves to model the spatial antagonism. Opponent Gabor's features consist of the monochrome part of features: $\eta_{k\phi j} = \sqrt{\sum_{r} W_{k\phi j}^2(r)}$, where $W_{k\phi j}$ is the response to the Gabor filter of orientation ϕ and scale k, j is j-th spectral plane of the image. The opponent part of features is:

$$\psi_{k\phi\phi',j\ell} = \sqrt{\sum_{r} \left(\frac{W_{k\phi,j}(r)}{\eta_{k\phi,j}} - \frac{W_{k\phi',\ell}(r)}{\eta_{k\phi',\ell}} \right)^2}$$

for all spectral planes j, ℓ with $j \neq \ell$ and $|\phi - \phi'| \leq 1$. (Opponent features could also be expressed as the correlation between spectral plane responses.) The distance between feature vectors is measured by $L_{2\sigma}$ metric, where each feature is normalized by its standard deviation (estimated from all datasets) before L_2 metric is computed.

3.3. Local binary patterns

Local Binary Patterns (Ojala et al., 2002) 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 circle's radius. The sampled point values are thresholded by the central pixel value and the pattern number is formed:

$$LBP_{P,R} = \sum_{s=0}^{P-1} \operatorname{sgn}(Y_s - Y_c) 2^s,$$
(21)

where *sgn* is the sign function, Y_s is the grey value of the sampled pixel, and Y_c is the grey value of the central pixel. Subsequently, the histogram of patterns is computed. Because of the 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^{*u*} comprise only a subset of patterns. All LBP histograms were normalized to have unit L_1 norm. As the authors suggested, the similarity between LBP features LBP_{8,1+8,3}, which are combination of features with radii 1 and 3, and uniform version LBP^{*u*}_{16,2} on radius 2. They were computed either on gray images or each spectral plane of the colour image and concatenated. LBP features exist in various modifications:

136 👄 P. VÁCHA AND M. HAINDL

Colour histogram	384	$LBP_{16,2}^{u}$, gray	243
Gabor, gray	48	LBP_{162}^{u} , colour	729
Gabor, colour	144	LBP _{8,1+8,3} , gray	512
Opponent Gabor	252	LBP _{8,1+8,3} , colour	1536
2D CAR-KL	1515	2D CAR-KL ($K = 4$)	1212
2D CAR	1515		

Table 1. The sizes of feature vectors of compared textural features.

(Ahonen et al., 2009; Fu & Wei, 2008; Heikkilä et al., 2009; Khellah, 2011; Liao et al., 2009; Nanni et al., 2012; Zhang et al., 2010), but they have similar behaviour; hence we chose two of their variants as representatives of the whole group as any comparison cannot be considered an exhaustive investigation without the LBP strategy.

3.4. Discussion

We excluded fashionable neural net features due to their uncompetitiveness on often restricted test data in practical applications. They need to be more understood, wasteful, dependent on the net topology, and thus cannot be regarded as well-defined textural features. Moreover, we use only one to six training images, which are insufficient for neural net robust learning. The MRF features outperformed deep Convolutional Neural Networks (CNN) in the bark recognition problem even on extensive training data, as demonstrated in Remeš and Haindl (2019). This result is understandable because MRF features are descriptive, while neural net features are discriminative. Similar results were presented in the extensive comparison of the multilayer NN marble textures classification with 17 LBP feature variants and three key-point texture descriptors types in Sidiropoulos et al. (2021). In their results, the CNN features never outperform all these alternatives. Another comparison where NN - ScatNet, PCANet, FV-AlexNet, and RandNet do not outperform LBP features on Outex, CUReT, ALOT, and KTHTIPS data can be consulted in Liu et al. (2016). However, it would be interesting to include tests with a low number of training samples, which would reveal the robustness of features to various conditions as performed in Burghouts and Geusebroek (2009), Vácha et al. (2011).

Table 1 shows sizes of feature vectors of compared textural features. The fastest features to compute are colour histogram and LBP, followed by five times slower 2DCAR and nine times slower both Gabor and opponent Gabor features.

4. Experiments

We tested the scale sensitivity of the selected textural features on two databases:

- (i) University of East Anglia (UEA) Uncalibrated Image Database was introduced by Finlayson et al. (2000) and consists of patterns under different illumination spectra,
- (ii) wooden BTF measurements from the extensive UTIA BTF database, introduced by Haindl et al. (2015), compose of material images under varying illumination directions.

