



ORIGINAL ARTICLE

Open Access



# Detection and visualization of encoded local features as anatomical predictors in cross-sectional images of Lauraceae

Sung-Wook Hwang<sup>1</sup>, Kayoko Kobayashi<sup>2</sup> and Junji Sugiyama<sup>1,3\*</sup>

## Abstract

This paper describes computer vision-based quantitative microscopy and its application toward better understanding species specificity. An image dataset of the Lauraceae family that consists of nine species across six genera was investigated, and structural features were quantified using encoded local features implemented in a bag-of-features framework. Of the algorithms used for feature detection, the scale-invariant feature transform (SIFT) achieved the best performance in species discrimination. In the bag-of-features framework with the SIFT features, each image is represented by a histogram of codewords. The codewords were further analyzed by mapping them to each image to visualize the corresponding anatomical elements. From this analysis, we were able to classify and quantify the modes of aggregation of different combinations of cell elements based on clustered codewords. An analysis of the term frequency–inverse document frequency weights revealed that blob-based codewords are generally shared by all species, whereas corner-based codewords are more species specific.

**Keywords:** Computer vision, Image recognition, SIFT, Optical micrograph, Visual codebook, Wood anatomy

## Introduction

The primary task to be considered for the utilization and research of wood is to correctly identify species. A variety of methods have been established, starting with visual inspection, in which the efforts of wood anatomists are concentrated, to chemometrics, X-ray computed tomography and DNA-based techniques. Although visual inspection based on the anatomical characteristics of wood is the most accurate and reliable identification method, it faces the difficulty of training a sufficient number of highly qualified technicians because the method is entirely specialized. Other techniques also have several limitations such as procedural complexity, field applicability, data sensitivity, and high cost, making them suitable for special purposes rather than general purpose.

With the advent of machine learning and its remarkable development, various problems can be handled by automated systems. The same is true for wood identification. Any kind of data generated from the wood identification methods mentioned above can be applied to machine learning. Image data representing morphological information of wood are more suitable for machine learning than other types of data in various aspects, including applicability, reproducibility, and scalability. Computer vision-based local feature detection is a representative method for extracting morphological information from an image. A local feature consists of information about neighboring pixels in an image and represents local structures or patterns.

The use of local features enables the distinction of different types of wood cells from cross-sectional images [1]. Wood anatomists predict taxa from anatomical features such as wood fibers, axial parenchyma cells, vessels, rays and their arrangements from cross-sectional images. In contrast, computer vision detects spatial and spectral distributions of features such as blobs, corners,

\*Correspondence: sugiyama.junji.6m@kyoto-u.ac.jp

<sup>1</sup> Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Kyoto 611-0011, Japan

Full list of author information is available at the end of the article

and edges, and returns a class. Such local features can effectively capture morphological clues for wood anatomists [1]. Taking these facts into consideration opens up the possibility of interpreting computer vision in terms of wood anatomy. However, wood is not only structurally complex, but many species have similar characteristics. Furthermore, because of the huge number of species, large-scale databases are required, even if only major species are being covered. Local features are often implemented within a bag-of-features (BOF) framework [2] to represent such complicated datasets. The BOF approach is adapted from the bag-of-words (BOW) method [3], which is a method for retrieving information from documents. In BOW, the number of words in a document is counted and a frequency histogram is generated for all words. The histogram allows us to identify the keywords of the document and retrieve it from a number of documents. BOF borrows the same concept to classify images, and text words in BOW are replaced with local features detected in images.

The basic principle of the BOF model with local features for wood recognition is to learn the properties of the detected features and the difference in the number of features of each species. If such differences are discerned, it is also possible to quantify the anatomical elements indirectly. While the information provided by this approach may differ from that produced by the currently established anatomy, the concept of pursuing species specificity in terms of features is fundamentally the same in both approaches. In recent years, some studies have reported that wood species can be recognized using BOF-based models [4, 5]. However, because they used macroscopic images and focused on species classification, these approaches are difficult to interpret from the perspective of the anatomical and morphological characteristics of wood.

All experiments in this study were implemented within the BOF framework. Local features were extracted from the cross-sectional images by the scale-invariant feature transform (SIFT), speeded up robust features (SURF), oriented features-from-accelerated-segment-test (FAST) and rotated binary-robust-independent-elementary-feature (BRIEF) (ORB), or accelerated-KAZE (AKAZE) algorithms. We then selected the best local feature extraction algorithm for the Lauraceae image dataset based on their recognition performances. The quantification and species specificity of structural elements were investigated by analyzing the local features and codewords of the selected algorithm. The codewords, which are analogous to the terms and glossary of anatomical descriptions, have been extensively analyzed. This paper discusses the potential for computer vision to improve or

at least corroborate the currently established anatomy of wood.

## Methods

### Dataset

The Lauraceae image dataset was employed to analyze the structural elements. The dataset contains nine species from six genera of the Lauraceae family, and is composed of 1019 cross-sectional optical micrographs, with each species present in more than 20 images (Table 1). The 8-bit grayscale images have a size of  $900 \times 900$  pixels and a pixel resolution of  $2.94 \mu\text{m}$ . This resolution is close to the theoretical resolving power of our optic system. To enable the recognition models to select the best feature detection algorithm, the dataset was divided into training and test sets at a two-to-one ratio, as presented in Table 1, and both sets consist of images collected from independent individuals.

### Feature extraction

Local features were extracted from the cross-sectional optical micrographs using the well-known SIFT [6], SURF [7], ORB [8], and AKAZE [9] algorithms. Their features are basically invariant to scale, rotation, and limited affine changes. The algorithms were implemented by OpenCV [10], and all their parameters except for those of the ORB algorithm follow OpenCV's defaults. For ORB, the default maximum number of features that can be detected from an image is 500, which is extremely small for cross-section images. We therefore employed ORB with the maximum number of features set to 15,000 and 30,000 (referred to as ORB (15,000) and ORB (30,000), respectively).

**Table 1** List of the species in the Lauraceae image dataset and division of the dataset into training and test sets

Species	Number of wood blocks	Number of images		
		Training set	Test set	Total
<i>Cinnamomum camphora</i>	21	149	72	221
<i>Cinnamomum japonicum</i>	15	103	55	158
<i>Laurus nobilis</i>	3	14	6	20
<i>Lindera glauca</i>	3	20	10	30
<i>Lindera umbellata</i>	10	70	30	100
<i>Litsea coreana</i>	18	120	60	180
<i>Machilus japonica</i>	5	30	20	50
<i>Machilus thunbergii</i>	21	140	70	210
<i>Sassafras tzumu</i>	3	32	18	50
Total	99	678	341	1019

**Codebook generation**

In this study, we implemented a BOF framework to analyze structural elements and recognize wood species (Fig. 1). The descriptors of all the extracted features were grouped into a specified number ( $k$ ) of clusters, with the centroids from each group extracted as codewords, which are also called visual words (Fig. 1c). The  $k$  codewords formed the codebook, which is analogous to a dictionary in text analysis. The number of codewords that are used to create the codebook is arbitrary and is determined by  $k$ , which is how many clusters are used to classify the image features.

