



HAL
open science

Automatic identification and characterization of radial files in light microscopy images of wood

Guilhem Brunel, Philippe Borianne, Gérard Subsol, Marc Jaeger, Yves Caraglio

► **To cite this version:**

Guilhem Brunel, Philippe Borianne, Gérard Subsol, Marc Jaeger, Yves Caraglio. Automatic identification and characterization of radial files in light microscopy images of wood. *Annals of Botany*, 2014, 114 (4), pp.829-840. 10.1093/aob/mcu119 . hal-01076468

HAL Id: hal-01076468

<https://sde.hal.science/hal-01076468v1>

Submitted on 15 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TECHNICAL ARTICLE: PART OF A SPECIAL ISSUE ON FUNCTIONAL–STRUCTURAL
PLANT MODELLING

Automatic identification and characterization of radial files in light microscopy images of wood

Guilhem Brunel^{1,2,*}, Philippe Borianne², Gérard Subsol³, Marc Jaeger² and Yves Caraglio²

¹CIRAD – UMR A51 AMAP, France, ²University of Montpellier 2, France and ³LIRMM – CNRS,
University of Montpellier 2, France

* For correspondence. E-mail guilhem.brunel@cirad.fr

Received: 31 October 2013 Returned for revision: 9 January 2014 Accepted: 31 March 2014 Published electronically: 2 July 2014

- **Background and Aims** Analysis of anatomical sections of wood provides important information for understanding the secondary growth and development of plants. This study reports on a new method for the automatic detection and characterization of cell files in wood images obtained by light microscopy. To facilitate interpretation of the results, reliability coefficients have been determined, which characterize the files, their cells and their respective measurements.
- **Methods** Histological sections and blocks of the gymnosperms *Pinus canariensis*, *P. nigra* and *Abies alba* were used, together with histological sections of the angiosperm mahogany (*Swietenia* spp.). Samples were scanned microscopically and mosaic images were built up. After initial processing to reduce noise and enhance contrast, cells were identified using a ‘watershed’ algorithm and then cell files were built up by the successive aggregation of cells taken from progressively enlarged neighbouring regions. Cell characteristics such as thickness and size were calculated, and a method was developed to determine the reliability of the measurements relative to manual methods.
- **Key Results** Image analysis using this method can be performed in less than 20 s, which compares with a time of approx. 40 min to produce the same results manually. The results are accompanied by a reliability indicator that can highlight specific configurations of cells and also potentially erroneous data.
- **Conclusions** The method provides a fast, economical and reliable tool for the identification of cell files. The reliability indicator characterizing the files permits quick filtering of data for statistical analysis while also highlighting particular biological configurations present in the wood sections.

Key words: Functional–structural plant modelling, image processing, microscopic wood images, radial cell file, reliability coefficients, cellular structure, light microscopy, *Pinus canariensis*, *P. nigra*, *Abies alba*, *Swietenia*.

INTRODUCTION

Tree development is the result of both primary growth, involving the elongation and branching of the axes, and secondary growth, consisting of their thickening over time. Many approaches have been developed to study the structural and functional aspects of plants, particularly when used to address issues of carbon sequestration and wood energy. These approaches, however, whether examining secondary growth and its inter- or intra-specific variations, or the relationships between primary and secondary growth, are often based on only fragmentary studies due to the fact that data acquisition is often costly (Rossi *et al.*, 2009).

The description of primary growth can be more or less extensive thanks to retrospective approaches (Barthélémy and Caraglio, 2007; Krause *et al.*, 2010) while dendrochronological studies are generally performed only on portions of the ligneous plane (coring). Here, the more detailed the observation level, the shorter the part of the ligneous plane described by the data. Secondary growth originates from cambial activity, the source of the wood (xylem), and thus from the different cells that make up the xylem (ligneous plane).

The biological typing (Lachaud *et al.*, 1999) of cells can be approached by examining their shape, size and spatial distribution,

independent of their organization. The findings can then be used to determine the anatomical changes that characterize xylem production and differentiation phases (Thibeault-Martel *et al.*, 2008; Guédon *et al.*, 2013) without truly explaining wood production or growth mechanisms.

Indeed, the mechanisms of cambial growth can be understood only by studying cell pattern rhythmicity and cell disruption or modification in space and time. A spatio-evolutionary perspective must therefore be introduced to discuss these issues. Environmental fluctuations and their effects on the differentiation of wood elements (originating from the divisions of cambial cells) require production to be monitored in specific cellular organizations. Here, for the sake of simplicity, we consider two organizations: first, the growth ring that reflects cell production over a given time (Heinrich, 2007) and, secondly, the cell file that reflects the activity of the initial cell over time (Rozenberg *et al.*, 2004).

Nicolini *et al.* (2003) highlighted the importance of secondary growth organization by studying cell motifs, spatial rhythmicity and the variability of successive growth rings to characterize plant space occupation strategies. This approach does not, however, address the mechanisms of cell differentiation or the construction of cell motifs and structures, which are time-dependent mechanisms. The precise, dynamic monitoring of cambial

production at the initial cells (Rossi *et al.*, 2006; Rathgeber *et al.*, 2011) is technically limited to a small portion of the ligneous plane. Thus, the study of cell files, which result from cambial production and its local and overall fluctuations, is a promising approach to understanding the establishment, differentiation and temporal rhythmicity of cells (Gindl, 2001).

Biological questions concerning interactions between the development and growth of trees under environmental constraints (Moëll and Donaldson, 2001) may be addressed by: (1) determining the contribution made to conduction and storage by the different cell types in the ligneous plane; (2) determining the variability of cell characteristics (vascularization elements, fibres, ray cells, vertical parenchyma); and (3) breaking down this variability by identifying that which can be attributed to ontogeny and that which can be attributed to an environmental response.

To interpret and compare internal observations with observations of external morphology, various parts of the plant's architecture must be studied simultaneously (Barthélémy and Caraglio, 2007). In particular, observations should distinguish between the constitutive anatomical features of the plant (i.e. trunk, branch) over time, and generate a sufficiently large sample for the use of statistical tools.

However, these biological questions and their associated characterizations are highly constrained due to cumbersome protocols and the difficulties inherent in acquiring large datasets. Indeed, a precise understanding of growth mechanisms and their modelling has mostly been achieved by searching for invariants in large observation samples, thus restoring intra- and inter-species variability.