In both experiments, all images were scaled down to 95%, 90%, 85%, ..., 50% of their original size, and regions with the exact resolution were cropped. Consequently, the

image of scale 50% covers double the size of the original texture image, but with half of the details than scale 100% (see examples in Figure 5 and Figure 11). The training set contains only images with original scales, and the classification accuracy was tested for all scales separately.

Training images per each material were randomly selected from the training set, and the remaining images were classified using the Nearest Neighbour (1-NN) classifier. The number of training images went from 1 to 6, and the results were averaged over 10³ of random selections of training images. Even single training samples were randomly selected so they could have different illumination conditions for each material, making recognition more challenging.

4.1. University of East Anglia uncalibrated image database

The UEA dataset contains 28 textile designs captured with six different devices (4 colour cameras and two colour scanners), and images for cameras were illuminated with three different illumination spectra, which sums up to 394 images in total. Examples of images are displayed in Figure 3, variations of their appearance in Figure 4, and different scales in Figure 5. UEA images are supposed to include even non-linear relations of their values caused by different processing in acquisition devices (gamma correction), and no light calibration was performed; see (Finlayson et al., 2000) for more details. Since the UEA database images include some scale variations, we have corrected this and rescaled the images to have the same scale and resolution. In total, we used 4312 images with 332×275 resolution.

Table 2 summarizes recognition accuracy for the compared features. The 2D CAR features are superior for all test numbers of random training images per material. The classification accuracy of 2D CAR-KL averaged over all scale variations goes from 48.4% for one training sample to 70.2% for six training samples per class. The standard deviation is less than 4 for one training sample and less than 3 for six training samples for all features. The 2D CAR model achieved slightly better results without the Karhunen-Loève transformation, however, we include 2D CAR-KL for more detailed analysis since it had better classification accuracy for experiments reported in Haindl and Vacha (2015). Moreover, the 2D CAR-KL model on K = 4 levels of the Gaussian pyramid achieved lower accuracy than the standard K = 5 levels because the images have sufficient resolution. The only comparable features are opponent Gabor features that achieved similar performance as



Figure 3. Examples of patterns included the UEA database.



Figure 4. The appearance of patterns from the UEA database with varying illumination spectra (3 columns on the left) and additional acquisition devices (2 columns from the right).



Figure 5. The appearance of patterns from the UEA database with varying scales, from the left, the scale factor is 50%, 60%, 75%, 90%, and 100%.

2D CAR-KL with slightly lower accuracy for one training sample. Colour histograms suffer from their sensitivity to colour changes, resulting in low performance (from 15.3% to 32.6%). Even though the colour histograms are robust to scale variation (because they do not describe spatial relations) they are unable to recognize materials under different illumination spectra. LBP features did not perform satisfactorily as well.

Detailed comparison of scale variation is displayed in Figure 6, where we can see as classification accuracy significantly increases if scales of training and test samples are closer to each other. The only exception is colour histogram features, which cannot recognize the same materials on the same scale due to insufficient robustness to the illumination spectra changes (regardless of the number of random training samples per class). The

No. of training samples	1	3	6
Colour histogram	15.3	23.2	32.6
Gabor, gray	29.6	44.5	53.2
Gabor, colour	33.9	46.2	59.7
Opponent Gabor	44.0	61.0	70.1
LBP ^u ₁₆₂ , gray	18.4	35.0	45.8
LBP ^u _{16,2} , colour	17.7	31.2	41.2
$LBP_{8,1+8,3}^{10,22}$, gray	13.9	28.8	38.2
LBP _{8,1+8,3} , colour	14.1	28.2	38.1
2D CAR-KL ($K = 4$)	43.1	58.2	67.0
2D CAR-KL	48.4	62.8	70.2
2D CAR	52.1	66.2	73.0

 Table 2. Classification accuracy [%] averaged over all scales and illumination conditions on the UEA dataset.

Note: Columns display results for the increasing number of training samples per class.



Figure 6. The illustration of the classification accuracy [%] progresses with decreasing scale differences among training and test sets (UEA dataset). On the left, for one training sample, and on the right, for six training samples per class.



