Pattern Recognition Letters 125 (2019) 612–617

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Bark recognition using novel rotationally invariant multispectral textural features☆

Václav Remeš, Michal Haindl*

The Institute of Information Theory and Automation of the Czech Academy of Sciences, 182 08 Prague, Czech Republic

ARTICLE INFO

Article history:
Received 17 February 2019
Revised 20 June 2019
Accepted 27 June 2019
Available online 27 June 2019

MSC:
41A05
41A10
65D05
65D17

Keywords:
Bark recognition
Tree taxonomy clasification
Spiral Markov random field model

ABSTRACT

We present novel rotationally invariant fully multispectral Markovian textural features applied for the efficient tree bark recognition. These textural features are derived from the novel descriptive multispectral spiral wide-sense Markov model. Unlike the alternative bark recognition methods based on various gray-scale discriminative textural descriptions, we benefit from fully descriptive color, rotationally invariant bark texture representation. The proposed methods significantly outperform the state-of-the-art bark recognition approaches regarding classification accuracy. Both our classifiers outperform convolutional neural network ResNet even on the largest public bark database BarkNet which contains 23 000 high-resolution images from 23 different tree species.

1. Introduction

Automatic bark recognition is a challenging but practical plant taxonomy application which allows fast and non-invasive tree recognition irrespective of the growing season, i.e., whether a tree has or has not its leaves, fruit, needles, or seeds or if the tree is healthy growing or just a dead stump. Automatic bark recognition makes identification or learning of tree species possible without any botanical expert knowledge through, e.g., using a dedicated mobile application. Manual identification of a tree's species based on a botanical key of bark images is a tedious task which would typically consist of scrolling through a book. Since bark cannot be described as easily as leaves or needles [6,20], the user has to go through the whole bark encyclopedia looking for the corresponding bark image. An automatic classifier can even overcome human experts. The article [6] reports 56.6% and 77.8% classification accuracy on the Austrian Federal Forest (AFF) dataset for two human experts, while two published automatic classifiers [18] and [14] reach better results than the top human expert.

The advantage of bark-based features is their relative stability during the corresponding tree's lifetime. Single shrubs or trees have specific bark which can be advantageously used for their identification. It enables numerous ecological applications such as plant resource management or fast identification of invading tree species. Industrial applications can be in sawmills or detection of bark beetle tree infestation.

1.1. Alternative bark recognition methods

An SVM type of classifier and gray-scale LBP features are used in [1]. Their dataset is a collection of 40 images per species, and there are 23 species, i.e., a total of 920 bark color images of local, mostly dry subtropical-climate, shrubs, and trees (acacias, agaves, opuntias, palms). The classifier exploited in [10] is a radial basis probabilistic neural network. The method uses Daubechies 3rd level wavelet-based features applied to each color band in the YC_bC_r color space. A similar method [9] with the same classifier uses Gabor wavelet features. Both approaches use the same test set which contains 300 color bark images. Gabor banks features with a narrow-band signal model in 1-NN classifier were proposed in [5]. The test set has eight species with twenty-five samples per tree category. The author also demonstrates a significant, but expectable, performance improvement when color information was added. The 1-NN and 4-NN classifier [21] represent bark textures by the run length, Haralick's co-occurrence matrix based, and histogram features. These methods are verified on a limited dataset of 160 samples from nine species. Authors in [3] propose a rota-

☆ Handle by Associate Editor Prof. G. Sanniti di Baja.

* Corresponding author.
E-mail addresses: remes@utia.cas.cz (V. Remeš), haindl@utia.cas.cz (M. Haindl).

https://doi.org/10.1016/j.patrec.2019.06.027
0167-8655/

tionally invariant statistical radial binary pattern (SRBP) descriptor to characterize a bark texture. Four types of multiscale LBP features (Multi-Block LBP (MBLBP) with a mean filter, LBP Filtering (LBPF), Multi-Scale LBP (MSLBP) with a low pass Gaussian filter, and Pyramid-based LBP (PLPB) with a pyramid transform) are used in [2]. Two bark image datasets (AFF [6], Trunk12 [19]) were used to evaluate the multi-scale LBP descriptors based bark recognition. The authors observed that multi-scale LBP provides more discriminative texture features than basic and uniform LBP and that LBPF gives the best results over all the tested descriptors on both datasets. The paper [17] proposes a combination of two types of texture features, the gray-level co-occurrence matrix metrics and the long connection length emphasis ([17]) binary textural features. Eighteen tree species in 90 images are classified using the k-NN classifier. The support vector machine classifier and multiscale rotationally invariant LBP features are used in [18]. The multi-class classification problem is solved using the one versus all scheme. The method is verified on two general texture datasets and the AFF bark dataset [6]. A comparison of the usefulness of the run-length method (5 features), co-occurrence correlation method (100) features for the bark k-NN classification into nine categories with 15 samples per category is presented in [21]. The method [6] uses support vector machine classifier with radial basis function kernel applied with four (contrast, correlation, homogeneity, and energy) gray-level co-occurrence matrices (GLCM), SIFT based bag-of-words, and wavelet features. The bark dataset (AFF bark dataset) consists of 1183 images of the eleven most common Austrian trees (Section 4). Color descriptor based on three-dimensional adaptive sum and difference histograms was applied on BarTex textures in [15,16].