Progress made in the preparation of complete cross-sections of axes for ligneous plane observation from a macroscopic to a microscopic level without using histological sections (polishing), combined with modern image acquisition equipment (stage for programmed movements, vibratome, etc.), today provide access to quality information on broad areas of the ligneous plane. This technology, however, is hampered by its limited ability to manage and process data. Acquiring a slice 10 cm in diameter at $\times 200$ magnification requires a mosaic of about 1000 images. On average, each image contains 500 cells and the manual evaluation of each image requires about 40 min using an image editor (ImageJ, Photoshop). Cellular arrangements are identified by studying several successive rings in several growth units, and here the number of images required is too large for manual processing. Usually, counts are made on three files per section (Rossi *et al.*, 2006) to save time and reduce costs.

Automating the study of cell files would make it easier to establish tighter links to the functional and ecological aspects of the species by having a better appreciation of the range of phenomena variability. Wood provides a continuous record of all the developmental changes undergone by a tree (ontogeny, the least known) and a record over time of its environment (dendrochronology and dendroclimatology, the most studied). One particular concern with automated methods is the need to assess and quantify the reliability of their results. This means establishing reliability indicators for the results produced.

The automated identification of cell structures is one of the new challenges to be met in studies of the structural biology of plants (Quelhas *et al.*, 2011), and requires a multidisciplinary approach. In matters of (bio)-imaging, cell segmentation is a problem that has been widely discussed in the literature

(Baggett *et al.*, 2005; Fourcaud *et al.*, 2008). For example, Park and Keller (2001) combined four conventional image analysis approaches to segment cells: watershed algorithm (Vincent and Soille, 1991), snakes (McInerney and Terzopoulos, 1999), multi-resolution analysis (Jeacocke and Lovell, 1994) and dynamic programming. In the specific context of wood, cell segmentation is generally combined with the determination of cell type, i.e. its biological characterization. For example, Wang *et al.* (2008) and Marcuzzo *et al.* (2009) sorted cells using the support vector machines (SVM) classification. This is a supervised incremental training method where the major challenge is to create the training data set that is sufficiently representative to recognize the intrinsic variability of individuals.

In studies of cellular organization, certain authors have used geometric models based on topology. This also provides a description of the cell's surrounding environment. For example, the studies of Jones and Bishof (1996) and Kennel *et al.* (2010) were based on a graph that was orientated so as to extract cell files from images of gymnosperms. More specifically, cells are extracted by applying the watershed algorithm to the converted greyscale image. The adjacency graph of the cells is then built from the basin diagram. Cell typing obtained by the supervised classification method (CART; Breiman *et al.*, 1984) can thus be used to extract tracheid alignments from the graph. This method is sensitive to the reinforcement training used to classify biological types.

From a technical perspective, the searches we conducted failed to find any software solutions suitable for the automatic identification of cell files. Specialized commercially available tools such as WinCell (Hitz *et al.*, 2008) are intended for use in the analysis of wood cells, but do not recognize or characterize cellular organizations. Users have only limited or no possibility of improving their functions. By contrast, commercial platforms such as Visilog (Travis *et al.*, 1996) offer a rich environment, but are not sufficiently specialized, because although it is possible to create macro functions (by assembling and configuring basic ones), it is impossible to add new basic functions.

Open source platforms such as ImageJ (Clair *et al.*, 2007) are an interesting option as their source codes may be enriched by special functions, and specialized macros may be developed. At the same time, they facilitate method sharing and dissemination. It is thus possible to develop new functions for automated processing. We thus preferred this solution for our technical development.

In this paper we describe a cell file detection method that is based solely on cell geometry (size and shape), not on biological type, thus avoiding the bias and uncertainty inherent to the creation of reinforcement learning representing biological variability. Reliability indices are employed to determine the accuracy of the results produced for both the geometrical parameters characterizing the cell and its components (lumen and wall), and the identification of cellular alignments. Results are presented from application of the process to real images of wood slices, obtained by a plugin implementation in ImageJ (Schneider *et al.* 2012), a Java-based image analysis freeware.

MATERIALS AND METHODS

Study species

For the purposes of our study we examined histological sections and blocks of gymnosperms, specifically *Pinus canariensis*,

TABLE 1. Sanded wood discs protocol.

	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Abrasive type	Sic Foil 320	Sic Foil 500	Sic Foil 800	Sic Foil 1200	Sic paper 4000	MD MOL
Disc rotation (r.p.m.)	150	150	150	150	150	150
Rotating sample holder	150	150	150	150	150	150
Direction of rotation	>>	>>	>>	>>	>>	>>
Vertical force (N)	120	120	120	120	120	120
Duration of step (min)	2	2	2	2	2	5
Liquid lubricant	Water	Water	Water	Water	Water	DP blue
Abrasive product						DP spray P3

P. nigra and *Abies alba*. Results were extrapolated to angiosperms by studying histological sections of *Khaya ivorensis*.

Any detailed study of cell typology requires an examination on multiple planes. In our study, we therefore considered only three biological cell types: tracheids, which provide support and are the main cells present in the radial cut; vessels; and rays. For each cell we calculated its circularity, height, wall thickness, and surface area of the cell and its lumen.

Preparation protocol and microscopy

Two preparation methods were evaluated to verify the general applicability of the processing and analytical techniques used.

The conventional preparation method for the study of cell files consists of creating a mosaic of histological sections from wood cores. In practice, this approach does not guarantee the juxtaposition of samples or tissue integrity. It is also relatively time-intensive as it takes approx. 1 h to prepare a mosaic of ten stained histological sections. Thus, this protocol is better suited to the study of relatively smaller areas of observation. The study of larger areas relies instead on the observation and measurements of discs of sanded wood. We thus developed a protocol that combined the sanding and polishing of wood discs to facilitate automated measurements on the entire ligneous plane.

Thin histological cuts. Producing a single cross-section of 20–25 μm thickness using a vibratome takes approx. 5–15 min depending on the case. The sections are then stained to increase the wall/lumen contrast, with the dye binding to the wall. The cleaning and staining steps take 30 min. The full protocol therefore requires approx. 1 h per cut. This method is able to produce histological section mosaics on several rings, depending on core dimensions.

Sanded wood discs. Samples were taken from pine felled in April 2013. Wood discs (about 2.5 cm thick) were cut from the centre of each growth unit. At this stage, it is important, although not always easy, to obtain samples in which the two sides are as parallel as possible. Samples were dried outdoors in the shade, sheltered from the rain, from May to July. A belt sander (Festool BS 105) was then passed over the samples using successively finer grit paper (60, 120 and 220).

The following equipment was used for the pre-polishing and polishing steps: a semi-automatic polisher fitted with a 30-cm-diameter magnetic plate (Struers); abrasive discs (Sic foil or paper depending on the grit used); wool polishing cloth (MD MOL); alcohol-based lubricant (DP blue); and diamond spray (spray DP). The discs were placed in the sample holder

and fixed with double-sided tape. Disc processing is outlined in Table 1.