The  $k$ -means clustering algorithm is a popular method of generating a codebook, but it can lead to memory problems because of the high computational cost involved in processing massive datasets [11, 12]. Therefore, we used the mini-batch  $k$ -means algorithm as an alternative [13]. This algorithm does not use all data, but instead repeatedly takes a subset of fixed size, drastically reducing the computational cost. Mini-batch  $k$ -means

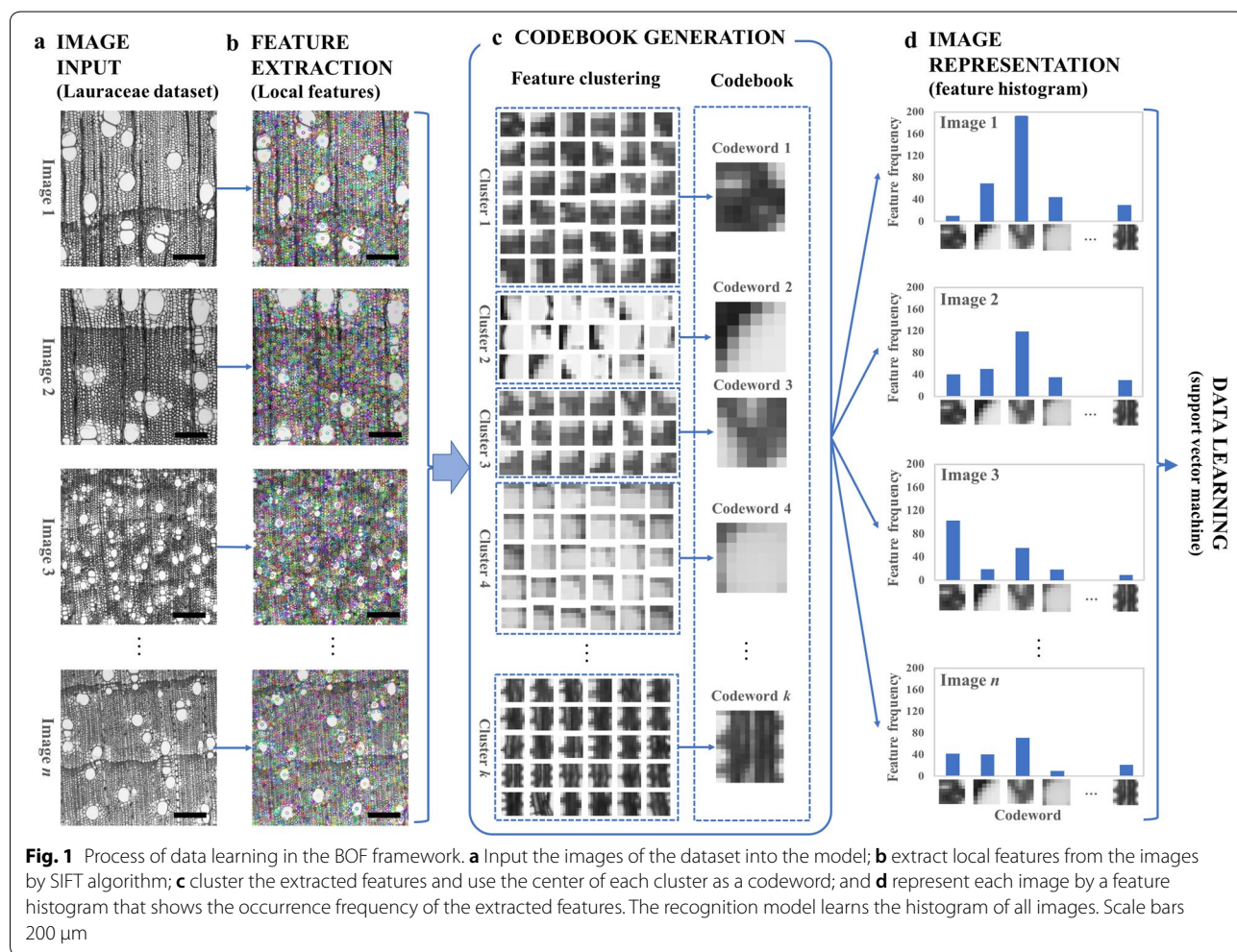
clustering was implemented by the  $k$ -means++ algorithm [14] with a processing batch size of 100. The number of clusters ( $k$ -size) of 500 and 1000 were used to generate codebooks.

**Image representation**

Once a codebook has been completed, it can be used to represent each image. In other words, the image is represented as a histogram of the occurrence frequency of codewords in the codebook (Fig. 1d). All images have individual histograms with  $k$  bins. This is called vector quantization. The key idea of our approach is that the histograms are then used for anatomical analysis and wood recognition.

**Data learning and wood recognition**

The support vector machine (SVM) classifier with a radial basis function (RBF) kernel was used to learn the data for wood recognition. For classification, SVM finds the hyperplane with as wide a margin as possible between the



input data of different classes. Whether linear or nonlinear models, SVM always classifies data linearly. Nonlinear models that map data into high-dimensional feature space to create hyperplanes are effective in classifying data that are difficult to distinguish linearly. The RBF kernel SVM uses a Gaussian kernel to map data into high-dimensional spaces [15]. The training set histograms were analyzed as feature vectors and input to the support vector machine classifier to learn the class boundaries. To optimize the parameters, we set up grid searches with a logarithmic grid from  $10^{-3}$  to  $10^3$  for gamma and from  $10^{-8}$  to  $10^{-2}$  for C. Gamma is a Gaussian kernel parameter for nonlinear classification, and C is a parameter that controls the cost of misclassification on the training data.

The F1 score, the harmonic mean of precision and recall, was used as a metric to evaluate the performance of the recognition models, which is given by

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Positive (FP)'}}$$

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)'}}$$

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}'}$$

where TP is a test result in which the model correctly predicts a positive class, FP an error predicted as a positive class, and FN an error predicted as a negative class. Because the Lauraceae dataset has quite imbalanced classes (Table 1), it is inappropriate to evaluate the model based on recognition accuracy. The F1 score tells us how precise and robust the model is. Based on the F1 score, we selected the best feature extraction algorithm for the Lauraceae images and used it to analyze the structural elements. In this study, the computational cost was not a consideration in algorithm selection.