Authors in [4] collected by far the largest public bark dataset BarkNet 1.0. The BarkNet 1.0 dataset contains more than 23 000 high-resolution bark images from 23 different Canadian tree species. The color bark images are classified using the convolutional neural network ResNet pre-trained on ImageNet.

The majority of the published methods suffer from neglecting spectral information and using discriminative and thus approximate textural features only. Few attempts to use multispectral information [9,10,12,14,21] independently apply monospectral features on each spectral band or apply the color LBP features [8,13] or convolutional neural network [4]. Most methods use private and very restricted bark databases, thus the published results are mutually incomparable and of limited information value.

The rotationally invariant bark recognition method [14] is based on novel 2D spiral Markovian texture features computed for each spectral band, then approximated by the multivariate Gaussian distribution. The 1-NN classifier uses the symmetrized Kullback-Leibler divergence, i.e., Jeffrey's divergence as the distance measure. The presented approach generalizes these monospectral 2D Markovian spiral textural features into fully multispectral 3D spiral textural features. In the 2DSCAR model article [14] we already introduced matrix notation which can be adapted with tiny modifications also for the presented 3DSCAR model-based textural features.

2. Spiral Markovian texture representation

The adaptive spiral 3D causal auto-regressive random (3DSCAR) field model is a generalization of the 3DCAR model [7]. The model's functional contextual neighbor index shift set is denoted I_r^{cs}. The model can be defined in the following matrix equation:

$$Y_r = \gamma Z_r + e_r, \tag{1}$$

where $\gamma = [A_1, \ldots, A_\eta]$ is the parameter matrix, $\eta = cardinality(I_r^{cs})$, $r = [r_1, r_2]$ is spatial index denoting history of movements on the lattice I, e_r denotes the driving white Gaussian noise vector with zero mean and a constant but unknown covariance matrix Σ, and Z_r is a neighborhood support vector of multispectral pixels Y_{r-s} where $s \in I_r^{cs}$.

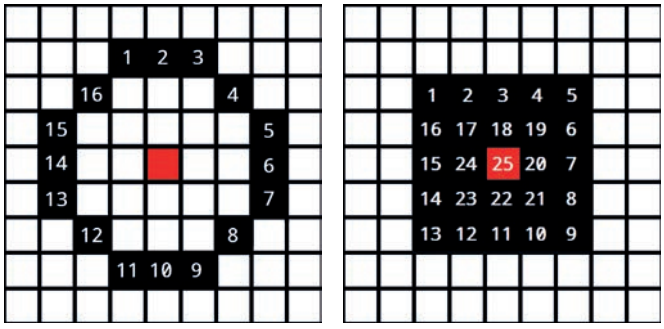

Fig. 1. The paths of the two "spirals" in an image. Left: octagonal, right: rectangular. The numbers designate the order in which the pixels r, i.e., I_r^{cs} neighborhoods are traversed and the red square means the center pixel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

All 3DSCAR model statistics can be efficiently estimated analytically [7]. The Bayesian parameter estimation (conditional mean value) $\hat{\gamma}$ can be accomplished using fast, numerically robust and recursive statistics [7], given the known 3DSCAR process history $Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \ldots, Y_1, Z_t, Z_{t-1}, \ldots, Z_1\}$:

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

$$V_{t-1} = \tilde{V}_{t-1} + V_0, \tag{3}$$

$$\tilde{V}_{t-1} = \begin{pmatrix} \sum_{u=1}^{t-1} Y_u Y_u^T & \sum_{u=1}^{t-1} Y_u Z_u^T \\ \sum_{u=1}^{t-1} Z_u Y_u^T & \sum_{u=1}^{t-1} Z_u Z_u^T \end{pmatrix} = \begin{pmatrix} \tilde{V}_{yy(t-1)} & \tilde{V}_{zy(t-1)}^T \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{zz(t-1)} \end{pmatrix}, \tag{4}$$

where V_0 is a positive definite initialization matrix (see [7]). To simplify notation, we introduce a new traversing order multi-index t of the sequence of multi-indices r which is based on the selected model movement in the lattice I (e.g., $t_{16} = \{t_{16} + (1; -1), t_{16} + (2; -1), \ldots, t_{16} + (-1; 1)\}$ for Fig. 1-left). The optimal causal functional contextual neighbourhood I_r^{cs} can be solved analytically by a straightforward generalization of the Bayesian estimate in [7]. To simplify our experiments we did not optimize the neighbourhood I_r^{cs} but used its fixed form Fig. 2. The model can be easily applied also to numerous synthesis applications. The 3DSCAR model pixel-wise synthesis is simple direct application of (1) for any 3DSCAR model.