Figure 7. The classification accuracy [%] for all combinations of scales among training and test sets on the UEA dataset, one training sample per class was used.

2D CAR-KL features again achieved the best results. Classification accuracy for one training sample starts at 23.1% for half scale factor and goes to 64.9% for the same scale (15% better than alternative features). Opponent Gabor features were slightly better for the highest difference in the scale factor. A similar situation applies to six training samples, where classification accuracy goes from 36.9% to 90.0% for 2D CAR-KL features. However, opponent Gabor features performed better with a significant difference in scale factor (0.5 - 0.7).

4.1.1. Across scales

Figure 7 shows classification accuracy across different training and test scale combinations, with a single training sample averaged over 10³ of random selections. As expected, classification accuracy decreases with a more considerable difference in scale factors. It is worth noting that the last rows of images in Figure 7 correspond to the 2D CAR-KL and opponent Gabor graphs on the left in Figure 6. Interestingly, recognition accuracy on the diagonal decreases as the scale factor goes from 0.5 to 1. This decrease may be caused by the fact that images with a scale factor of 0.5 (left column in Figure 5) cover a larger area of the original material (although being subsampled), so they contain comprehensive information, and the extracted features can be more discriminative.

This discrepancy is more apparent in Figure 8, which compares the original training on a scale factor of 1 and training on a scale swapped to a factor of 0.5. We can see a 15% difference in recognition accuracy between training on a scale of 1 and 0.5 for 2D CAR-KL and opponent Gabor features when the test scale is the same as the training scale. The difference in recognition accuracy decreases with the test scale moving further from the training scale, and there is no difference for training on a scale factor of 0.5 with a test scale factor of 1 or vice versa.

4.2. Wood UTIA BTF database

This study's Wood UTIA BTF database contains veneers from sixty-five varied European, African, and American wood species, examples of images are shown in Figure 9. The UTIA BTF database¹ was measured using the high-precision robotic gonioreflectometer



Figure 8. Classification accuracy [%] on the UEA dataset with one training sample, on the left for the training sample with the scale factor of 1 and on the right with the scale factor of 0.5.

¹http://btf.utia.cas.cz/



Figure 9. Examples of wood veneers included the Wood UTIA BTF database.

described in Haindl et al. (2012), which consists of independently controlled arms with a camera and light. Its parameters, such as angular precision of 0.03 degrees, the spatial resolution of 1000 DPI, or selective spatial measurement, classify this gonioreflectometer as a state-of-the-art device. They measured each wood sample in 81 viewing positions times 81 illumination positions, resulting in 6561 images per sample, 4 TB of data. Because of substantial storage requirements, we took only images for one camera position (top view), and we selected 35 from 81 illumination directions (1 image with a tilt of 0 degrees, 12 images with 30 deg, ten images with 60 deg, and 12 images with 75 deg). The images uniformly represent the space of possible illumination directions; see example images in Figure 10. Images with different scales are displayed in Figure 11. In total, we used 25,025 images with 816 \times 802 resolution.

Table 3 summarizes recognition accuracy for the compared features. The 2D CAR features are superior for all test numbers of random training images per material. The classification accuracy of 2D CAR-KL averaged over all scale variations goes from 45.4% for one training sample to 69.4% for six training samples per class. The standard deviation is less than 4 for one training sample and less than 2 for all features. The 2D CAR-KL model on K= 4 levels of the Gaussian pyramid achieved lower accuracy than the standard K = 5 levels as the images have sufficient resolution. The best alternative is opponent Gabor features. However, their accuracy is more than 20% points lower than 2D CAR-KL. Neither colour histogram nor LBP features performed satisfactorily since the recognition accuracy is less than 13.3% and 25.6% for 1 and 6 training samples, respectively. This is because binarized LBP micropatterns are sensitive to illumination direction, as confirmed by Vácha and Haindl (2012). Additionally, authors (Haindl & Vácha, 2017) show that LBP features are susceptible to even minor scale variations.

The detailed comparison of scale variation is displayed in Figure 12. The classification accuracy increases as the scales of training and test samples are closer (except for



Figure 10. The illustration of the appearance of four veneers from the Wood UTIA BTF database in varying illumination directions. The left column is illuminated from the surface normal, and the direction of illumination tilt increases to the right: 0, 30, 60, 60, and 75 degrees, illumination azimuth is 0, 90, 180, 252, and 345 degrees, respectively.