#### Codeword clustering and assignment to anatomical elements

We clustered the codewords generated by the feature extraction algorithm that was shown to have the best discrimination based on recognition performance. For anatomical analysis, codewords generated from the local features of the selected algorithm were clustered with those with similar descriptor characteristics. Agglomerative hierarchical clustering was performed on the codewords using the Euclidian distance through Ward's method [16]. The codewords were further analyzed by mapping them onto each image to visualize the corresponding anatomical structure. The number of local features (keypoints) of the codewords contained in each

species was then evaluated to indirectly quantify anatomical elements.

#### Feature weights

In a codebook, some codewords may be common to many species, whereas others are more species specific. It is possible to classify such uncommon codewords using the term frequency-inverse document frequency (TFIDF) weighting [17], which is given by

$$w_{(i,j)} = \frac{tf_{i,j}}{\sum_i tf_{i,j}} \cdot \log \left( \frac{N+1}{df_j+1} \right),$$

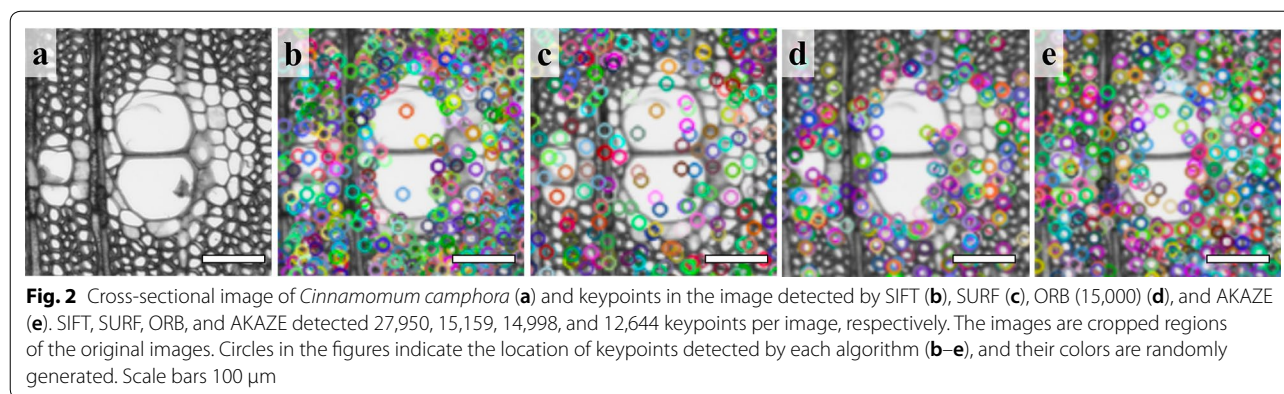
where  $w$  is the TFIDF weight,  $tf_{i,j}$  the frequency of feature  $j$  in image  $i$ ,  $df_j$  the number of images containing feature  $j$ , and  $N$  the number of images. The first and second terms in the formula represent the term frequency (TF) and inverse document frequency (IDF) weights, respectively. A codeword with a high TFIDF score indicates a rare and unique feature present in a small number of species, whereas a low score denotes a feature shared by many species. The score of common features detected in all images becomes zero. The IDF and TFIDF scores respectively provide us with the rarity and species specificity of each codeword. The program for all the experiments was written in Python 3.5.2 with various external packages.

## Results and discussion

### Comparison of local features

SIFT detected the largest number of keypoints in various types of wood cells, including vessels, axial parenchyma cells, rays, and wood fibers (Fig. 2b). Because the SIFT algorithm is basically designed to detect corners and blobs as local features, cell corners and cell lumina are primarily detected as keypoints in cross-sectional images, exposing the aggregation of numerous wood cells with lumina. In fact, SIFT detected them more aggressively than the other algorithms. A cell corner can be detected when the pixel resolution is more than  $2.94 \mu\text{m}$  and was found to be a key feature that determines the discriminative power of SIFT [1]. SURF also detected various cells more densely in vessel lumina and rays, but the final number of keypoints is slightly over half that of SIFT (Fig. 2c). This means that SURF misses a significant number of wood cells that SIFT does not.

ORB revealed a different aspect of the wood images than that revealed by SIFT and SURE, in that most keypoints were detected in blobs (Fig. 2d). This trend was common in both ORB (15,000) and ORB (30,000). Both ORB methods were implemented by combining the FAST keypoint detector [18] and the BRIEF descriptor [19] to improve performance. Although ORB uses FAST, which is an algorithm that detects corners as features, it



**Table 2 F1-scores of wood recognition models established using local features detected by the SIFT, SURF, ORBs, and AKAZE algorithms**

Algorithm	Codebook size	
	500	1000
SIFT	0.771	0.792
SURF	0.422	0.279
ORB (15,000)	0.545	0.639
ORB (30,000)	0.607	0.633
AKAZE	0.616	0.607

did not detect any corners in the cross-sectional images, which are made up of numerous corners. This is because FAST certainly detects corners, but its keypoints, unlike those of other algorithms, represent a pixel surrounded by a corner, not a pixel corresponding to a corner.

AKAZE, with the least number of keypoints, found various types of wood cells (Fig. 2e). It detected features in both cell corners and lumina, but those in the lumina were significantly more numerous. In contrast to SIFT and SURF, which employ Gaussian filtering for feature detection, AKAZE adopts nonlinear diffusion filtering to preserve image details and remove noise. However, the algorithm works similarly to SIFT for the detection of most wood fibers, and similarly to SURF for the presence of many keypoints in vessel lumina.