2.1. Spiral models

The 3DSCAR model's movement r on the lattice I takes the form of circular or spiral-like paths as seen in Fig. 1. The causal neighborhood I_r^c has to be transformed to be consistent for each direction in the traversed path to. The paths used can be arbitrary as

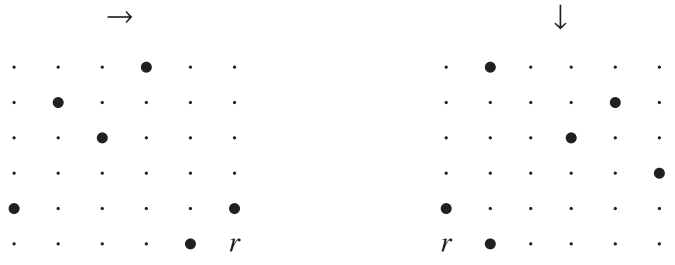

Fig. 2. The fixed causal functional contextual neighbourhood I_r^{cs}. Left: rightwards direction, right: downwards direction.

long as they keep transforming the causal neighborhood into I_r^{cs} in such a way that all neighbors of a control pixel r have been visited by the model in the previous steps. We shall call all these paths as spirals further on. We present two types of paths - octagonal (Fig. 1 on the left) and a rectangular spiral (Fig. 1 - right). During our experiments, they exhibited comparable results with the octagonal path being faster thanks to its consisting of fewer pixels for the same radius.

After the whole path is traversed, the parameters for the center pixel (shown as red square in Fig. 1) of the spiral are estimated. Contrary to the standard CAR model [7], since this model's equations do not need the whole history of movement through the image but only the given one spiral, the 3DSCAR models can be easily parallelized. If the spiral paths used have a circular shape, the 3DSCAR models exhibit rotational invariant properties thanks to the CAR model's memory of all the visited pixels. The spiral neighborhood I_r^{cs} (Fig. 1-right) is rotationally invariant only approximately. Additional contextual information can be easily incorporated if every initialization matrix $V_0 = V_{t-1}$, i.e., if this matrix is initialized from the previous data gathering matrix.

2.2. Rotationally invariant multispectral features

For feature extraction, we analyzed the 3DSCAR model around pixels with the vertical and horizontal stride of 2 to speed up the computation. The following illumination invariant features initially derived for the 3DCAR model [7] were adapted for the 3DSCAR:

$$\alpha_1 = 1 + Z_r^T V_{zz}^{-1} Z_r, \tag{5}$$

$$\alpha_2 = \sqrt{\sum_r \left(Y_r - \hat{\gamma} Z_r\right)^T \lambda_r^{-1} \left(Y_r - \hat{\gamma} Z_r\right)}, \tag{6}$$

$$\alpha_3 = \sqrt{\sum_r (Y_r - \mu)^T \lambda_r^{-1} (Y_r - \mu)}, \tag{7}$$

where μ is the mean value of vector Y_r and

$$\lambda_{t-1} = V_{yy(t-1)} - V_{zy(t-1)}^T V_{zz(t-1)}^{-1}. \tag{8}$$

The inversion matrix $V_{zz(t-1)}^{-1}$ is updated in its Cholesky factor to guarantee numerical stability of all computed model statistics [7]. As the texture features, we also used the estimated trace of γ parameters, the posterior probability density [7]

$$p(Y_r \,|\, Y^{(r-1)}, \hat{\gamma}_{r-1}) = \frac{\Gamma\left(\frac{\beta(r)-\eta+3}{2}\right)}{\Gamma\left(\frac{\beta(r)-\eta+2}{2}\right) \pi^{\frac{1}{2}} \left(1 + X_r^T V_{x(r-1)}^{-1} X_r\right)^{\frac{1}{2}} |\lambda_{(r-1)}|^{\frac{1}{2}}} \times \left(1 + \frac{(Y_r - \hat{\gamma}_{r-1} X_r)^T \lambda_{(r-1)}^{-1} (Y_r - \hat{\gamma}_{r-1} X_r)}{1 + X_r^T V_{x(r-1)}^{-1} X_r}\right)^{-\frac{\beta(r)-\eta+3}{2}}, \tag{9}$$

