Coniferous Trees Needles-Based Taxonomy Classification

Michal Haindl & Pavel Žid
Pattern Recognition Dep.
The Institute of Information Theory and Automation
of the Czech Academy of Sciences
Prague, Czechia
{haindl,zid}@utia.cas.cz

***Abstract*—This paper introduces multispectral rotationally invariant textural features of the Markovian type applied for the effective coniferous tree needles categorization. Presented texture features are inferred from the descriptive multispectral spiral wide-sense Markov model. Unlike the alternative texture recognition methods based on various gray-scale discriminative textural descriptions, we take advantage of the needles texture representation, which is fully descriptive multispectral and rotationally invariant.**

The presented method achieves high accuracy for needles recognition. Thus it can be used for reliable coniferous tree taxon classification. Our classifier is tested on the open source needles database Aff, which contains 716 high-resolution images from 11 diverse coniferous tree species.

***Index Terms*—Coniferous needles categorization, Tree taxonomy recognition, Spiral Markov random field model**

I. Introduction

Automatic tree taxon classification is a challenging but practical plant taxonomy application which permits non-invasive quick tree categorization. Automatic needles categorization allows identification or learning of coniferous tree species possible without specific botanical knowledge using, for example, a dedicated mobile application. Tree's species identification based on manual listing in a botanical key of needles images is a tedious task which would typically requires a book scrolling. Even if needles can be more easily described than for example bark texture [1], [2], the observer still has to search through the whole needles printed or electronic encyclopedia looking for the corresponding needles image. An automatic recognition application can even overcome human scholars. The authors [1] reports 56.6% and 77.8% classification accuracy on the Austrian Federal Forest (AFF) dataset for two human specialits, while two published automatic classifiers [3] and [4] achieve more accurate results than the top human specialist.

The benefit of needles-based features is their relative stability during the corresponding tree's lifetime. Single coniferous shrubs or trees have distinct needles which can be advantageously used for their identification. Numerous practical applications might be ecological such as plant resource management or fast identification of invading tree species. Industrial applications can be in sawmills automatization or detection of bark beetle tree infestation.

Unfortunately, the only needles database we are aware of is AFF [1] as well as the only partially reported results on this database. The authors [1] first segmented the image to find the needles endings and then calculated the skeleton of the needles. From the needles endings, they inferred features like the eccentricity, solidity, the moment invariants, and curvature features. However, they did not find any solution to identify the tree species from images of the needles where the needles grow in clusters. Thus they could classify only five fir images with precision 100 % and nine spruce images with precision 77 %.

II. Spiral Markovian texture representation

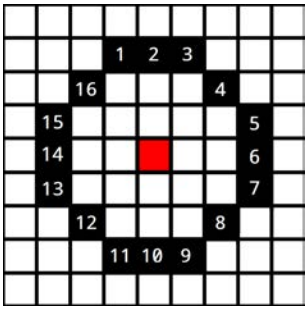

Fig. 1. The octagonal spirals path. The numbers mark the order in which the pixels r, i.e., I_r^{cs} neighborhoods are traversed, and the center pixel is marked as the red square.

The spiral 2D causal auto-regressive adaptive random (2DSCAR) field model is a generalization of our directional 2DCAR model [5] to the rotationally invariant form, which was introduced in [4] for bark classification application. The model's contextual neighbor index shift set (Fig. 2) denoted I_r^{cs} is functional. The model for d spectral bands can be defined in the following matrix form:

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

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

978-1-7281-4187-9/19/$31.0©2019 IEEE

where $\gamma = [A_1, \ldots, A_\eta]$ is the parameter matrix, $A_i = diag[a_{i1}, \ldots, a_{id}]$ $\forall i$, $\eta = cardinality(I_r^{cs})$, $r = [r_1, r_2]$ is spatial multi-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 $s \in I_r^{cs}$.

All 2DSCAR model statistics can be efficiently estimated analytically as proven in [5]. The Bayesian parameter estimation (conditional mean value) $\hat{\gamma}$ can be accomplished using fast, numerically robust and recursive statistics [5], given the known 2DSCAR 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 [5]). We introduce a new traversing order multi-index t of the sequence of multi-indices r, to simplify notation, which depends on the selected model movement in the underlying lattice I (e.g., $t_{16} = \{t_{16} + (1; -1), t_{16} + (2; -1), \ldots, t_{16} + (-1; 1)\}$ for Fig. 1). The optimal functional causal contextual neighbourhood I_r^{cs} (Fig. 2) can be solved analytically by a straightforward generalization of the Bayesian estimate derived in [5]. We did not optimize the neighbourhood I_r^{cs} but used its fixed form Fig. 2 to simplify and speed up our experiments. However, if this neighborhood is optimized, we can expect further accuracy improvement. The model can be easily applied also to various synthesis and restoration applications. The 2DSCAR model pixel-wise synthesis is a direct application of the equation (1) fed from a Gaussian noise generator for any 2DSCAR model.