### Performance of local feature extractors on wood recognition

#### Wood recognition performance of local features

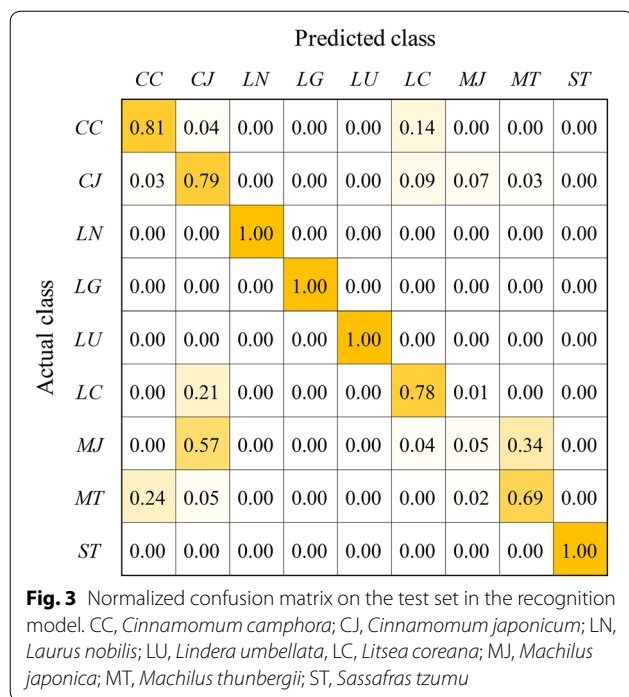
The F1 scores of each recognition model using the local features detected by the SIFT, SURF, ORBs, and AKAZE algorithms are presented in Table 2. For both codebook sizes, SIFT achieved the best F1 scores. The ORBs were ranked next, but their performance gaps with SIFT are quite large. Although ORB (30,000) detects twice as

many keypoints as ORB (15,000), its F1 score did not surpass that of ORB (15,000). SURF produced such poor recognition performance that it would not be possible to perform further analyses based on its results.

Hwang et al. [1] reported that cell corners and cell lumina should contain important information for wood classification, given that a reduction in the number of keypoints detected in the cell corners and lumina results in a decrease in classification accuracy. In a cross-section of wood, cell lumina represent the characteristics of single cell types, whereas cell corners are a more important factor for wood recognition because they contain information not only about single cells but also about aggregates of various types of cells. Therefore, the high discriminative power of SIFT seems to be closely related to its cell corner detection capability, which is superior to that of the others. In a study by Tareen and Saleem [20], which compared the image matching performance of the well-known local feature extraction algorithms SIFT, SURF, KAZE, AKAZE, ORB, and BRISK (binary robust invariant scalable keypoints), they reported that SIFT achieved the best performance if the computational cost was not taken into account. The same results were reported by Karami et al. [21]. Given these results, we selected SIFT as the optimal algorithm for further wood anatomy analysis.

#### Recognition error

A normalized confusion matrix of the recognition model using the SIFT features is presented in Fig. 3. The confusion matrix is a table that compares the actual and predicted classes to measure the performance of the established model. As seen in Fig. 3, *Laurus nobilis*, *Lindera glauca*, *Lindera umbellata*, and *Sassafras tzumu* were recognized perfectly even though the number of images used to train the model was small. For *Machilus japonica*, in contrast, 91% of its test images were



misrecognized as *Cinnamomum japonicum* and *Machilus thunbergii*.

Table 3 shows the International Association of Wood Anatomist (IAWA) list of microscopic features for hardwood identification [22]. It also presents anatomical features in cross-sections for the misrecognized species; in fact, their cross-section anatomical features are very similar. *M. japonica* shares most of its anatomical features with *C. japonicum* and *M. thunbergii* and even covers all the features of *M. thunbergii*. In comparison with *C. japonicum*, *M. japonica* has vessels that are larger in tangential diameter, but theoretically, the SIFT algorithm detects features that are invariant in scale, so the difference in lumen diameters or cell size is not considered to be an important factor for recognition. For *M. thunbergii*, a quarter of its test images were misrecognized as *C. camphora*, whereas *C. camphora* images were never recognized as *M. thunbergii*. This means that although they have similar anatomical features, *C. camphora* has sufficient species-specific features that distinguish it from others, but *M. thunbergii* does not.

**Codewords decomposed into anatomical elements**

The number of codewords influences the discrimination of local features. To determine the codebook size (*k*) that maximizes the discriminative power of the SIFT features, we performed threefold cross-validation on the training set with various numbers of codewords (from 100 to 1900). The minimum cross-validation error was

produced at codebook size of 1000, and similar levels were obtained at larger sizes. When the codebook size is excessively large, there is a high possibility of introducing artifacts into the histogram or overfitting, and computation is inefficient. In consideration of the possible problems, therefore, we used a codebook with 1000 codewords for further analyses.

Figure 4 shows a dendrogram of 1000 codewords forming six clusters and a visualization of some codewords belonging to each cluster. The codewords are mainly classified into two groups, corner-based (as in clusters 1, 3, and *k* in Fig. 1c) and blob-based groups (as in clusters 2 and 4 in Fig. 1c), which are further classified into four and two sub-clusters (Fig. 4a).

Together with the visual assignment of codewords, approximately 60% of the codebook refers to corners and blobs generally present in various types of wood cells, and they form their own clusters (clusters III and VI in Fig. 4a). In the corner-based group, vessels, axial parenchyma cells, and rays, wood fibers, and vessel boundaries can be divided into different clusters. The corner-based codewords in vessels are included in two clusters (clusters I and IV in Fig. 4a). The first cluster includes corner-based codewords present in the rather long, smooth and thicker cell walls, which are independent from the cell type but more often detected in rays, whereas the codewords in the fourth cluster are always present in the vessel cell walls that connect with the other type of adjacent cells. In the blob-based group, the codewords are included in two clusters. These clusters are divided by the location of SIFT keypoints in the cell lumen, namely, the center or portions near edges, regardless of the cell types. These results show that the corner-based codewords have different characteristics for anatomical elements compared with those based on blobs, and they vary depending on the type of adjacent wood cells.