$\beta(r) = r + \eta - 2$, and the absolute error of the one-step-ahead prediction

$$Abs(GE) = \left| E\left\{ Y_r \,|\, Y^{(r-1)} \right\} - Y_r \right| = |Y_r - \hat{\gamma}_{r-1} X_r|. \tag{10}$$

3. Bark classifier

We first subsample the images to the height of 300px (if the image is larger), keeping aspect ratio to speed up the feature extraction part. This subsampling ratio depends on an application data, i.e., a compromise between the algorithm efficiency and its recognition rate. The features are extracted as described in Section 2. The feature space is assumed to be approximated by the multivariate Gaussian distribution, the parameters of which are then stored for each training sample image.

$$\mathcal{N}(\theta|\mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} e^{\left(-\frac{1}{2}(\theta-\mu)^T \Sigma^{-1} (\theta-\mu)\right)}. \tag{11}$$

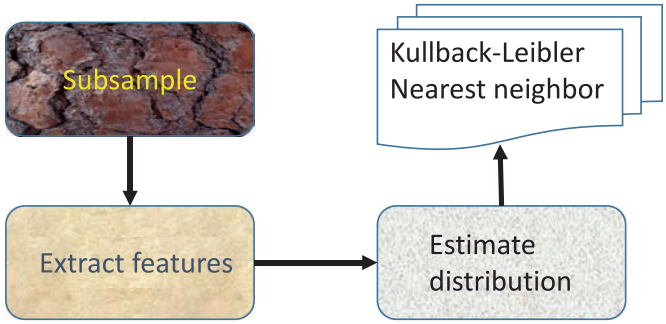

Fig. 3. Flowchart of our classification approach.

During the classification stage, the parameters of the Gaussian distribution are estimated for the classified image as in the training step (the flowchart of our approach can be seen in Fig. 3). They are then compared with all the distributions of the training samples using the Kullback-Leibler (KL) divergence. The KL divergence is a measure of how much one probability distribution diverges from another. It is defined as:

$$D(f(x)||g(x)) \stackrel{def}{=} \int f(x) \log \frac{f(x)}{g(x)} \, dx. \tag{12}$$

For the Gaussian distribution data model, the KL divergence can be solved analytically:

$$D(f(x)||g(x)) = \frac{1}{2}\left(\log \frac{|\Sigma_g|}{|\Sigma_f|} + tr(\Sigma_g^{-1}\Sigma_f) - d\right) + \frac{1}{2}(\mu_f - \mu_g)^T \Sigma_g^{-1} (\mu_f - \mu_g). \tag{13}$$

We use the symmetrized variant of the Kullback-Leibler divergence known as the Jeffreys divergence

$$D_s(f(x)||g(x)) = \frac{D(f(x)||g(x)) + D(g(x)||f(x))}{2}. \tag{14}$$

The class of the training sample with the lowest divergence from the image being recognized is then selected as the final result. The advantage of our approach is that the training database is heavily compressed through the Gaussian distribution parameters (as we extract only about 40 features, depending on the chosen neighborhood, we just need to store 40 numbers for the mean and 40×40 numbers for the covariance matrix) and the comparison with the training database is extremely fast, enabling us to compare hundreds of thousands of image feature distributions per second on an ordinary computer.

4. Experimental data

The proposed method is verified on four publicly available bark databases. Examples of images of the datasets can be seen in Fig. 4. We have used the leave-one-out approach for the classification rate estimation. Thus the number of training images varies between 6 for AFF to 2723 for the largest BarkNet class Quercus rubra. V_0 is initialized to be the identity matrix. All bark images in our experiments were extremely resized to our disadvantage to have a comparable amount of information with the [4] method which uses 224×224 textures only.

4.1. AFF dataset

The AFF bark dataset provided by Österreichische Bundesforste, Austrian Federal Forests (AFF) [6], is a collection of the most common Austrian trees. The dataset contains 1182 bark samples (960×1325) belonging to 11 classes, the size of each class varying between 7 and 213 images. AFF samples are captured at different scales, and under different illumination conditions.