After the last step, the samples were removed from the holder and rinsed with deionized water. To avoid damage, care was taken not to touch the fragile polished side of the samples. To protect from dust, the samples were wrapped in a laboratory fibre paper. Images were acquired the next day.

Scanning. Scanning, which is performed within 24 h of sample preparation, consists of producing a mosaic of images representing the observation area. The time between sample preparation and scanning must be as short as possible such that the natural drying of the wood does not produce microdeformations affecting the uniformity or ‘flatness’ of the surface, thus compromising the sharpness of the images produced.

In our study, the transverse planes were scanned by an Olympus DP71 LCD camera mounted on an Olympus BX51 microscope equipped with a movable stage, thus allowing, by gradual movements of the slide, the creation of a mosaic of images.

A magnification of $\times 200$ was used to obtain significant cell surfaces while maintaining a large enough study area.

The position of the mosaic image is determined by the x and y movements of the stage. Here, the maximum viewing area was $10 \times 10 \text{ cm}^2$, corresponding to a mosaic of about 1500 images. Each image was defined by 1600×1200 pixels, which were encoded on 3 bytes in the RGB system. The images may be blurred locally or peripherally due to optical distortions or variations in sample thickness.

Image analysis

The image analysis process followed a three-step sequence, starting with cell identification, then cell geometrical characterization and finally file detection. Cell typing and a determination of the reliability of the results obtained are detailed below.

Each image processing step was itself divided into several tasks, as described below.

Cell segmentation. This first step aimed to identify the different cells in the original image. It was divided into three tasks: image noise reduction preprocessing, contrast enhancement by filtering and colour to greyscale conversion, and cell identification from basins generated by the watershed algorithm.

Image noise reduction. At the magnifications used, the images acquired showed ‘salt and pepper’-type noise, caused by heat from the lamp. This noise was attenuated using a median filter applied on a 5×5 -pixel neighbourhood.

Contrast enhancement. The micrographs showed alternating light areas, corresponding to the lumen, and dark areas, corresponding to the wall. To enhance the contrast between different areas in the image and highlight the walls and lumens, we used a Difference of Gaussian (DoG) operator (Haddad and Akansu, 1991). This filter acts as a band-pass, only letting through the frequency range corresponding to the lumen and filtering out the noise present at higher frequencies. The first image was built from a slightly blurred Gaussian filter application with smoothing factor being selected close to average wall size. The second image was built from a heavily blurred Gaussian filter application with smoothing factor tenfold that used for the first image. The heavily blurred image was then subtracted from the slightly blurred image. These operations were performed on colour images, with negative values being set to zero. This process boosts the intensity of the lumen while substantially reducing wall intensities.

From colour to greyscale images. The choice of colour system is an important aspect of colour image processing. Several studies have shown that the ‘best system’ depends on image contents (Busin et al., 2009). For our study, we preferred the RGB system in which colours correspond to the wavelengths that stimulate the three cones of the human eye. This system can define all the hues but not all the saturations.

However, if the watershed algorithm is to be applied, a colour to greyscale image conversion is required, given that our intention was to avoid basing automated processing on human perception. Instead to use the relative luminance formula, which give less weight to blue and more weight to red and green components, we preferred to apply a simple average of the three colour channels.

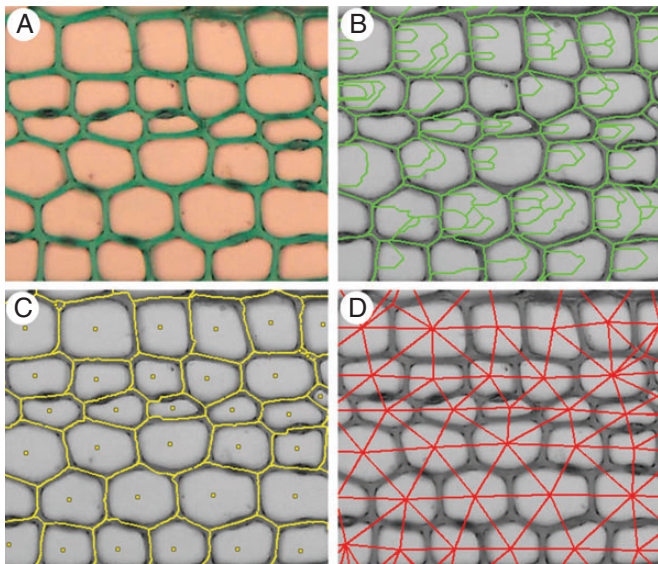


FIG. 1. Cell individualization for a Mahogany cross section. (A) Native image. (B) Watershed segmentation results with crest lines (in green) crossing the lumen. (C) cleaned watershed image with crest lines replaced by curvilinear edges (in yellow). The edges correspond to the middle lamella while the geometric centre of the watershed (extra points in yellow) correspond to the biological cells. (D) Dual watershed image, in which the red lines connect each cell centre with its neighbours (adjacency graph). Note that the degree of connections is greater than 4, showing the staggered pattern of the cells.

Cell extraction and identification. Cell extraction from the grayscale image was based on the watershed algorithm. This powerful algorithm (Vincent and Soille, 1991) considers the image as a landscape, with altitude is given by grey values. Here, the lower points in the topographic relief, corresponding to dark regions, act as catchment basins (as if a drop of water were to fall onto the topography) and are separated by a watershed, represented by the lighter pixels.

These catchment basins are thus obtained from a partition of the image. The crest lines constitute the intercellular border, corresponding to the middle lamella (Fig. 1C). The watershed algorithm is known to produce an exaggerated partitioning, resulting in lines that *abnormally* intersect cells. This phenomenon is due to the fact that each local minimum in the image produces a potential catchment basin. Typically, two methods are used to reduce or even eliminate the exaggerated partitioning: *a priori* filtering of the local minima values (Gilmore and Kelley, 1995) used to initialize the watershed, or *a posteriori* merging of similar adjacent basins in terms of both mean and minimal intensity (Beucher, 2012). Neither method guarantees perfect cleaning of the watershed lines. Here, we propose a method based on the biological observation that the lumen of the cell is empty. Hence, if a watershed line crossing a lumen shows a characteristic intensity profile whose overall maximum is greater than a threshold experimentally set from average lumen values, then it will be deleted. Conversely, a watershed line near the middle lamella contains only values below the low-intensity threshold (Brunel et al., 2012).

After cleaning the watershed. An adjacency graph of the basins was used to describe the neighbouring relationships between the different basins. This is a simple graph conventionally defined by a set of vertices and a set of edges. More precisely, each edge of the graph connects the geometric centres of two neighbouring basins, i.e. incidents at the same crest line.