Figure 11. The appearance of two veneers from the Wood UTIA BTF database in varying scales, from the left, the scale factor is 50%, 60%, 75%, 90%, and 100%.

histogram features). The best results were again achieved by 2D CAR-KL features, where classification accuracy for one training sample starts at 22.7% for half scale and goes to 60.9% for the same scale. This improvement is more than 10% better than the opponent Gabor features for all scale factors. A similar situation holds for six training samples, where

No. of training samples	1	3	6
Colour histogram	10.6	19.7	28.6
Gabor, gray	11.9	20.2	25.1
Gabor, colour	18.0	28.6	36.8
Opponent Gabor	24.0	36.4	44.3
LBP_{162}^{u} , gray	7.2	11.0	13.3
LBP ^u _{16,2} , colour	10.9	16.0	19.6
$LBP_{8,1+8,3}$, gray	10.8	15.8	19.3
LBP _{8,1+8,3} , colour	13.3	21.0	25.6
2D CAR-KL ($K = 4$)	39.9	55.7	64.3
2D CAR-KL	45.4	61.4	69.4
2D CAR	32.8	47.1	55.7

 Table 3. Classification accuracy [%] averaged over all scales and illumination angles on the Wood UTIA

 BTF dataset.

Note: Columns display results for the increasing number of training samples per class.



Figure 12. Classification accuracy [%] progresses with decreasing scale differences among training and test sets (Wood UTIA BTF). On the left, for one training sample, and on the right, for six training samples per class.

classification accuracy goes from 31.5% to 91.0% for 2D CAR-KL features, again more than 10% better than opponent Gabor features.

4.2.1. Across scales

Figure 13 shows classification accuracy across different training and test scale combinations with a single training sample (averaged over 10³ of random selections). As expected, classification accuracy decreases with a more considerable difference in scale factors. It is worth noting that the last rows of images in Figure 13 correspond to the 2D CAR-KL and opponent Gabor graphs on the left in Figure 12. Similarly to the UEA dataset, recognition accuracy on the diagonal decreases as a scale factor goes from 0.5 to 1. The reason is that images with a scale factor of 0.5 (left column in Figure 11) contain more comprehensive information, and the extracted features are more discriminative.

The comparison between the original training on scale factor of 1 and training on scale swapped to factor of 0.5 is displayed in Figure 14. The difference in recognition accuracy between training on a scale of 1 and 0.5 is 9% for 2D CAR-KL, and 6% for opponent Gabor



Figure 13. The classification accuracy [%] for all combinations of scales among training and test sets on the Wood UTIA BTF dataset, one training sample per class was used.



Figure 14. Classification accuracy [%] on the Wood UTIA BTF dataset with one training sample, on the left for the training sample with the scale factor of 1 and on the right with scale factor of 0.5.

features. Again, the difference in recognition accuracy decreases with the test scale moving further from the training scale.

4.2.2. Illumination tilt

The additional experiment utilizes different illumination angles in Wood UTIA BTF and splits classification accuracy for different illumination tilts. The single training sample was fixed to the illumination from a normal surface direction (0-degree tilt), and the remaining images were classified. The classification accuracy is averaged for each illumination tilt: 30, 60, and 75 degrees (12, 10, and 12 images). Training and test sets have the same scaling factor: 1. The results are displayed in Table 4, where classification accuracy decreases as illumination direction moves further from training sample illumination. The last column's average results roughly correspond to the left graph in Figure 12 for test scale factor 1. As the scale variation is absent, the results of LBP and opponent Gabor features are comparable.

Illumination tilt [deg]	30	60	75	avg.
Colour histogram	5.6	1.7	1.5	3.0
Gabor, gray	26.9	10.3	5.5	14.2
Gabor, colour	30.2	13.4	5.4	16.3
Opponent Gabor	50.1	26.1	14.9	30.8
LBP_{162}^{u} , gray	46.8	16.3	9.2	24.1
LBP ^u ₁₆₂ , colour	52.7	20.0	12.9	28.5
$LBP_{8,1+8,3}$, gray	55.4	22.1	12.1	29.9
LBP _{8,1+8,3} , colour	55.8	25.2	13.6	31.5
2D CAR-KL ($K = 4$)	76.4	54.6	32.4	54.5
2D CAR-KL	83.3	62.9	42.1	62.8
2D CAR	74.7	49.7	25.8	50.1

Table 4. Classification accuracy [%] is shown for different illumination tilts (declination angle from the surface normal) without any scale variation (Wood UTIA BTF).