**Rarity and species specificity of codewords**

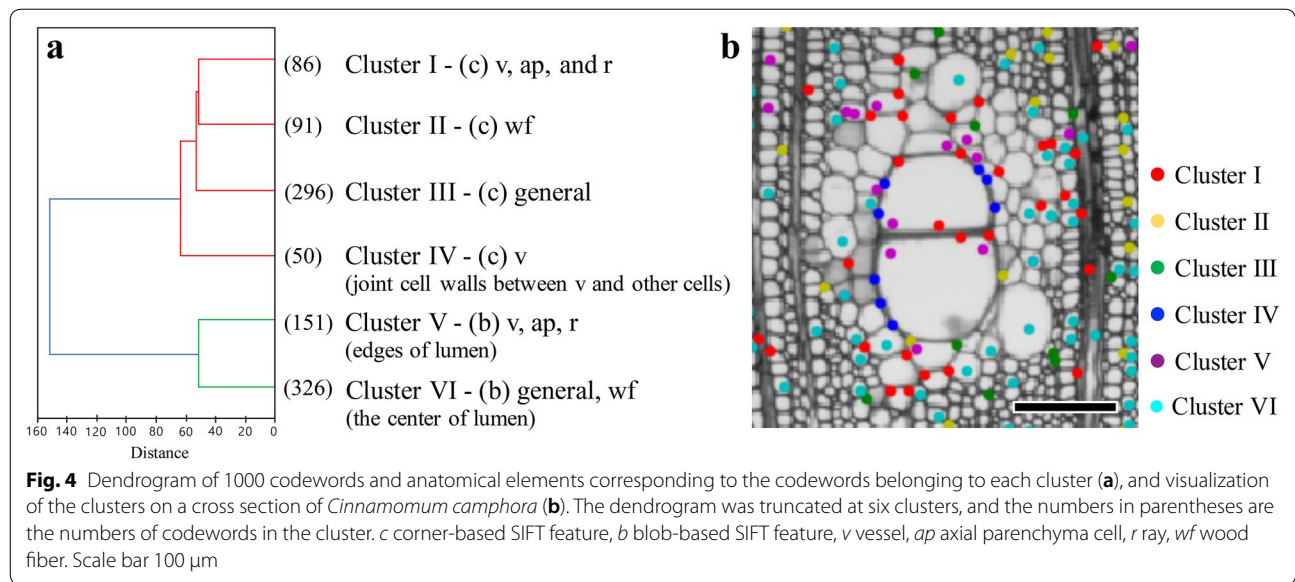
The 20 codewords with the highest IDF values and their anatomical elements as well as the average number of SIFT keypoints for each codeword by species are presented in Table 4. If some artifacts introduced into an image by chance or the noise generated during image processing are detected as local features, they are likely to be converted into codewords with high IDF values even though they are not informative. To avoid such problems, we excluded codewords with less than 10 keypoints in all species.

The IDF reflects how common or rare a codeword is. Of the 1000 codewords, 452 have IDF values of 0, indicating that 45% of the codewords are common features in the Lauraceae dataset. Hence, the 55% of codewords

**Table 3 Microscopic features in the cross-section of misrecognized species**

Growth ring boundary	Vessel		SV	TD (µm)	Frequency (/mm <sup>2</sup> )	Tyloses and deposits	Fiber		Parenchyma		Ray frequency (/mm <sup>2</sup> )
	VP	VP					WT	PAP	AAP	PAP	
<i>M. japonica</i>	Diffuse	Diffuse	-	50-100, 100-200	5-20, 20-40	Common	Very thin, thin to thick	diffuse	Scanty paratracheal, vasicentric, confluent	4-12	
<i>C. japonicum</i>	Diffuse	Diffuse	-	50-100	-	Common	Very thick	diffuse	Vasicentric, confluent	4-12	
<i>M. thunbergii</i>	Diffuse	Diffuse	-	50-100, 100-200	5-20, 20-40	Common	Very thin, thin to thick	diffuse	Confluent	4-12	
<i>C. camphora</i>	Diffuse or indistinct	Diffuse or semi-ring	Angular	100-200	5-20	Common	Very thin, thin to thick	-	Vasicentric, confluent	4-12	

VP vessel porosity, SV solitary vessel outline, TD tangential diameter of lumina, WT wall thickness, AAP apotracheal axial parenchyma, PAP paratracheal axial parenchyma



**Table 4** Twenty codewords with the highest IDF and their corresponding anatomical elements, and the average number of SIFT keypoints belonging to each codeword by species

Codeword number	IDF value	Feature type	Anatomical elements	Cluster	Average number of SIFT keypoints in the codeword								
					CC	CJ	LN	LG	LU	LC	MJ	MT	ST
751	0.180	c	ap	I	10.1	3.5	0.6	0.8	0.7	3.2	2.3	5.3	7.6
820	0.137	c	ap	I	3.1	3.0	1.3	0.9	2.1	2.8	1.8	3.1	10.9
313	0.120	c	v	IV	3.6	3.5	4.1	11.1	2.6	4.8	5.8	4.7	2.9
878	0.114	c	ap	I	4.5	4.0	3.9	1.6	1.9	4.5	2.7	4.5	17.5
847	0.109	c	General	III	20.0	12.2	4.4	1.4	1.3	8.7	4.9	14.9	21.8
923	0.097	c	ap	I	3.3	2.9	4.4	13.6	3.9	4.9	4.7	4.6	1.9
332	0.089	b	General	III	4.3	3.1	9.9	17.3	8.4	7.0	4.2	6.1	4.4
682	0.089	c	r	I	4.0	3.5	21.6	9.0	9.6	14.2	2.1	5.9	8.7
850	0.086	c	v	IV	8.9	6.8	1.9	0.8	1.3	5.4	2.9	6.1	23.1
294	0.086	c	ap	I	39.2	9.9	1.9	2.1	2.0	9.5	7.0	16.1	22.3
760	0.083	c	wf	II	2.7	2.8	13.1	5.3	5.5	8.6	2.6	4.7	4.3
518	0.080	c	General	III	4.3	4.3	7.9	32.8	7.4	5.8	4.8	4.6	1.6
619	0.075	c	General	III	5.0	6.2	6.4	2.4	3.5	6.6	4.3	6.7	11.6
948	0.075	c	General	III	15.7	15.7	1.7	1.3	1.0	6.1	5.6	15.2	19.9
225	0.073	b	r, ap	V	8.1	7.1	7.1	42.1	7.0	7.6	10.5	5.4	2.8
890	0.070	c	r	I	5.3	4.5	46.9	47.0	7.1	18.1	3.1	8.5	9.9
293	0.067	c	General	III	18.8	12.3	2.6	3.4	1.5	7.7	5.9	12.0	10.6
151	0.064	b	v	V	31.2	19.9	26.9	1.3	6.0	9.2	0.8	40.7	59.6
809	0.064	c	r <sup>a</sup>	II	10.4	6.2	5.0	9.4	5.6	5.8	2.5	7.9	5.4
635	0.062	c	r	I	54.0	170.3	156.4	0.2	29.8	27.5	2.7	94.3	70.4

c, corner-based codeword; b, blob-based codeword; ap, axial parenchyma cells; v, vessels; r, rays; wf, wood fibers; general, observed in more than three different types of wood cells; CC, *Cinnamomum camphora*; CJ, *Cinnamomum japonicum*; LN, *Laurus nobilis*; LG, *Lindera glauca*; LU, *Lindera umbellata*; LC, *Litsea coreana*; MJ, *Machilus japonica*; MT, *Machilus thunbergii*; ST, *Sassafras tzumu*. <sup>a</sup>Peripheral cells in contact with rays, mainly wood fibers

with values greater than 0 contain species-specific information. In addition, although the number of codewords and keypoints representing corners in axial parenchyma