Geometric properties at the cell scale. Each cell was then individualized. Cells are composed of two anatomical structures: the wall and the lumen. On the original colour image, these structures are automatically divided into two classes by an unsupervised two-means classification algorithm (Forgy, 1965), also equivalent to Otsu’s thresholding operator. The bright pixels class corresponds to the lumen and the dark pixels class to the wall. Surface areas are calculated simply by counting the pixels in each class.

Cell geometrical properties at larger scales: the global cell arrangement orientation. When considering a basin in the adjacency graph, it can be seen that the cells are arranged in a staggered manner, and the edges of the graph are orientated in three directions. At the full graph scale, the most frequent orientation corresponds to the cellular alignments. This is defined as the main direction. We will use the method originally described by Jones and Bishof (1996) and reused by Brunel et al. (2012), studying the distribution (via a histogram) of the angle that each edge of the adjacency graph forms with the horizontal. In theory, due to the arrangement of the cells, the distribution should be trimodal and show at least one marked amplitude mode. The mode of the greatest amplitude indicates the most represented orientation in the adjacency graph, i.e. the main

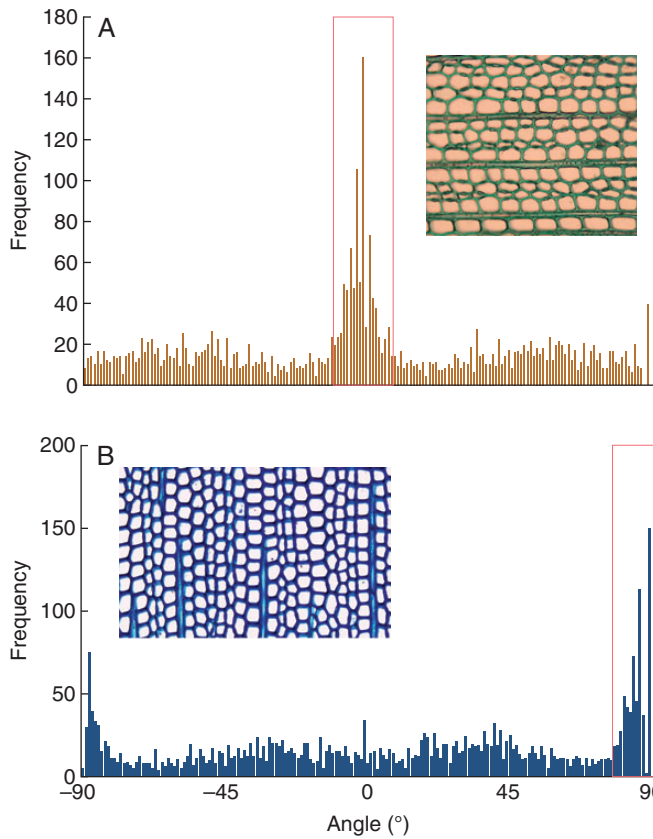


FIG. 2. Histograms showing the angle distribution on the full adjacency graph; each angle is defined by an adjacency graph edge direction on the x axis. (A) The main mode centred on 0° corresponds to horizontal files. (B) The main mode is split on the $[-90^\circ, +90^\circ]$ interval borders.

orientation of the files (Fig. 2). This mode is retrieved by searching the histogram maximum amplitude.

File detection

The process used to identify cell files in an image is based on the following assumptions: (1) files are pair-wise alignments of similar cells (in terms of size and shape) and (2) cell alignments are independent of the orientation of the image.

File identification is based on a double scale constructive approach, applied to the adjacency graph. The first step builds cell alignments, i.e. candidate cell files under spatial and geometric constraints. The second step manages the case of alignment recoveries. Finally, the resulting isolated alignments are linked to build cell file fusions.

(1) *Building alignments by applying spatial and geometric constraints.* This task extracts the longer rectilinear straight paths of ‘geometrically similar’ vertices from the adjacency graph. In other words, it detects a vertex sequence along which the surface of the underlying basin varies gradually. These alignments are constructed step by step, by successive aggregations of vertices. More specifically, it means finding which vertex w of the graph must be added at the end of the v line under construction to complete it.

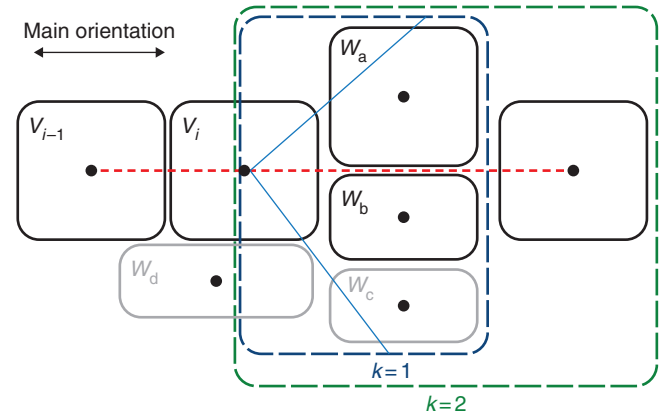


FIG. 3. The angulation (direction) constraint. The red dashed line corresponds to the file axis. The vertices (v_{i-1}, v_i) define the current alignment. The blue and green dashed line boxes show the study neighbourhood at different ranks. Cells w_a and w_b are the candidates located at a geodesic distance of 1. Vertex w_d is not considered because it is above the angular threshold (blue line) and vertex w_c is not considered because vertex w_b is closer to the file axis.

The angulation (direction) constraint. Let us define the vertex neighborhood $N_k(v_i)$ of rank k as the set of vertices located at a geodesic distance k from vertex v_i . Let us then define the angulation $A(v_i, w)$ as the angle defined from the complement to 180° of (v_{i-1}, v_i, w) where v_i and v_{i-1} are the last two vertices of the built alignment and w is the candidate vertex. The angulation can be used to choose the best candidate w that minimizes the angular offset from the reference direction, i.e. the current cell alignment. In no cases may this angle be greater than a threshold which is given by a decreasing linear function (to 35° at 10°) that itself is dependent upon the geodesic distance. For the first two file cells the angulation value is defining according to the main orientation. The vertex solution w is selected from among the w_k vertices of the $N_k(v_i)$ set and is the vertex most geometrically similar and nearest to the file axis. In fact, for a given geodesic distance k , at most two vertex candidates are considered, i.e. the two closest vertices (w_a and w_b) lying on either side of the axis formed by the last two geometric centres of the cells of the file under construction (Fig. 3).

The geometrical similarity constraint. The candidate most geometrically similar to vertex v_i is retained. The geometrical comparison is performed based on their mutual underlying basins. Let GS_{pq} be the geometric similarity index between two vertices p and q that is given by the normalized difference between the surface areas S_p and S_q

$$GS_{pq} = \frac{|S_p - S_q|}{(S_p + S_q)} \quad (1)$$

The more the index tends toward 0, the more similar the cell surface areas. This index is particularly well suited to softwoods in which the shape and surface area of the tracheid varies very little. For angiosperms, the index is a good indicator of rough size ruptures while permitting continuous and progressive surface area variations.