A. Spiral models

The 2DSCAR model moves (r) on the circular path on the lattice I as is illustrated in Fig. 1. The causal neighborhood I_r^c has to be transformed to be consistent for each direction in the traversed path, as denoted in Fig. 2. The paths used can be arbitrary as long as they keep transforming the causal neighborhood into I_r^{cs} in such a way that the model has visited all neighbors of a control pixel r. Thus these neighbors are known from the previous steps. We shall call all these causal paths as spirals further on. In this paper we present the octagonal type of path - (Fig. 1), however, alternatively, a spiral path can be used as well. The parameters for the center pixel (the red square in Fig. 1) of the spiral are estimated after the whole path is completed. Since this model's equations do not need the whole history of movement through the image but only the local neighborhood of a single spiral, the 2DSCAR models can be easily parallelized. This memory restriction is advantageous in comparison to the standard directional CAR models [5]. The 2DSCAR models exhibit rotational invariant properties for the circular shape spiral paths, thanks to the CAR model's memory of all the visited pixels. Additional prior contextual information can be easily incorporated if every initialization matrix $V_0 = V_{t-1}$, for example, this matrix can be initialized from the previous data gathering matrix.

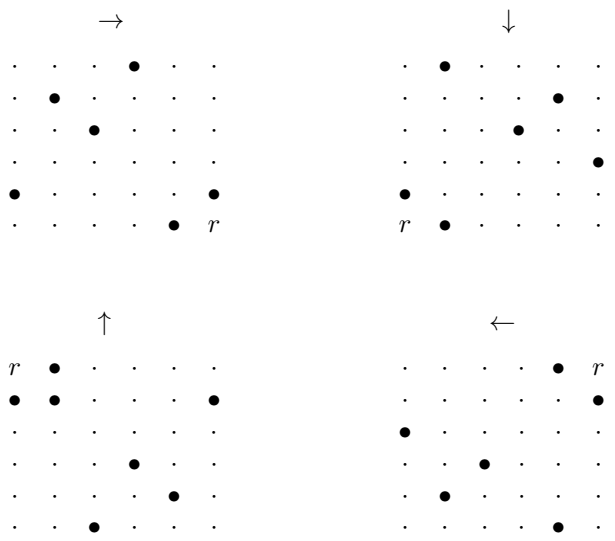

Fig. 2. The applied fixed causal functional contextual neighbourhood I_r^{cs} in four directions. Upper left: rightwards, upper right: downwards, bottom left upwards, bottom right leftwards direction, respectively.

B. Multispectral rotationally invariant features

We analyzed the 2DSCAR model around all pixels with the vertical and horizontal stride of 2 to speed up the computation for feature extraction. The following $\alpha_1, \alpha_2, \alpha_3$ illumination invariant features initially derived for the 3DCAR model [5] were adapted for the 2DSCAR model:

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

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

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

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 data gathering matrix $V_{zz(t-1)}^{-1}$ is updated in its square-root Cholesky factor to guarantee numerical stability for computed model statistics [5]. Additional used texture features are also the estimated trace of γ parameters, the posterior probability density [5]

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

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

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

III. Needles classifier

The algorithm starts with image subsampling to the height of 600px (for larger images), while keeping aspect ratio, to speed up the feature extraction part. This subsampling ratio depends on application data; it is a compromise between the algorithm efficiency and its recognition rate. Every pixel has extracted features, as described in Sec. II. The resulting feature space indexed on the lattice I is assumed to be govern by the multivariate Gaussian distribution. The estimated Gaussian parameters then represent every training image sample:

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

In the classification step, the Gaussian distribution parameters are estimated for the classified image in the same way. The classified image parameters are then compared with all the distributions from the training samples set using the Kullback-Leibler (KL) divergence. The KL divergence is a probability distribution non-symmetric similarity measure between two distributions; it is defined as:

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

The KL divergence for the Gaussian distribution data model can be solved analytically:

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

We use the Jeffreys divergence which the symmetrized variant of the Kullback-Leibler divergence:

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

The selected class is the class of a training image with the lowest Jeffreys divergence from the tested image. The primary benefit of our method is the significant compression of the training database into the Gaussian distribution parameters (as we extract only about 40 features, depending on the chosen neighborhood, we need to store 40 real numbers for the mean and 40×40 numbers for the covariance matrix). The subsequent comparison with the training database is thus extremely fast, enabling us to compare hundreds of thousands of image feature distributions per second on an ordinary computer.