Note: The training sample is illuminated from the surface normal direction.

5. Conclusion

The results indicate that Markovian illumination invariant texture features (2D CAR), based on the Markovian descriptive model, are the most robust textural features for realistic texture classification under natural conditions when learning and classifying textures differ in scale and illumination properties. The 2D CAR features outperformed alternative tested textural features, i.e. the Gabor, opponent Gabor, variants of LBP, or colour histogram texture features. 2D CAR statistical features are analytically derived from the underlying descriptive textural model and can be efficiently, recursively, and adaptively learned. Their additional advantage is their numerically robust estimation. The method's correct recognition accuracy improvements are between 27% and 44%, compared to the LBP features and up to 25% compared to the opponent Gabor features (the second-best alternative). The worst are colour histograms, with an accuracy decrease between 35% and 43%. The colour Markovian textural features were also successfully applied elsewhere in recognition of wood veneers using a smartphone camera, reported by Haindl and Vacha (2015), or tree taxonomy categorization based on bark or coniferous tree needles as shown in Remeš and Haindl (2019). The presented results apply to recognition with bounded scale variation. The full-scale-invariant textural features should be considered for extreme expected scale variation. However, fully invariant features usually lose some discriminability. Thus, each application must carefully balance invariance with expected variability and discriminability.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The Czech Science Foundation project GAČR 19-12340S supported this research.