In practice, if the smallest GS_{pq} value is lower than 0.5, the candidate is assigned to the current alignment. If the geometric similarity index of the vertices is greater than 0.5, the candidate is

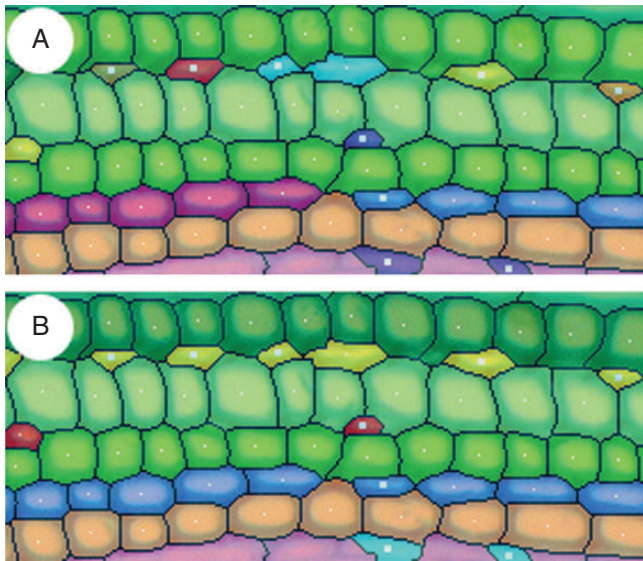


FIG. 4. Linking isolated sections. (A) Result of linkage processing without connection. (B) Result with connection. One colour is used per file.

rejected, the geodesic distance k is incremented and the process is repeated. The neighbourhood explored is progressively enlarged to a geodesic width of 5, at which point the construction of the alignment stops. This neighbourhood width was an experimental limit set arbitrarily by the wood anatomists involved in the project.

Before validating the alignment as a cell file section, a feedback control test is applied as the construction process should be reversible.

This feedback uses the main orientation of the file as a reference to reduce file drift. When a vertex w is chosen, a check is performed that the reciprocal is true, i.e. that we obtain current vertex v when falling back from vertex w using the main orientation of the files as a directional reference. If reciprocity is not confirmed, the construction of the file stops. If it is confirmed, the vertices located on the shortest path linking v to w are added to the cell alignment by the Dijkstra algorithm (Cormen et al., 2001). The criteria employed are described in the next section. Finally, vertex w becomes the new extremity of the alignment and the search process is repeated until the alignment construction stops.

Once stopped, the detected alignment builds a candidate cell file. The two next steps (recovery management and fusion) then concern the cell file level.

(2) *Managing recoveries.* During construction, some vertices may be assigned to multiple files due to the presence of biological organizations (vessels and channels) that disrupt the staggered arrangement. The assignment of these vertices to the most optimal file is solved by the shortest path search algorithm in the adjacency graph. The Euclidean distance between the geometrical centres of the watershed basin is used as a criterion. This distance sums the angle and shape considered in file building. The greater the angular deviation, the greater the Euclidean distance. The greater the size difference between successive cells, the greater the Euclidean distance. The shortest path in terms of Euclidean distance corresponds to the best alignment. As above, we use the Dijkstra algorithm here, and the

corresponding vertices are then removed from the other files, which will be fractionated.

(3) *Linking isolated cell file sections.* Due to the presence of intrusions, cell tears or limitations of the watershed algorithm, the file identification process may cause isolated file sections to appear (Fig. 4). Here, merging can be used to concentrate several parts of the files using simple topological rules to establish whole files, based on file adjacency (Brunel et al., 2012). Let us assume that two files are adjacent when two of their cells are adjacent.

The merging process is then based on two further assumptions: (1) each file crosses the image from one side to the other and (2) the files do not intersect. A file section does not fulfil the first assumption, i.e. one or both tip(s) do(es) not match with the underlying image border.

From those assumptions, we deduce a property resulting from file adjacency.

Given two file sections $T2$ and $T2'$ adjacent to files $F1$ and $F3$, it is highly likely that $T2$ and $T2'$ are two parts of a single $F2$ file. In other words, $T2$ and $T2'$ are located between $F1$ and $F2$.

Note that, under the exposed assumptions, restricted cases of file merging concern only sections between two lines crossing the image from side to side. In other cases, it was decided to leave the situation at the expert's discretion and evaluation.

This method creates files or sections of files with high confidence (Fig. 5, right). Note that isolated cells or artefacts (due to segmentation errors or to poor image resolution preventing an acceptable segmentation) cannot be part of any file and no attempt is made to connect them to a file. Complex cases are left to the discretion of the user.

Typing

Cell typing is the ultimate step as used to classify the different cells present (fibrous, tracheids, vessels, rays, etc.). It is based on the geometric and densitometric characterization of the catchment basins associated with the cells' adjacency graph vertices. Unlike Kennel et al. (2010), we decided to discard supervised classification due to the difficulties encountered in providing reinforcement training that is both sufficiently complete and representative of biological variability.

A decision tree (Fig. 6) was established with wood anatomists and is based on the following observations that are sequentially tested:

- (1) The perimeter of the cells, denoted P , can be used to differentiate 'large' structures from generally 'small' cells.
- (2) The circularity coefficient (Zunic and Hirota, 2008) is given by a weighted ratio between the surface area and the squared perimeter -4π (area/perimeter²). Circularity ranges between 0.0 and 1.0: the circularity of a circle is 1, and is far less than for a starfish footprint. Values may not be valid for very small particles due to the definition of perimeter and area in a square grid.