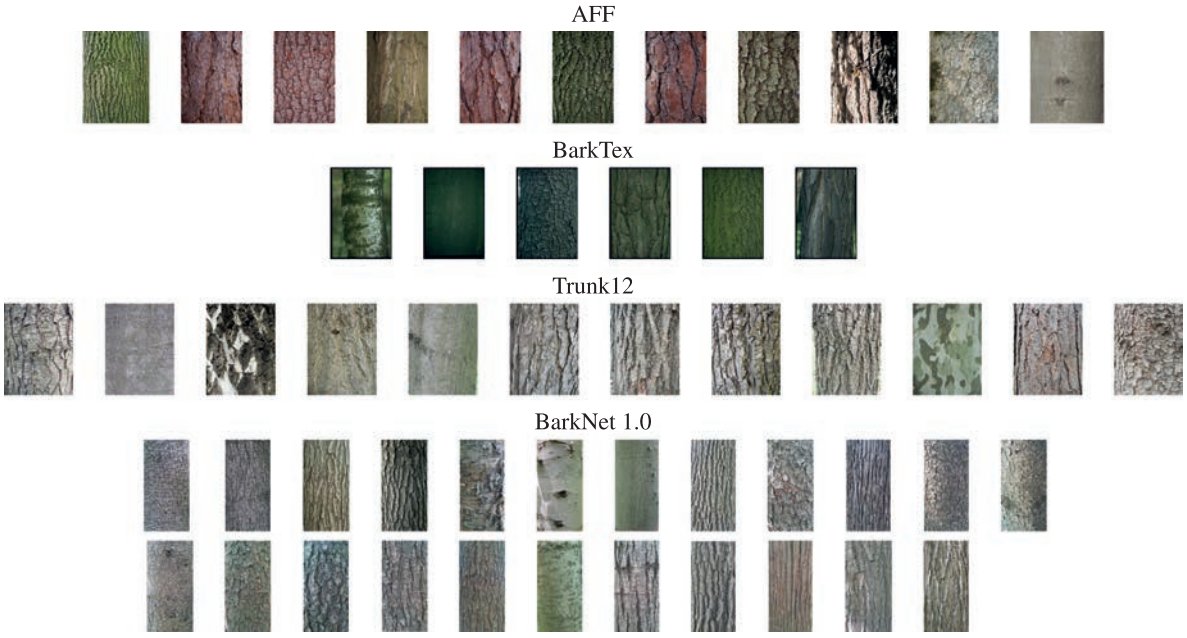

Fig. 4. Examples of images from the individual datasets. Top to bottom (rightwards): AFF (ash, black pine, fir, hornbeam, larch, mountain oak, Scots pine, spruce, Swiss stone pine, sycamore maple, beech), BarkTex (betula pendula, fagus silvatica,picea abies, pinus silvestris, quercus robur, robinia pseudacacia), Trunk12 (alder, beech, birch, ginkgo biloba, hornbeam, horse chestnut, chestnut, linden, oak, oriental plane, pine, spruce), BarkNet (Betula alleghaniensis, Betula papyrifera, Quercus rubra, Picea glauca, Picea mariana, Picea abies, Picea rubens, Acer platanoides, Acer rubrum, Acer saccharum, Fraxinus americana, Fagus grandifolia, Larix laricina, Ulmus americana, Ostrya virginiana, Populus grandidentata, Populus tremuloides, Pinus strobus, Pinus rigida, Pinus resinosa, Tsuga canadensis, Abies balsamea, Thuja occidentalis).

4.2. Trunk12 dataset

The Trunk12 dataset (3000×4000, [19], http://www.vicos.si/Downloads/TRUNK12) contains 393 images of tree barks belonging to 12 different trees that are found in Slovenia. The number of images per class varies between 30 and 45 images. Bark images are captured under controlled scale, illumination and pose conditions. The classes are more homogeneous than those of AFF regarding imaging conditions.

4.3. BarkTex dataset

The BarkTex dataset [11] contains 408 samples from 6 bark classes, i.e., 68 images per class. The images have small (256×384) resolution and they have unequal natural illumination and scale.

4.4. BarkNet 1.0 dataset

The BarkNet 1.0 dataset [4] contains 23 bark classes (Betula alleghaniensis, Betula papyrifera, Quercus rubra, Picea glauca, Picea mariana, Picea abies, Picea rubens, Acer platanoides, Acer rubrum, Acer saccharum, Fraxinus americana, Fagus grandifolia, Larix laricina, Ulmus americana, Ostrya virginiana, Populus grandidentata, Populus tremuloides, Pinus strobus, Pinus rigida, Pinus resinosa, Tsuga canadensis, Abies balsamea, Thuja occidentalis) recorder by three cellphones Nexus 5, Samsung Galaxy S5, Samsung Galaxy S7 and camera Panasonic Lumix DMC-TS5 in the vicinity of Quebec, Canada. This dataset contains over 23,000 variable high-resolution bark images each one together with the tree diameter at breast height information. Images were captured from variable distances between 20–60 cm away from the trunk without perspective distortion and under alternating weather conditions which ranged from sunny to light rain. The number of images per class differs from 64 (Populus grandidentata) to 2724 (Quercus rubra).