ORCID

Michal Haindl D http://orcid.org/0000-0001-8159-3685

146 🕒 P. VÁCHA AND M. HAINDL

References

- Ahonen, T., Matas, J., He, C., & Pietikainen, M. (2009). Rotation invariant image description with local binary pattern histogram fourier features. In *SCIA* (pp. 61–70). Springer-Verlag, Berlin Heidelberg
- Bell, S., Upchurch, P., Snavely, N., & Bala, K. (2015). Material recognition in the wild with the materials in context database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3479–3487). IEEE
- Bovik, A. (1991). Analysis of multichannel narrow-band filters for image texture segmentation. *IEEE Transactions on Signal Processing*, 39(9), 2025–2043. https://doi.org/10.1109/78.134435
- Burghouts, G. J., & Geusebroek, J. -M. (2009). Material-specific adaptation of color invariant features. *Pattern Recognition Letters*, 30(3), 306–313. https://doi.org/10.1016/j.patrec.2008.10.005
- Finlayson, G., Schaefer, G., & Tian, G. (2000). *The UEA uncalibrated colour image database*. Technical Report SYS-C00, School of Information System, University of East Anglia, Norwich, UK.
- Fu, X., & Wei, W. (2008). Centralized binary patterns embedded with image Euclidean distance for facial expression recognition. In *Natural Computation, 2008. ICNC '08. Fourth International Conference on* Vol. 4, (pp. 115–119).
- Gibert, X., Patel, V. M., & Chellappa, R. (2015). Material classification and semantic segmentation of railway track images with deep convolutional neural networks. In *2015 IEEE International Conference on Image Processing (ICIP)* (pp. 621–625). IEEE.
- Grigorescu, S. E., Petkov, N., & Kruizinga, P. (2002). Comparison of texture features based on gabor filters. *IEEE Transactions on Image Processing*, *11*(10), 1160–1167. https://doi.org/10.1109/TIP.2002. 804262
- Haindl, M. (2012). Visual data recognition and modeling based on local markovian models. In L. Florack, R. Duits, G. Jongbloed, M. C. Lieshout, and L. Davies (Eds.), *Mathematical methods for signal and image analysis and representation*, volume 41 of *Computational imaging and vision*, chapter 14, (pp. 241–259). Springer London. https://doi.org/10.1007/978-1-4471-2353-8_14
- Haindl, M. (2023). *Bidirectional texture function modeling*, chapter 28, (pp. 1023–1064). Springer International Publishing, Gewerbestrasse 11, 6330 Cham, Switzerland.
- Haindl, M., & Filip, J. (2013). *Visual texture*. Advances in computer vision and pattern recognition. Springer-Verlag London.
- Haindl, M., Filip, J., & Vávra, R. (2012). Digital material appearance: The curse of tera-bytes. *ERCIM News*, *90*, 49–50.
- Haindl, M., Mikeš, S., & Kudo, M. (2015). Unsupervised surface reflectance field multi-segmenter. In G. Azzopardi and N. Petkov, (Eds.), *Computer analysis of images and patterns*, Vol. 9256 of *Lecture notes in computer science*, (pp. 261 – 273). Springer International Publishing.
- Haindl, M., & Vacha, P. (2015). Wood veneer species recognition using Markovian textural features. In
 G. Azzopardi and N. Petkov, (Eds.), *Computer analysis of images and patterns*, volume 9256 of *Lecture notes in computer science*, (pp. 300–311). Springer International Publishing.
- Haindl, M., & Vácha, P. (2017). Scale sensitivity of textural features. In C. Beltrán-Castañón, I. Nyström, and F. Famili (Eds.), Progress in pattern recognition, image analysis, computer vision, and applications: 21st Iberoamerican Congress, CIARP 2016, Lima, Peru, November 8–11, 2016, Proceedings, Vol. 10125 of Lecture notes in computer science, (pp. 84 92). Gewerbestrasse 11, Cham, CH-6330, Switzerland. Springer International Publishing AG.
- Han, J., & Ma, K.-K. (2007). Rotation-invariant and scale-invariant gabor features for texture image retrieval. *Image and Vision Computing*, *25*(9), 1474–1481. https://doi.org/10.1016/j.imavis.2006. 12.015
- Heikkilä, M., Pietikäinen, M., & Schmid, C. (2009). Description of interest regions with local binary patterns. Pattern Recognition, 42(3), 425–436. https://doi.org/10.1016/j.patcog.2008.08.014
- Hlaing, C. S., & Zaw, S. M. M. (2018). Tomato plant diseases classification using statistical texture feature and color feature. In 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS) (pp. 439–444). IEEE.
- Jain, A. K., & Healey, G. (1998). A multiscale representation including opponent color features for texture recognition. *IEEE Transactions on Image Processing*, 7(1), 124–128. https://doi.org/10. 1109/83.650858