The thresholds used in the decision tree are re-evaluated for each image by automated classification of their numerical values. The classifier used is an implementation of two-means clustering. This conventional data mining method consists of

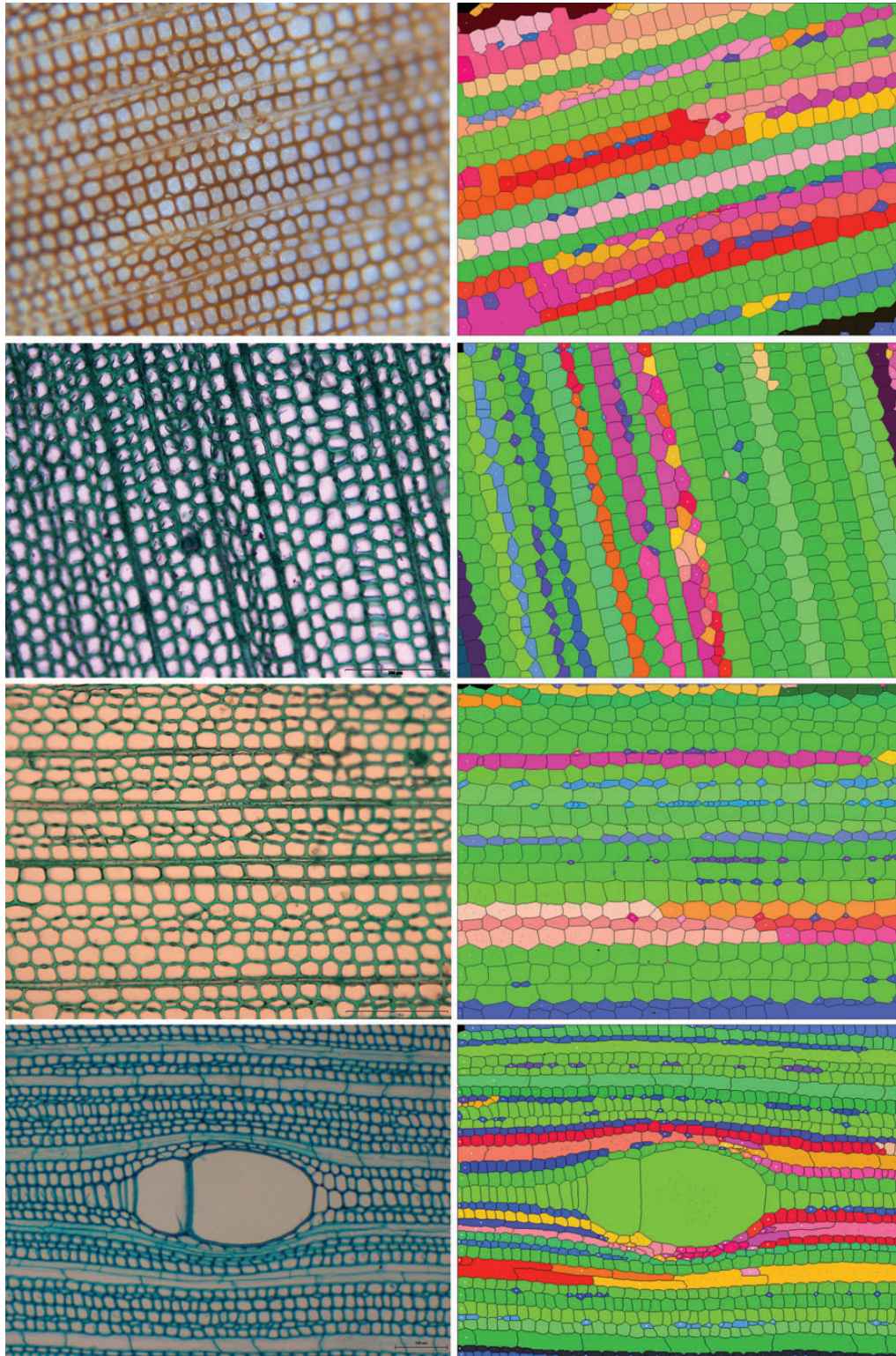


FIG. 5. Left. The native image; the images show different the preparations (sanded wood and stained histology sections), species (*Pinus*, fir, *Pinus*, mahogany) and clades (gymnosperm and angiosperm) that were processed. Right. Automatic identification of cell files; random colouring is used to enhance the visibility of the identified files (these colours should not be confused with the reliability index colour).

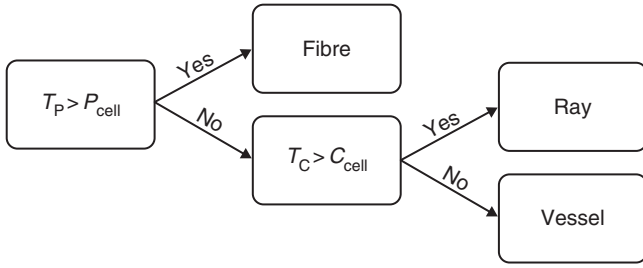


FIG. 6. Decision Tree. T_p corresponds to the threshold of the perimeter and T_c the threshold of circularity. The thresholds are automatically rated by a two-means clustering applied to all values from the image. Only cells with a lumen are processed by the decision tree.

allocating n observations in two clusters to minimize the intra- and maximize inter-class standard deviation. The threshold is therefore assigned by the median value between the upper bound of the weakest group and the lower bound of the strongest group.

Beyond the identification of alignments and biological typing, the anatomist requires a set of numerical results characterizing the shape, size and nature of more or less complex biological structures (walls, lumen, cells, files, etc.). These characteristics are defined by parameters that are automatically evaluated. In the context of mass processing, it is advantageous to be able to qualify the accuracy of these evaluations. A certainty index is thus assigned to each one calculated, whether it focuses on the cells, their components or their alignments.

Reliability

Intuitively, the computed measures will be more or less close to the true value depending on the cumulative errors stemming from image quality, biological configuration and algorithmic approximations. It is therefore important to assign each result a reliability estimation.

A file f must be sufficiently long to be significant; its length L_f must be greater than threshold L . L is defined from the length L_f distribution analysis. For each image the threshold is re-evaluated to reflect the biological characteristics inherent to each of them as a file must be composed of similar cells in terms of shape, size and appearance. For the sake of simplicity, these three criteria are reduced to the single height of cells defined from all points in the normal direction of the file. This simplification stems from the maturation mechanism of tracheids, for which cell variations only concern cell extension, not thickening. Clearly, this mechanism cannot be applied to broadleaved trees, but in practice this single criterion appears to be sufficient to characterize the continuous variation observed in the cell files of broadleaved trees. As a result, the heights H and H_{j+1} of consecutive cells in file f are compared. The reliability coefficient R_f of file f is described by a product of standardized terms ranging from 0 to 1.

$$R_f = \left(1 - \max\left(\frac{L - L_f}{L}, 0\right)\right) \prod_{j=0}^{n-1} \left(1 - \frac{|H_j - H_{j+1}|}{H_j + H_{j+1}}\right) \quad (2)$$

The closer the coefficient to 1, the more reliable file f .