5. Results

5.1. 2DSCAR results

We have improved (83.6% [14]) the accuracy to 89.1% on the AFF dataset (Table 1), and achieved 91.7% on the BarkTex database and 92.9% on the Trunk12 dataset (Table 3). In all the three tables, the name of the row indicates the actual tree type whereas the column indicates the predicted class. The comparison with other methods is presented in Table 4. We can see that our approach vastly outperforms all compared methods on the BarkTex and Trunk12 datasets and has the second best results on the AFF dataset. However, the result of the method [18] on the AFF dataset was achieved on quadruple textures. Thus this comparison is biased and comparable results would be probably worse for their method. The BarkNet 1.0 dataset accuracy result is 90.4% on all 23 classes. The sensitivity for all classes is between $84 - 99$ [%] with median value 90% and precision $77 - 100$ [%] with median value 90%, which are better results than the ResNet classifier achievement.

5.2. 3DSCAR results

The 3DSCAR model was applied to the BarkNet 1.0 dataset and compared with results obtained using convolutional neural network ResNet pre-trained on ImageNet [4]. The neural net requires huge training set, contrary to our 3DSCAR model, which can be reliably learned on a single image. Thus the ResNet did not classify three classes with the smallest number of images (Populus grandidentata - 64, Acer platanoides - 70, Pinus rigida - 123). This is not a problem for our method as can be seen in Table 5. The sensitivity for all 23 classes is between $72 - 96$ [%] with median value 86% and precision $68 - 97$ [%] with median value 85%. The ResNet classifier is highly dependent on the number of training images. If the number of training images is in the range $\langle 476; 2179 \rangle$ per class they reach the accuracy 85% but for smaller training set ($\langle 52; 242 \rangle$) the accuracy significantly decreases to 64%. The accuracy re-

Table 1
AFF bark dataset results of the 2DSCAR method (MO - Mountain oak, SP - Scots pine, SSP - Swiss stone pine, SM - Sycamore maple).

	Ash	Beech	Black pine	Fir	Horn-beam	Larch	MO	SP	Spruce	SSP	SM	Sensitivity %
Ash	23	0	0	0	0	0	0	0	0	0	0	95,8
Beech	0	6	0	0	0	0	0	0	0	0	0	85,7
B. pine	0	0	149	1	0	8	0	8	0	1	0	94,9
Fir	0	0	0	101	0	1	0	0	4	0	0	85,6
Horn.	0	0	0	0	32	0	0	0	0	0	1	97,0
Larch	0	0	4	10	0	163	0	30	2	0	0	85,3
MO	0	0	0	0	0	0	65	0	0	4	3	95,6
SP	0	0	4	3	0	19	0	143	3	0	0	79,0
Spruce	0	0	0	3	0	0	1	0	195	0	0	95,6
SSP	1	0	0	0	0	0	2	0	0	82	3	94,3
SM	0	1	0	0	1	0	0	0	0	0	5	41,7
Precision %	100	100	89,2	95,3	97	78	90,3	83,1	98	93,2	71,4	**Accuracy 89,1**

Table 2
BarkTex dataset results of the 3DSCAR method (BP - Betula pendula, FS - Fagus silvatica, PA - Picea abies, PS - Pinus silvestris, QR - Quercus robur, RP - Robinia pseudacacia).

	BP	FS	PA	PS	QR	RP	Sensitivity %
Betula pendula	55	1	4	1	4	0	80,9
Fagus silvatica	2	67	0	0	0	0	98,5
Picea abies	6	0	60	0	4	0	88,2
Pinus silvestris	1	0	0	64	8	2	94,1
Quercus robur	4	0	4	2	50	3	73,5
Robinia pseudacacia	0	0	0	1	2	63	92,7
Precision %	84,6	97,1	85,7	85,3	79,4	95,5	**Accuracy 88,0**

Table 3
Trunk12 dataset results of the 2DSCAR method (A - Alder, Be - Beech, Bi - Birch, Ch - Chestnut, GB - Ginkgo biloba, H - Hornbeam, HC - Horse chestnut, L - Linden, OP - Oriental plane, S - Spruce).