IV. Experimental AFF needles data

The proposed method is verified on the only publicly available needles database AFF [1]. The AFF bark and needles datasets provided by Österreichische Bundesforste, Austrian Federal Forests (AFF) [1], is a collection of the most common Austrian trees. The larger bark dataset contains 1182 bark samples (960×1325) belonging to 11 classes, the size of each class varying between 7 and 213 images. The separate AFF needles database contains 716 needles samples from six Austrian conifer trees, and each class is divided into two subclasses. The first subclass is trees on which one needle grows separately on the branch, and the second subclass is the trees on which the needles grow on clusters at the branch. The needles joint database contains 647 images of needles from six tree classes. AFF sample dataset is further complicated due to images captured at different scales and under different illumination conditions.

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 for the AFF database varies between 8 for fir or Scots pine and 135 for the largest spruce class. V_0 is initialized to be the identity matrix. All needles pictures in our experiments are resized to $1:5$ (600×400) for single scale experiments and the other scale was resized to $1:10$ (300×200).

V. Results

The first experiment is the classification of all separate eleven classes. Tab.I using single resolution level. We have reached 96.1% accuracy on the AFF separate dataset (Tab. I). The sensitivity for all needles classes is between $89-100$ [%] with median value 97% and precision $88-100$ [%] with median value 96%.

The second experiment is the single resolution classification on the joint needles dataset (Tab.II). Its accuracy is 96 % which is 0.1 % lower than in the first experiment. The sensitivity for all classes is between $93-98$ [%] with median value 96% and precision $91-99$ [%] with median value 95%.

The third experiment is the classification of all separate eleven classes. Tab.III using double-resolution levels. We have reached 91.1% accuracy on the AFF separate dataset, what is 1 % lower than in the single scale experiment(Tab. I). The sensitivity for all classes is between $56-100$ [%] with median value 95% and precision $72-100$ [%] with median value 96%.

Finally, the last experiment is the double-resolution classification on the joint needles dataset (Tab.IV). Its accuracy is 95.4 %, what is 0.6 % lower than in the single resolution alternative (Tab.II). The sensitivity for all classes is between