By analogy, we can determine reliability coefficients for the geometrical parameters that characterize file cells. The expression for this coefficient will depend on the nature of each parameter and more specifically on the factors influencing that parameter. The surface area of the basin is independent of local image blur because it is obtained by watershed segmentation, while the surface area of the lumen is highly sensitive to local image blur because it is obtained by an intensity classification. The reliability coefficient for the lumen is therefore directly related to blur level. [Ladjal \(2006\)](#) defined a blur indicator that characterizes the spread of maximum intensity amplitudes, i.e. the speed at which the signal passes from the lowest to the highest intensity. This is defined by the following expression where $p(x)$ shows the pixel intensity of a basin and $p'(x)$ the variance.

$$\mathfrak{S}(p) = \frac{\max_{x \in \mathbb{R}} p(x) - \min_{x \in \mathbb{R}} p(x)}{\max_{x \in \mathbb{R}} |p'(x)|} = \frac{\text{amplitude}(I)}{\max(|\nabla I|)} \quad (3)$$

When the coefficient is high, blur is strong. The reliability coefficient for lumen surface area (I_{ls}) is defined below where threshold T_s is defined by studying lumen variation using a Gaussian blur. Surface areas above the threshold are not significant.

$$I_{ls} = \begin{cases} 0 & \text{if } \mathfrak{S}(p) > T_s \\ 1 - \frac{\mathfrak{S}(p)}{T_s} & \text{else} \end{cases} \quad (4)$$

Note that the accuracy of the surface areas and circularities are sensitive to the size of the objects measured; for example, the circularity of a discrete circle ([Andrès, 1994](#)) is meaningful only above 10 pixel rays.

Parameter computation

Our method produces several layers of results corresponding to different observation levels: (1) radial files are classified according to their length and their cellular self-similarity, (2) cells are characterized by geometric size, diameter, shape, etc., and (3) cell components (wall, lumen) are characterized by geometric parameters (size, thickness, diameter, etc.).

The method has been implemented in Java and integrated as a plug-in for ImageJ freeware. As an indication, the plug-in allows images consisting of 1600×1200 pixels to be processed in less than 20 s on a computer with an Intel Q720 processor clocked at 1.6 GHz.

RESULTS AND DISCUSSION

The study described herein aimed to compare the results produced by our method with those in an expert database containing 12 images whose cells had been manually cropped by wood anatomists. Our results were obtained without any specific user settings (Fig. 5): all parameters were automatically re-evaluated by the application based on the intensity histogram generated for each image. The different tests conducted allowed us to evaluate the limitations of our method, namely with regard to different biological contexts, different preparation protocols, and diverse and varied acquisition conditions. The tests were essentially

TABLE 2. Outline of some significant results: size of processed images, number of extracted cells, CPU time on a PC with an Intel Xeon at 2.3 GHz, and total Quality Index defined as the ratio of number of files automatically identified and reconstructed over total number of lines; for the purposes of the expert analysis, the cells were manually cropped using ImageJ software

Species	Size (pixels)	Expert		Plugin		% of files detected
		Cell numbers	Times (s)	Cells numbers	Times (s)	
Mahogany	1024 × 768	1302	2520	1359	14.3	86
Fir	1360 × 1024	787	1300	800	12.4	88
Black Pine 1	1600 × 1200	1794	2750	1873	23.2	92
Caribbean Pine	1360 × 1024	819	1500	828	11.5	91
Black Pine 2	1600 × 1200	1411	2450	1458	16.1	83

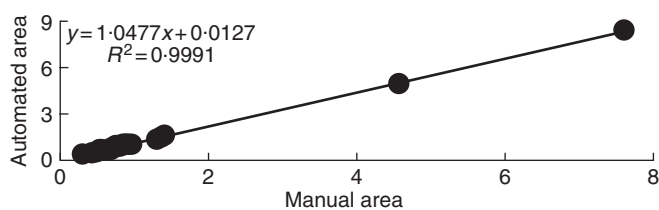


FIG. 7. Study of basin surfaces on mahogany. The x-axis shows the normalized areas obtained by the fully manual method. The y-axis shows those obtained by the automated method. The coefficient of determination is close to 1, showing a good fit.

motivated by simple questions: (1) is our method suitable for both angiosperms and gymnosperms? and (2) is our method sensitive to preparation protocols, to optical blur, to luminosity levels, etc.?

Qualitative performance

At this stage in our studies, the performance of our method was subject to a qualitative assessment by comparing – in a fairly global manner – the results produced with those of an expert appraisal. Of course, it is always possible to compare ‘processing’ time, number of cells detected or files identified (Table 2). For example, Table 2 shows that the automatic method detects a consistent number of cells about 80–100-fold more rapidly than the expert, and this regardless of preparation protocol, anatomical configuration or magnification used, and that, overall, our method works well (88 % of files detected on average) on images presenting a structured organization with a marked wall/lumen contrast.

Geometric precision

We have seen that files are created only from geometric and topological rules applied to the basin adjacency graph. It is therefore important to ensure the robustness of the basin detection, i.e. the method used must be precise in cell cropping and insensitive to image blurring.

Figure 7 compares about 60 normalized areas obtained by the manual and automated methods. The areas were divided by the median area to refine the regression. The coefficient of determination tends toward 1, showing that the areas are closely correlated. The slope of the regression is weakly greater than 1, indicating slight over-estimation by the automated method.

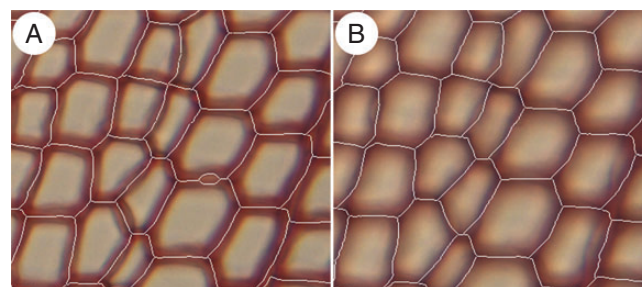


FIG. 8. Comparison of watershed results on a *Pinus* cross-section. (A) Crest lines from a sharp image. (B) Crest lines from a blurred image. Basin outlines are virtually identical between the sharp and blurred images. The sharp image has a small additional basin corresponding to wall detachment caused by cutting. This basin was deleted in the supernumerary removal phase.

The low shift of 0.0127 confirms the middle error of 5 % on right areas. The automated method appeared to over-estimate the values (or the expert appraisal under-estimated the values). The only certainty is that the automated method is repeatable.

The major limitation to the automatic file identification method stems from image content, namely the photometric characteristics of the image and the structural biological layouts.

On file identification, only geometric accuracy is studied. The geometric shift is deliberately not taken into account as it is considered to be negligible with respect to cell surface area. However, the study could feasibly be extended to geometric shift, namely to assess the geometric overlap between expert annotations and segmentation algorithms.

A study was conducted on the effect of optical blur, which is often present in images and is caused by peripheral deformations due to the microscope lens or sample flatness. The algorithmic stability of our method was evaluated by acquisitions of the same image area with different adjustments. A statistical evaluation of cell cropping variability was then performed. We did highlight the fact that the intercellular lines obtained by the watershed algorithm are generally unchanged by image blurring. They correspond, in fact, to gradient slope inversions: the image smoothing produced by the optical blur reduces the intensity dynamic of the image without removing the changes in intensity variations. Overall, the crest lines remain unchanged (Fig. 8).