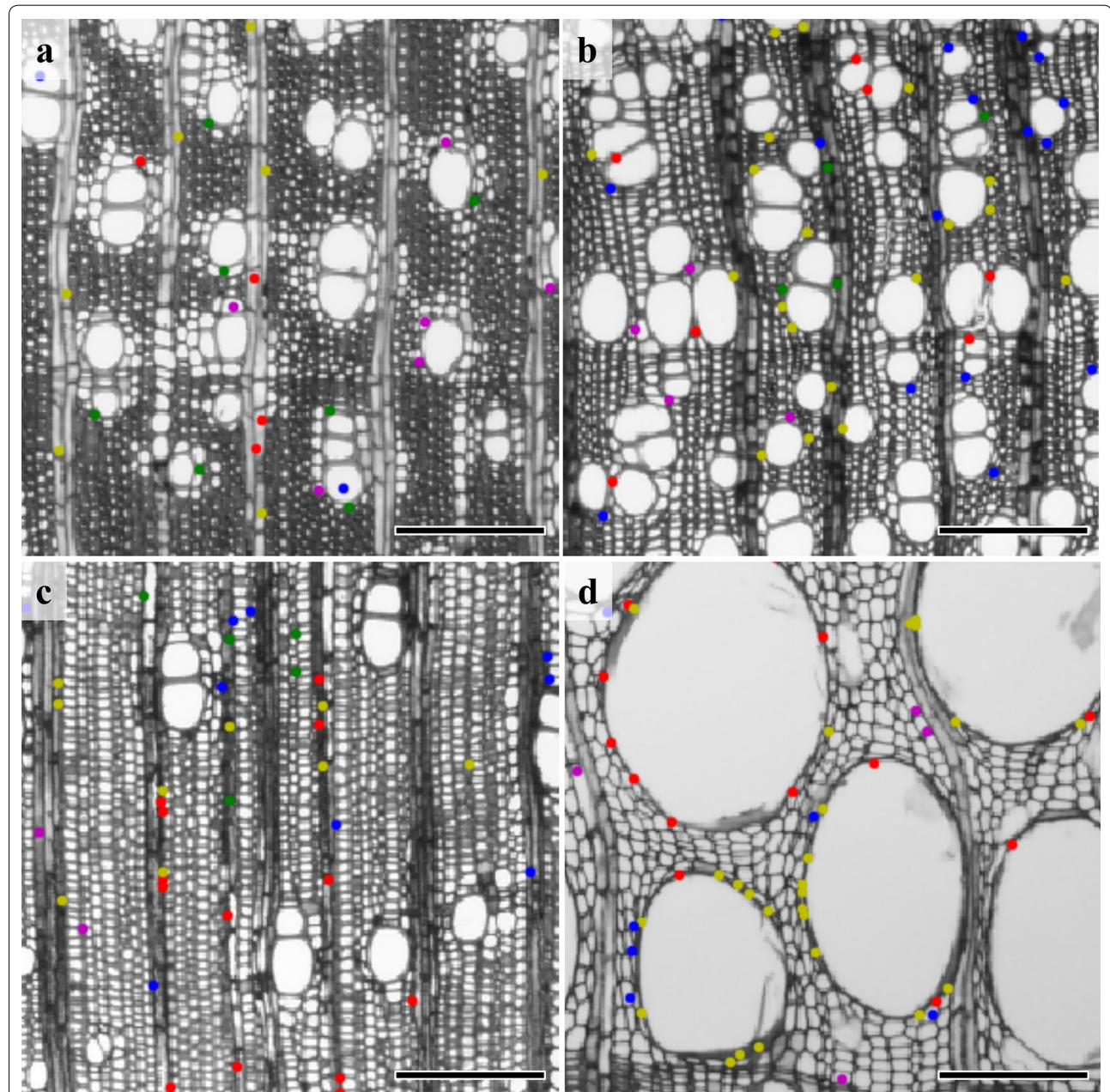
cells, vessels, and rays is relatively small, they are ranked in the top class in terms of IDF values. The IDF reveals that blob-based codewords are generally shared by many



species, whereas corner-based codewords are more species specific.

The TFIDF score can be used to predict the feature importance by IDF, taking into consideration the number of keypoints. Although codewords 751 and 820, which have the two highest IDF values, both represent corners in axial parenchyma cells, the former is abundant in *C. camphora* and the latter is abundant in *S. tzumu*

(Table 4). In addition, codeword 923, which represents the same element, is dominant in *L. glauca*. For vessels, codewords 313 and 850 are abundant in *L. glauca* and *S. tzumu*, respectively. Rays are also species dependent (codewords 682, 225, and 890), and *L. nobilis* in particular has many keypoints in the ray-related codewords. *L. glauca* rarely has keypoints in codeword 635, which is abundant in most species. Hence, this codeword



**Fig. 5** Visualization of SIFT keypoints belonging to the five codewords with the highest TFIDF scores from *Lindera glauca* (a), *Lindera umbellata* (b), *Laurus nobilis* (c), and *Sassafras tzumu* (d). The images show an enlarged part of the original images, and the scores run from high to low in the order red, yellow, green, blue, and purple. Scale bars 200  $\mu$ m

deficiency is in fact a species-specific characteristic of *L. glauca* in the Lauraceae dataset.

These results indicate that even within the same type of wood cells, codewords represent different local features and some of them are species dependent. Thus, the TFIDF score, which considers both IDF and the number of features simultaneously, is a useful tool for determining the differences in the anatomical elements of wood.