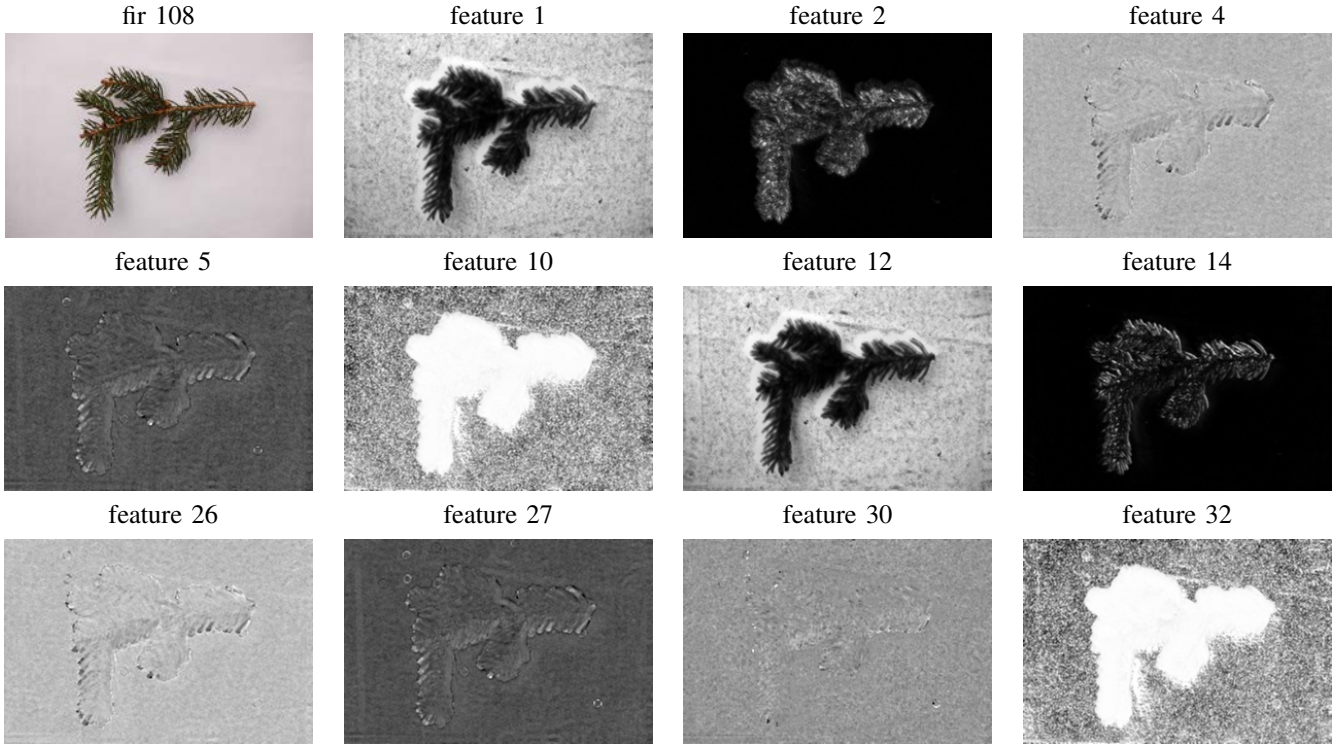

Fig. 3. Examples of the selected 2DSCAR features for fir 108 texture.

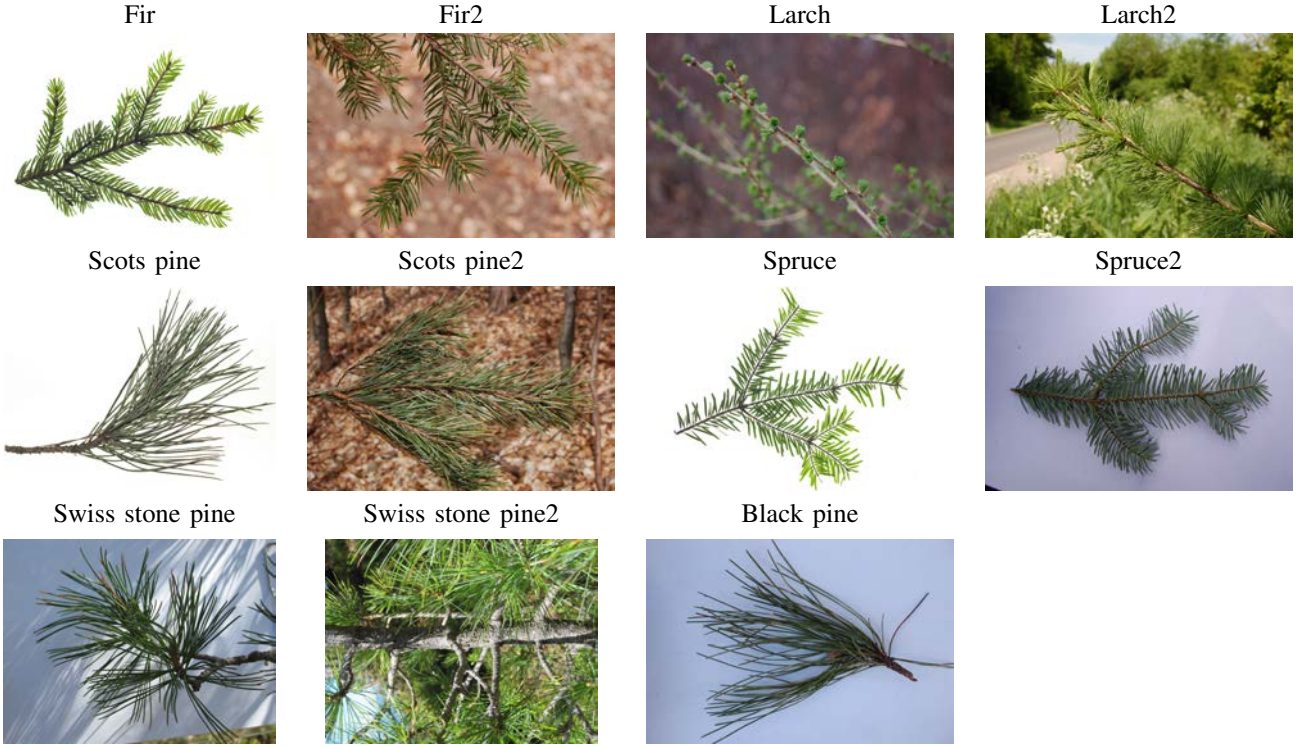

Fig. 4. Examples of needles images from the AFF dataset.