- Khellah, F. (2011). Texture classification using dominant neighborhood structure. IEEE Transactions on Image Processing, 20(11), 3270–3279. https://doi.org/10.1109/TIP.2011.2143422
- Li, Z., Liu, G., Jiang, H., & Qian, X. (2009). Image copy detection using a robust gabor texture descriptor. tor. In *Proceedings of the First ACM Workshop on Large-scale Multimedia Retrieval and Mining*, LS-MMRM '09 (pp. 65–72). New York, NY, USA. ACM.
- Liao, S., Law, M. W. K., & Chung, A. C. S. (2009). Dominant local binary patterns for texture classification. *IEEE Transactions on Image Processing*, 18(5), 1107–1118. https://doi.org/10.1109/TIP. 2009.2015682
- Liu, L., Chen, J., Fieguth, P., Zhao, G., Chellappa, R., & Pietikainen, M. (2018). A survey of recent advances in texture representation. *arXiv preprint arXiv:1801.10324*.
- Liu, L., Fieguth, P., Wang, X., Pietikäinen, M., & Hu, D. (2016). Evaluation of lbp and deep texture descriptors with a new robustness benchmark. In *European Conference on Computer Vision*, (pp. 69–86). Springer.
- Luo, Q., Su, J., Yang, C., Silven, O., & Liu, L. (2022). Scale-selective and noise-robust extended local binary pattern for texture classification. *Pattern Recognition*, 132, 108901. https://doi.org/10. 1016/j.patcog.2022.108901
- Manjunath, B. S., & Ma, W. Y. (1996). Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8), 837–842. https://doi.org/10.1109/ 34.531803
- Muzaffar, A. W., Riaz, F., Abuain, T., Abu-Ain, W. A. K., Hussain, F., Farooq, M. U., & Azad, M. A. (2023). Gabor contrast patterns: A novel framework to extract features from texture images. *IEE Access*, *11*, 60324–60334. https://doi.org/10.1109/ACCESS.2023.3280053
- Nanni, L., Lumini, A., & Brahnam, S. (2012). Survey on {LBP} based texture descriptors for image classification. *Expert Systems with Applications*, 39(3), 3634–3641. https://doi.org/10.1016/j.eswa. 2011.09.054
- Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987. https://doi.org/10.1109/TPAMI.2002.1017623
- Randen, T., & Husøy, J. H. (1999). Filtering for texture classification: A comparative study. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(4), 291–310. https://doi.org/10. 1109/34.761261
- Remeš, V., & Haindl, M. (2019). Bark recognition using novel rotationally invariant multispectral textural features. *Pattern Recognition Letters*, 125, 612–617. https://doi.org/10.1016/j.patrec.2019.06. 027
- Roy, S. K., Bhattacharya, N., Chanda, B., Chaudhuri, B. B., & Ghosh, D. K. (2018). Fwlbp: A scale invariant descriptor for texture classification. *arXiv preprint arXiv:1801.03228*.
- Santini, S., & Jain, R. (1999). Similarity measures. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(9), 871–883. https://doi.org/10.1109/34.790428
- Shivashankar, S., Kudari, M., & Hiremath, P. S. (2018). Galois field-based approach for rotation and scale invariant texture classification. *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, 10(9), 56–64. https://doi.org/10.5815/ijigsp
- Shu, X., Pan, H., Shi, J., Song, X., & Wu, X.-J. (2022). Using global information to refine local patterns for texture representation and classification. *Pattern Recognition*, *131*, 108843. https://doi.org/10. 1016/j.patcog.2022.108843
- Sidiropoulos, G. K., Ouzounis, A. G., Papakostas, G. A., Sarafis, I. T., Stamkos, A., & Solakis, G. (2021). Texture analysis for machine learning based marble tiles sorting. In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0045–0051). IEEE.
- Simon, P., & Uma, V. (2018). Review of texture descriptors for texture classification. In *Data Engineering and Intelligent Computing*, (pp. 159–176). Springer.
- Stricker, M. A., & Orengo, M. (1995). Similarity of color images. Vol. 2420, (pp. 381-392). SPIE.
- Vácha, P., & Haindl, M. (2012). Texture recognition using Robust Markovian features. In Salerno, E., Çetin, A., and Salvetti, O. (Eds.), *Computational intelligence for multimedia understanding*, volume 7252 of *lecture notes in computer science*, (pp. 126–137). Springer Berlin/Heidelberg. https://doi. org/10.1007/978-3-642-32436-9_11

- 148 🕒 P. VÁCHA AND M. HAINDL
- Vácha, P., & Haindl, M. (2022). Textural features sensitivity to scale and illumination variations. In Bădică, C., Treur, J., Benslimane, D., Hnatkowska, B., and Krótkiewicz, M., (Eds.), Advances in computational collective intelligence Vol. 1653 of Communications in computer and information science (pp. 237–249), Gewerbestrasse 11, Cham, CH-6330, Switzerland. Springer International Publishing.
- Vácha, P., Haindl, M., & Suk, T. (2011). Colour and rotation invariant textural features based on Markov random fields. *Pattern Recognition Letters*, 32(6), 771–779. https://doi.org/10.1016/j. patrec.2011.01.002
- Varma, M., & Zisserman, A. (2009). A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11), 2032–2047. https://doi.org/10.1109/TPAMI.2008.182
- Veerashetty, S., & Patil, N. B. (2020). Novel lbp based texture descriptor for rotation, illumination and scale invariance for image texture analysis and classification using multi-kernel svm. *Multimedia Tools and Applications*, 79(15–16), 9935–9955. https://doi.org/10.1007/s11042-019-7345-6
- Yang, P., Zhang, F., & Yang, G. (2018). Fusing dtcwt and lbp based features for rotation, illumination and scale invariant texture classification. *IEEE Access*, 6, 13336–13349. https://doi.org/10.1109/ ACCESS.2018.2797072
- Zhang, B., Gao, Y., Zhao, S., & Liu, J. (2010). Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor. *IEEE Transactions on Image Processing*, 19(2), 533–544. https://doi.org/10.1109/TIP.2009.2035882