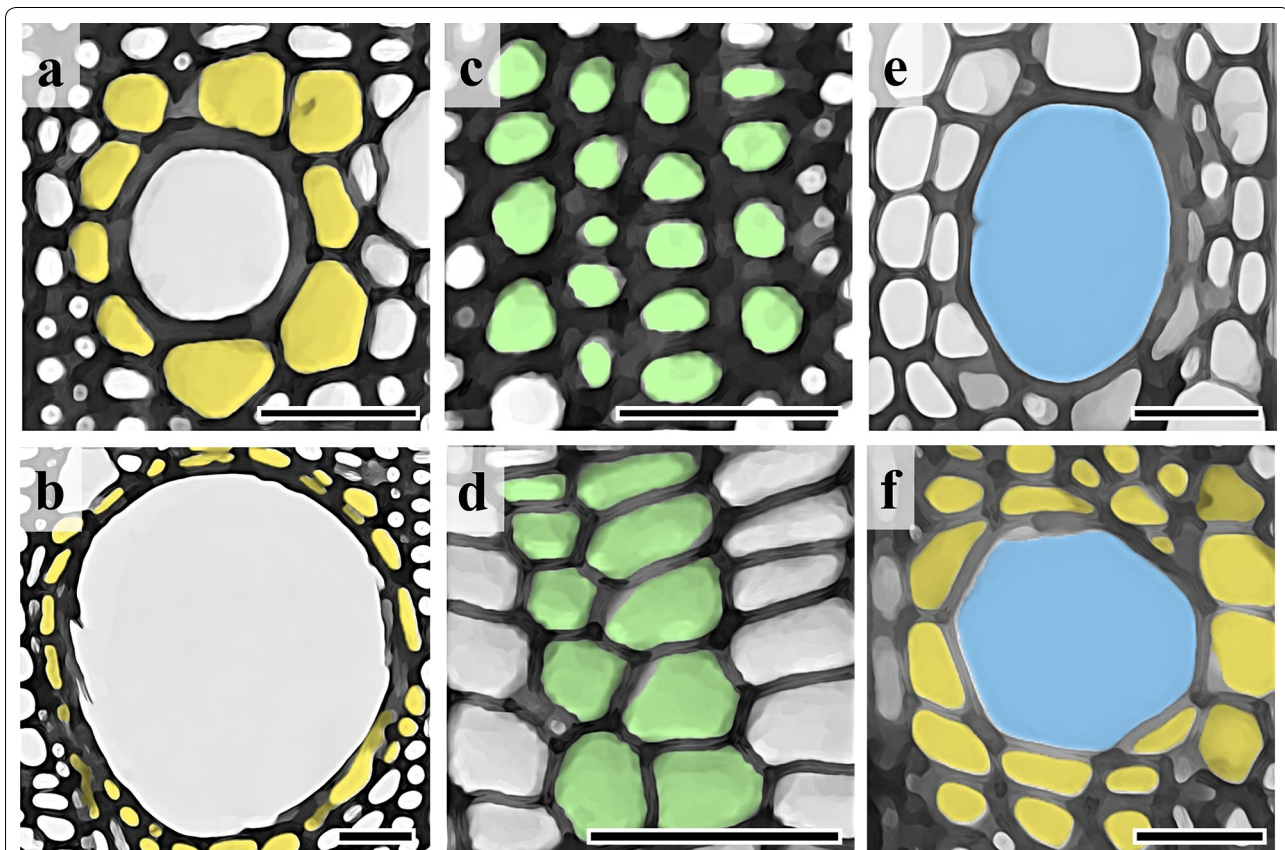
### Species-specific features

The codewords with the top five TFIDF scores of four species that were correctly recognized among all the species investigated are mapped and visualized on the original optical micrographs of the corresponding species (Fig. 5). In the image of *L. glauca* (Fig. 5a), the keypoints are distributed in the rays, axial parenchyma cells, and vessels, especially in their cell lumina. In contrast, in the images of the other three species, the keypoints are mainly distributed in corners in vessels (Fig. 5b, d), and rays (Fig. 5c). Such anatomical elements can be considered species-specific features, which are not shared by all species. In contrast, *M.*

*japonica*, the most misrecognized species, is the only species that does not have a specific wood cell that is species specific. This species has relatively high TFIDF scores in the corner- and blob-based features that are common to various wood cells.

For *S. tzumu*, which is the only ring porous species, large vessels in earlywood have high scores (Fig. 5d). Although the SIFT feature is invariant in scale, the reason for the high TFIDF scores of large vessels is that the formation of large vessels leads to morphological deformation of their adjacent cells, resulting in local features that are differentiated from the others (Fig. 6a, b). SIFT ignores individual cell type and size but detects a variety of the mode of aggregation of different cell elements.

Because the SIFT descriptors represent the gradient orientation of local area in the image, the shape and/or arrangement of wood cells is an important factor in determining species-specific features. Figure 6c, d illustrates different types of wood fibers in *C. camphora*. In the Lauraceae dataset, rounded wood cells are a feature common to all species, whereas polygonal cells are likely to be species specific with high TFIDF scores.



**Fig. 6** Illustrations of wood cells with different morphological characteristics in the same anatomical features. Vasicentric axial parenchyma cells of *Sassafras tzumu* in latewood (a) and earlywood (b). Wood fibers of *Cinnamomum camphora* with round (c) and polygonal shapes (d). Solitary vessels of *Machilus thunbergii* with a rounded cell wall (e) and *Cinnamomum camphora* with an angular wall (f). Scale bars 50  $\mu$ m

Figure 6e, f presents the solitary vessels of *M. thunbergii* and *C. camphora*, respectively. According to the TFIDF score, *C. camphora* has angular corners in vessels and axial parenchyma cells as species-specific features. Recalling the relationship between this species and *M. thunbergii* in wood recognition, *M. thunbergii* had a high probability of being misrecognized as *C. camphora*, whereas *C. camphora* does not, even though they have very similar anatomical features. The angular outline of solitary vessels present in *C. camphora* is the only anatomical difference in the cross-section based on the IAWA feature code that distinguishes *C. camphora* and *M. thunbergii* (Table 3), and the proposed method suggests it is species specific. Similar to the large vessels of *S. tzumu*, the solitary vessels with angular outlines of *C. camphora* influence the formation of vasicentric axial parenchyma cells, resulting in unique morphological characteristics.

In the BOF framework, codewords as anatomical predictors allowed us to quantify the aggregation of different combinations of wood cells, and the TFIDF score provided us with information on species-specific features. With a codebook and local features, we have moved one step closer toward understanding how the features used by computer vision for recognition relate to wood anatomy.

## Conclusion

A BOF model based on local features was established to investigate the Lauraceae image dataset. Of the well-known local feature extraction algorithms implemented, the SIFT algorithm was selected as the best for the dataset and was used for distinguishing anatomical elements. The resulting codebook allowed us to approach computer vision from the perspective of wood anatomy. In the same anatomical elements, the local features varied with respect to the type of adjacent cells, and according to the TFIDF weights, these are important features for recognition in some species. The local features encoded with codewords are promising anatomical predictors that were able to indirectly assess the anatomical characteristics of Lauraceae. Further efforts to remove the gap between human and computer vision, as well as to reduce the gap between informatics and wood anatomy, will be the focus of subsequent studies.