$89-100$ [%] with median value 95% and precision $92-98$ [%] with median value 96%.

In all the four tables, the name of the row indicates the actual tree type, whereas the column indicates the predicted

TABLE I
AFF NEEDLES DATASET RESULTS OF THE 2DSCAR METHOD.

	Fir	**Fir2**	**Larch**	**Larch2**	**Scots pine**	**Scots pine2**	**Spruce**	**Spruce2**	**Swiss stone pine**	**Swiss stone pine2**	**Black pine**	**Sensitivity [%]**
Fir	9	0	0	0	0	0	1	0	0	0	0	100
Fir2	0	121	1	1	0	0	0	0	0	0	0	98,4
Larch	0	0	16	0	0	0	0	0	0	0	1	88,9
Larch2	0	0	0	106	0	0	0	0	0	2	0	93,0
Scots pine	0	0	0	0	9	0	0	0	0	1	0	100
Scots pine2	0	0	0	0	0	111	0	2	0	0	3	96,5
Spruce	0	0	0	0	0	0	12	0	0	0	0	92,3
Spruce2	0	0	1	1	0	0	0	134	0	0	2	98,5
Swiss stone pine	0	0	0	0	0	0	0	0	20	0	0	100
Swiss stone pine2	0	2	0	5	0	0	0	0	0	50	0	94,3
Black pine	0	0	0	1	0	4	0	0	0	0	100	94,3
Precision [%]	90,0	98,4	94,1	98,2	99,0	95,7	100	97,1	100	87,7	95,2	**Accuracy 96,1**

class. The classification results for separate or joined databases are nearly identical. This is in sharp contrast with the reported results in [1] where the authors were not able to classify the trees on which the needles grow on clusters at the branch (the AFF second subclass). The double-resolution results have 1 % worse accuracy than their single resolution counterparts. This accuracy decrease can be explained with too high subsampling rate for such fine grain texture as the needles. We are not aware of any alternative result on needles classification even authors in [1] do not provide any numerical results. Hence we cannot compare our results with any alternative method. However, we compared the presented textural features with several alternative methods on the bark databases (BarkTex, Trunk12, AFF, BarkNet 1.0) in [6]. On these data our approach vastly outperformed all compared methods, including methods using deep convolution neural networks.

The conclusion, as mentioned above, can be expected because our features are derived from the generative random field type of model, and as such, it accumulates more information than the general discriminative textural features. Further comparative details of our generative features with their standard counterparts can be checked in our papers (Outex data [7], bidirectional texture function [8], wood mobile data [9] and some others).

VI. CONCLUSION

We present the method for coniferous tree taxon categorization based on needles classification. The classifier uses rotationally invariant monospectral Markovian textural features from all three spectral classes. Our features are analytically derived from the underlying descriptive textural model and can be efficiently, recursively, and adaptively learned. Our 2DSCAR features are rotationally invariant, exploits information from all spectral bands, and can be easily parallelized or made fully illumination invariant if the non-illumination invariant features are excluded (the posterior probability density and the absolute error of the one-step-ahead prediction). The classifier does not need extensive learning data contrary to the convolutional neural nets.

REFERENCES

[1] S. Fiel and 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.

[2] J. Wäldchen and P. Mäder, "Plant species identification using computer vision techniques: A systematic literature review," *Archives of Computational Methods in Engineering*, vol. 25, no. 2, pp. 507–543, Apr 2018. [Online]. Available: https://doi.org/10.1007/s11831-016-9206-z

[3] M. Sulc and 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.