	A	Be	Bi	Ch	GB	H	HC	L	Oak	OP	Pine	S	Sensitivity %
Alder	33	0	1	0	0	0	0	0	0	0	0	0	97,1
Beech	0	29	0	0	0	1	0	0	0	0	0	0	96,7
Birch	0	0	36	1	0	0	0	0	0	0	0	0	97,3
Chestnut	2	0	0	24	0	0	0	0	4	0	2	0	75,0
Ginkgo biloba	0	0	0	0	30	0	0	0	0	0	0	0	100
Hornbeam	0	2	0	0	0	28	0	0	0	0	0	0	93,3
Horse chestnut	0	0	1	0	0	1	27	3	0	0	1	0	81,8
Linden	0	0	0	1	0	0	4	25	0	0	0	0	83,3
Oak	1	0	0	0	0	0	0	0	29	0	0	0	96,7
Oriental plane	0	0	0	1	0	0	1	0	0	30	0	0	93,8
Pine	0	0	0	0	0	0	0	0	0	0	30	0	100
Spruce	1	0	0	0	0	0	0	0	0	0	0	44	97,8
Precision %	89,2	93,5	94,7	88,9	100	93,3	84,4	89,3	87,9	100	90,9	100	**Accuracy 92,9**

Table 4
Accuracy comparison with the state-of-the-art. '-' denotes lack of results in the particular article on the given dataset.

Dataset %	2DSCAR [14]	3DSCAR	[3]	[6]	[18]	[8]	[12]	[13]	[16]	[15]	[4]
AFF	89.1	88.0	60.5	69.7	**96.5**	-	-	-	-	-	-
BarkTex	**91.7**	88.0	84.6	-	-	81.4	84.7	81.4	82.1	89.6	-
Trunk12	**92.9**	81.9	62.8	-	-	-	-	-	-	-	-
BarkNet	**90.4**	85.0	-	-	-	-	-	-	-	-	(64) 85

sult for the AFF dataset is 88%, BarkTex database (Table 2) is 88%, and the Trunk12 dataset is 81.9%, respectively. As expected, the 3DSCAR feature results are worse than their 2DSCAR alternative, due to not robust estimation of the spectral correlation because the learning resized images are purposely too small. Nevertheless, even this insufficient learning data allow to achieve better accuracy (Table 4 – 85%) than the convolutional neural network ResNet (64%).

The computational load of the alternative methods in Table 4 cannot be easily compared due to missing information in the corresponding papers. We can guess at this point that the learning step of the methods based on neural nets [4,10] is significantly more demanding than ours, while the classification stage will be comparable. Our unoptimized implementation is divided into three steps - feature extraction and their subsequent storage for every image, parametric Gaussian model estimation,

Table 5
Presented method (3DSCAR) classification results for three bark species which have too limited training sets in BarkNet for the alternative ResNet classifier.

tree	precision [%]		sensitivity [%]	
	2DSCAR	3DSCAR	2DSCAR	3DSCAR
Populus grandidentata	92	87	94	91
Acer platanoides	100	96	83	76
Pinus rigida	77	68	83	76

and classification. The first step is the most computationally demanding taking seconds per image due to the unneeded storage of many large parametric files instead of directly computing the Gaussian models; the two subsequent steps are negligible in comparison with the feature extraction step.

6. Conclusion

We present novel rotationally invariant fully multispectral Markovian textural features and apply them to the tree bark recognition. These statistical features are analytically derived from the underlying descriptive textural model and can be efficiently, recursively, and adaptively learned. The classifiers based on these features outperform the state-of-the-art alternative methods on four public bark databases. The 2DSCAR model-based classifier is probably the second best on the AFF database even in its combined monospectral form, which neglects mutual spectral correlations between spectral bands. Our classifiers outperform convolutional neural networks ResNet even on the by far largest public bark database BarkNet which contains 23 000 high-resolution images from 23 different tree species. Our 2DSCAR/3DSCAR methods are rotationally invariant, benefit from information from all spectral bands and can be easily parallelized or made fully illumination invariant. Both our classifiers are fully analytical and much faster than the convolutional neural net alternative, especially in their learning stage. They do not need large learning data, and they outperform the convolutional neural nets even if these learning data are restricted. The choice between 2DSCAR and 3DSCAR based features depends on the number of available learning data. 3DSCAR features will benefit from larger learning set, while for limited learning data available, the 2DSCAR features computed separably from each spectral band is the recommendable option. We have also executed our method without any modification on the AFF dataset's images of needles and leaves, with results exceeding 94% accuracy. This will be a subject of our further research.

Conflict of interest

None.