## Abbreviations

SIFT: Scale-invariant feature transform; BOF: Bag-of-features; BOW: Bag-of-words; SURF: Speeded up robust features; FAST: Features-from-accelerated-segment-test; BRIEF: Binary-robust-independent-elementary-feature; ORB: Oriented FAST and rotated BRIEF; AKAZE: Accelerated-KAZE; SVM: Support vector machine; RBF: Radial basis function; TFIDF: Term frequency-inverse document frequency; BRISK: Binary robust invariant scalable keypoints.

## Acknowledgements

We are indebted to Professor Hiroaki Kawashima, Graduate School of Information Science, Kyoto University, for patient guidance, encouragement, and advice throughout this study. Most of the micrograph datasets were prepared by Ms. Keiko Kanai and Ms. Izumi Kanai, which is also acknowledged. We thank Kimberly Moravec, PhD, from Edanz Group (<http://www.edanzediting.com/ac>) for editing a draft of this manuscript. SH is a recipient of JSPS Postdoctoral Fellowship for Foreign Researchers.

## Authors' contributions

SH implemented the model, performed data production and analysis, and was a major contributor in writing the manuscript. KK supported data analysis and the technical aspects needed to implement the model. JS designed and supervised the work. All authors read and approved the final manuscript.

## Funding

This study was supported by Grants-in-Aid for Scientific Research (Grant Numbers 25252033, 17F17402, 18H05485) from the Japan Society for the Promotion of Science, RISH Cooperative Research (database), and RISH Mission Research V to JS.

## Availability of data and materials

The datasets generated and analyzed during the current study are available in the Kyoto University Research Information Repository, URL: <http://hdl.handle.net/2433/245888>.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Research Institute for Sustainable Humanosphere, Kyoto University, Uji, Kyoto 611-0011, Japan. <sup>2</sup> Department of Biomaterial Science, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113-8657, Japan. <sup>3</sup> College of Materials Science and Engineering, Nanjing Forestry University, Nanjing 210037, China.

Received: 30 October 2019 Accepted: 27 February 2020

Published online: 07 March 2020

## References

- Hwang SW, Kobayashi K, Zhai S, Sugiyama J (2018) Automated identification of Lauraceae by scale-invariant feature transform. *J Wood Sci* 64:69–77
- Csurka G, Dance C, Fan L, Willamowski J, Bray C (2004) Visual categorization with bags of keypoints. In: Workshop on statistical learning in computer vision, ECCV 2004, Prague, 11–14 May 2004
- Joachims T (1998) Text categorization with support vector machines: Learning with many relevant features. In: Nédellec C, Rouveirol C (eds) Machine learning: ECML-98, Chemnitz, April 1998. ECML 1998. Lecture notes in computer science (lecture notes in artificial intelligence), vol 1398. Springer, Berlin, Heidelberg, pp 137–142
- Hu S, Li K, Bao X (2015) Wood species recognition based on SIFT keypoint histogram. In: Proceedings of the 2015 8th international congress on image and signal processing (CISP), Shenyang, 14–16 October 2015
- Barmpoutis P, Dimitropoulos K, Barbutis I, Grammalidis N, Lefakis P (2018) Wood species recognition through multidimensional texture analysis. *Comput Electron Agric* 144:241–248
- Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Intl J Comput Vis* 60(2):91–110
- Bay H, Tuytelaars T, Van Gool L (2006) SURF: Speeded up robust features. In: Leonardi A, Bischof H, Pinz A (eds) Computer Vision—ECCV 2006. ECCV 2006, Graz, May 2006. Lecture notes in computer science, vol 3951. Springer, Berlin, Heidelberg, pp 404–417
- Rublee E, Rabaud V, Konolige K, Bradski GR (2011) ORB: An efficient alternative to SIFT or SURF. In: Proceedings of the 2011 international conference on computer vision (ICCV 2011). Barcelona, 6–13 November 2011

9. Alcantarilla PF, Nuevo J, Bartoli A (2013). Fast explicit diffusion for accelerated features in nonlinear scale spaces. In: Proceedings of the British machine vision conference (BMVC 2013). Bristol, 9–13 September 2013
10. Itseez (2016) Open source computer vision library. <http://opencv.org/>. Accessed 14 Jun 2018
11. Bradley PS, Fayyad UM, Reina C (1998) Scaling clustering algorithms to large databases. In: Proceedings of the 4th international conference on knowledge discovery and data mining (KDD '98). New York, 27–31 August 1998
12. Shindler M, Wong A, Meyerson A (2011) Fast and accurate k-means for large datasets. In: Proceedings of the 25th international conference on neural information processing systems (NIPS '11). Granada, 16–17 December 2011
13. Sculley D (2010) Web-scale k-means clustering. In: Proceedings of the 19th international conference on world wide web (WWW '10). Raleigh, 26–30 April 2010
14. Arthur D, Vassilvitskij S (2007) k-means++: the advantages of careful seeding. In: Proceedings of the 18th annual ACM-SIAM symposium on discrete algorithms (SODA '07). New Orleans, 7–9 January 2007
15. Vert JP, Tsuda K, Schölkopf B (2004) A primer on kernel methods. In: x Vert JP, Tsuda K, Schölkopf B (eds) Kernel methods in computational biology. MIT press, Cambridge, pp 35–70
16. Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244
17. Szeliski R (2010) Computer vision: algorithms and applications. Springer-Verlag, London
18. Rosten E, Drummond T (2006) Machine learning for high-speed corner detection. In: Leonardis A, Bischof H, Pinz A (eds) Computer vision—ECCV 2006. ECCV 2006, Graz, May 2006. Lecture notes in computer science, vol 3951. Springer, Berlin, Heidelberg, pp 430–443
19. Calonder M, Lepetit V, Strecha C, Fua P (2010) BRIEF: Binary robust independent elementary features. In: Daniilidis K, Maragos P, Paragios N (eds) Computer vision—ECCV 2010. ECCV 2010, Heraklion, September 2010. Lecture notes in computer science, vol 6314. Springer, Berlin, Heidelberg, pp 778–792
20. Tareen SAK, Saleem Z (2018). A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. In: Proceedings of the 2018 international conference on computing, mathematics and engineering technologies (iCoMET 2018). Sukkur, 3–4 March 2018
21. Karami E, Prasad S, Shehata M (2015) Image matching using SIFT, SURF, BRIEF, and ORB: Performance comparison for distorted images. In: Proceedings of the 24th annual Newfoundland electrical and computer engineering conference (NECEC 2015), St. John's, 15 November 2015
22. Committee IAWA (1989) IAWA list of microscopic features for hardwood identification. *IAWA Bull* 10:219–332

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---