[4] V. Remeš and 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)*, ser. Lecture Notes in Computer Science, B. X., H. E., H. T., W. R., B. B., and R.-K. A., Eds., vol. 11004. Springer Nature Switzerland AG, August 2018, pp. 22 – 31. [Online]. Available: http://ssspr2018.buaa.edu.cn

[5] M. Haindl, "Visual data recognition and modeling based on local markovian models," in *Mathematical Methods for Signal and Image Analysis and Representation*, ser. Computational Imaging and Vision, L. Florack, R. Duits, G. Jongbloed, M.-C. Lieshout, and L. Davies, Eds. Springer London, 2012, vol. 41, ch. 14, pp. 241–259, 10.1007/978-1-4471-2353-8_14. [Online]. Available: http://dx.doi.org/10.1007/978-1-4471-2353-8_14

[6] V. Remeš and M. Haindl, "Bark recognition using novel rotationally invariant multispectral textural features," *Pattern Recognition Letters*, vol. 125, pp. 612 – 617, July 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/

[7] P. Vacha and M. Haindl, "Image retrieval measures based on illumination invariant textural mrf features," in *CIVR '07: Proceedings of the 6th ACM international conference on Image and video retrieval*. New York, NY, USA: ACM Press, 2007, pp. 448–454. [Online]. Available: http://doi.acm.org/10.1145/1282280.1282346

[8] P. Vácha and M. Haindl, "Illumination invariant and rotational insensitive textural representation," in *IEEE 16th Int. Conf. on Image Processing - ICIP 2009*, M. Bayoumi, Ed. IEEE, 2009, pp. 1333–1336.

[9] M. Haindl and P. Vacha, "Wood veneer species recognition using markovian textural features," in *Computer Analysis of Images and Patterns*, ser. Lecture Notes in Computer Science, G. Azzopardi and N. Petkov, Eds. Springer International Publishing, September 2015, vol. 9256, pp. 300 – 311. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-23192-1_25

TABLE II
AFF JOINED NEEDLES DATASET RESULTS OF THE 2DSCAR METHOD.

	Fir	**Larch**	**Scots pine**	**Spruce**	**Swiss stone pine**	**Black pine**	**Sensitivity [%]**
Fir	121	1	0	0	0	0	98,4
Larch	0	106	0	0	1	0	93,0
Scots pine	1	0	112	2	1	4	97,4
Spruce	0	2	0	133	0	3	97,8
Swiss stone pine	1	4	0	0	50	0	94,3
Black pine	0	1	3	1	1	99	93,3
Precision [%]	98,2	99,1	93,3	96,4	90,9	94,3	**Accuracy 96,0**

TABLE III
AFF NEEDLES DATASET RESULTS OF THE 2DSCAR METHOD AND TWO RESOLUTION LEVELS.

	Fir	**Fir2**	**Larch**	**Larch2**	**Scots pine**	**Scots pine2**	**Spruce**	**Spruce2**	**Swiss stone pine**	**Swiss stone pine2**	**Black pine**	**Sensitivity [%]**
Fir	5	0	0	0	0	0	1	0	0	0	0	55,6
Fir2	0	121	0	2	0	2	0	0	0	1	0	98,4
Larch	0	0	17	0	0	0	0	0	0	0	0	94,4
Larch2	0	1	0	108	0	1	0	0	0	3	0	94,7
Scots pine	2	0	0	0	6	0	0	0	0	1	0	66,7
Scots pine2	0	1	0	0	0	111	0	2	0	0	8	96,5
Spruce	2	0	0	0	3	0	13	0	0	0	0	100
Spruce2	0	0	1	0	0	0	0	136	0	0	2	100
Swiss stone pine	0	0	0	0	0	0	0	0	20	0	0	100
Swiss stone pine2	0	0	0	3	0	0	0	0	0	48	0	90,6
Black pine	0	0	0	1	0	1	0	0	0	1	96	90,6
Precision [%]	100	96,0	100	95,6	75,0	92,5	72,2	97,8	100	94,1	97,0	**Accuracy 95,1**

TABLE IV
AFF JOINED NEEDLES DATASET RESULTS OF THE 2DSCAR METHOD AND TWO RESOLUTION LEVELS.

	Fir	**Larch**	**Scots pine**	**Spruce**	**Swiss stone pine**	**Black pine**	**Sensitivity [%]**
Fir	121	3	1	0	1	0	98,4
Larch	1	107	2	0	1	1	93,9
Scots pine	1	0	108	0	0	9	93,9
Spruce	0	1	0	136	0	2	100
Swiss stone pine	0	3	0	0	51	0	96,2
Black pine	0	0	4	0	0	94	88,7
Precision [%]	96,0	95,5	91,5	97,8	94,4	95,9	**Accuracy 95,4**