Acknowledgments

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

References

[1] L.J. Blaanco, C.M. Travieso, J.M. Quinteiro, P.V. Hernandez, M.K. Dutta, A. Singh, A bark recognition algorithm for plant classification using a least square support vector machine, in: 2016 Ninth International Conference on Contemporary Computing (IC3), 2016, pp. 1–5, doi:10.1109/IC3.2016.7880233.
[2] S. Boudra, I. Yahiaoui, A. Behloul, A comparison of multi-scale local binary pattern variants for bark image retrieval, in: International Conference on Advanced Concepts for Intelligent Vision Systems, Springer, 2015, pp. 764–775.
[3] S. Boudra, I. Yahiaoui, A. Behloul, Statistical radial binary patterns (srbp) for bark texture identification, in: J. Blanc-Talon, R. Penne, W. Philips, D. Popescu, P. Scheunders (Eds.), Advanced Concepts for Intelligent Vision Systems, Springer International Publishing, Cham, 2017, pp. 101–113.
[4] M. Carpentier, P. Giguere, J. Gaudreault, Tree species identification from bark images using convolutional neural networks, 2018, arXiv:1803.00949.
[5] Z. Chi, L. Houqiang, W. Chao, Plant species recognition based on bark patterns using novel gabor filter banks, in: International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003, 2003, pp. 1035–1038Vol.2, doi:10.1109/ICNNSP.2003.1281045.
[6] S. Fiel, R. Sablatnig, Automated identification of tree species from images of the bark, leaves and needles, in: 16th Computer Vision Winter Workshop, Verlag der Technischen Universität Graz, 2011, pp. 67–74.
[7] M. Haindl, Visual data recognition and modeling based on local markovian models, in: L. Florack, R. Duits, G. Jongbloed, M.C. Lieshout, L. Davies (Eds.), Mathematical Methods for Signal and Image Analysis and Representation, volume 41 of Computational Imaging and Vision, chapter 14, Springer London, 2012, pp. 241–259, doi:10.1007/978-1-4471-2353-8_14.
[8] V.T. Hoang, A. Porebski, N. Vandenbroucke, D. Hamad, Lbp histogram selection based on sparse representation for color texture classification, in: VISIGRAPP (4: VISAPP), 2017, pp. 476–483.
[9] Z.K. Huang, Bark classification using rbpnn based on both color and texture feature, Int. J. Comput. Sci. Netw. Secur. 6 (2006) 100–103.
[10] Z.k. Huang, D.s. Huang, M.R. Lyu, T.m. Lok, Classification based on gabor filter using rbpnn classification, in: Computational Intelligence and Security, 2006 International Conference on, IEEE, 2006, pp. 759–762.
[11] R. Lakmann, Statistische Modellierung von Farbtexturen, 1998 Ph.d. thesis.. ftp://ftphost.uni-koblenz.de/de/ftp/pub/outgoing/vision/Lakman/BarkTex/.
[12] C. Palm, Color texture classification by integrative co-occurrence matrices, Pattern Recognit. 37 (2004) 965–976.
[13] A. Porebski, N. Vandenbroucke, D. Hamad, Lbp histogram selection for supervised color texture classification, in: ICIP, 2013, pp. 3239–3243.
[14] V. Remeš, M. Haindl, Rotationally invariant bark recognition, in: IAPR Joint International Workshop on Statistical Techniques in Pattern Recognition and Structural and Syntactic Pattern Recognition (S+SSPR 2018), Springer Nature Switzerland AG, 2018, pp. 22–31. http://ssspr2018.buaa.edu.cn. doi: 10.1007/978-3-319-97785-0_3.
[15] F. Sandid, A. Douik, Dominant and minor sum and difference histograms for texture description, in: 2016 International Image Processing, Applications and Systems (IPAS), 2016, pp. 1–5, doi:10.1109/IPAS.2016.7880136.
[16] F. Sandid, A. Douik, Robust color texture descriptor for material recognition, Pattern Recognit. Lett. 80 (2016) 15–23, doi:10.1016/j.patrec.2016.05.010.
[17] J. Song, Z. Chi, J. Liu, H. Fu, Bark classification by combining grayscale and binary texture features, in: Intelligent Multimedia, Video and Speech Processing, 2004. Proceedings of 2004 International Symposium on, IEEE, 2004, pp. 450–453.
[18] M. Sulc, J. Matas, Kernel-mapped histograms of multi-scale lbps for tree bark recognition, in: Image and Vision Computing New Zealand (IVCNZ), 2013 28th International Conference of, IEEE, 2013, pp. 82–87.
[19] M. Švab, Computer-vision-based tree trunk recognition, 2014 Ph.d. thesis. Ljubljana, Slovenia Faculty of Computer and Information Science, University of Ljubljana.
[20] J. Wäldchen, P. Mäder, Plant species identification using computer vision techniques: a systematic literature review, Arch. Comput. Method Eng. 25 (2018) 507–543, doi:10.1007/s11831-016-9206-z.
[21] Y.Y. Wan, J.X. Du, D.S. Huang, Z. Chi, Y.M. Cheung, X.F. Wang, G.J. Zhang, Bark texture feature extraction based on statistical texture analysis, in: Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004., 2004, pp. 482–485, doi:10.1109/ISIMP.2004.1434106.