Figure 9 compares 60 areas obtained from a sharp and a blurred image. The coefficient of determination tends toward 1 showing that the areas are closely correlated. The slope of the regression is slightly less than 1, indicating slight under-estimation for blurred

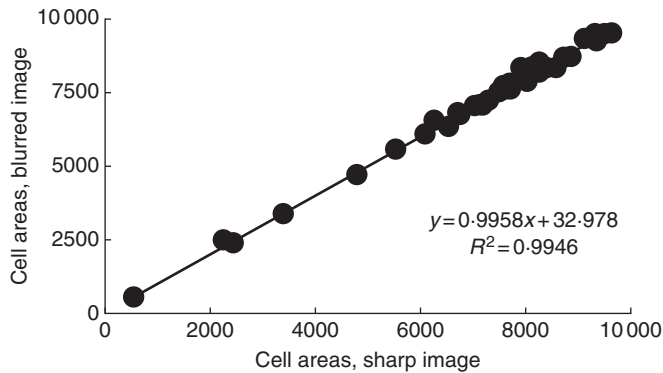


FIG. 9. Basin surface areas study on *Pinus*, with cell areas expressed as pixels. The *x*-axis shows the areas estimated from sharp images. The *y*-axis shows those estimated from simulated blurred images. The slope of the regression is slightly less than 1, indicating minor under-estimation of basin areas in the blurred images.

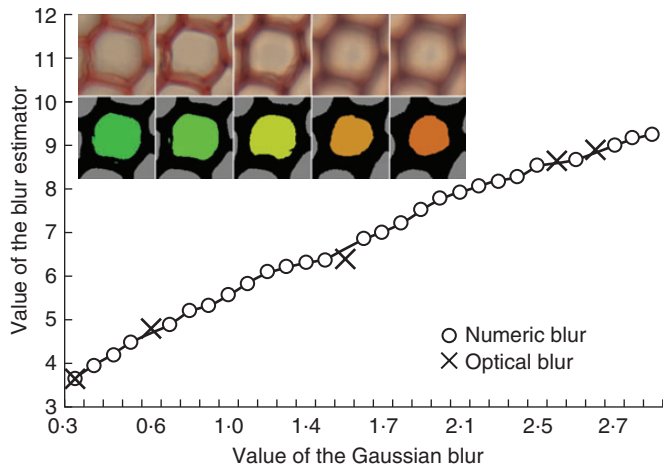


FIG. 10. The behaviour estimator shows the blur tendency and the red crosses correspond to the five optical microscopy blurs. The five lumen areas are shown with their corresponding optical blurred snapshots. The colour indicates the reliability coefficient, from green for true lumen area to orange for reduced lumen area.

images. The method therefore yielded almost identical results for both sharp and blurred images.

Reliability and blur

Image blurring is a source of potentially major bias when evaluating lumen areas and wall widths. As indicated before, the cell segmentation is based on a two-means clustering. This classification is sensitive to the intensity dynamics of the cellular points and is therefore modified if intensity dynamics are reduced by optical blur (Fig. 10). We observed that the surface area of the lumen for a given cell was reduced when blur increased, i.e. when intensity dynamics decreased.

The local blur indicator introduced earlier is used to estimate the error on the surface area measurements of cell constituents. The figure on the left shows the consistent behaviour of the blur

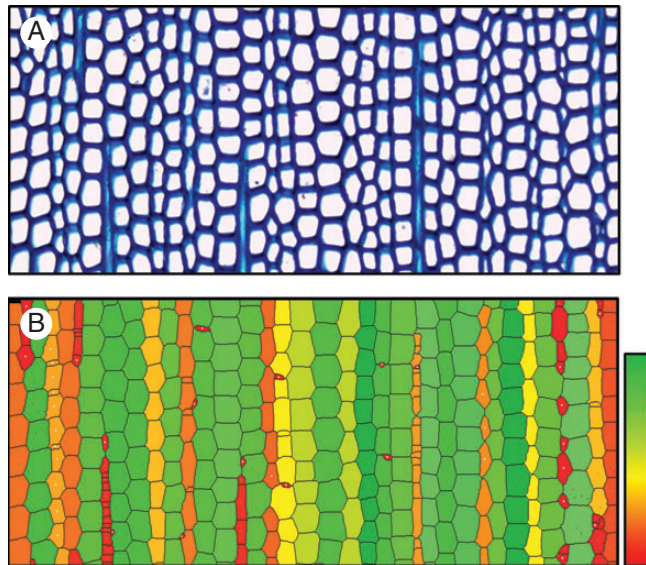


FIG. 11. (A) Cross-section of native *Abies*. (B) The automated identification of radial files. The colour value indicates the reliability coefficient: poorly reliable files appear in pale red while very reliable files appear in pale green.

estimator that was calculated from cells considered in images that were digitally disrupted by an incremental Gaussian blur.

Therefore, the local blur indicator can be used to filter the results and retain only the sharpest cells. It is used, in particular, to create mosaic images (Brunel et al., 2013), for retention of the most reliable basins. It could also possibly provide a correction factor for measures made on areas that are more or less subjected to optical blur.

Reliability of files

The overall reliability coefficient can be used for file characterization of files, in terms of both significant expression and geometric behaviour, i.e. a ‘good’ file is a sufficiently long line in which cell size varies gradually. This coefficient acts directly to filter the files (Fig. 11). Files with a good coefficient are shown in green while those with a poor coefficient are in red. The red–orange hues indicate configurations with cuts or sharp discontinuities. This estimator is well suited to softwoods with very organized and continuous cell arrangements. In angiosperms, it may be used for a rapid visualization of vessels and adjacent files. These files are useful for analysing the endogenous geometric constraints caused by vessels on files.

This indicator is ultimately a very integrated aspect of the files identified. It is more comprehensive than the performance indices mentioned in the first part of the results, and is also more complex to explain as it goes beyond file identification reliability to actually reflect file quality. For example, some alignments, shown in yellow in the image below, appear to be ‘visually correct’ while their coefficient is rather low. They are penalized by excessively elevated surface area variations. For example, the third orange file from the left looks to be good, but small basins can be seen within the file arising from over-segmentation caused by wall detachment. Files can be filtered for statistical analysis thanks to their cell coefficients.

General conclusions

The method described herein works correctly for images that show both high contrast between walls and lumens (without local color inversion) and a visible cellular organization. Under these conditions, all the obvious lines determined by qualified experts are correctly identified with significant time savings (a typical manually expert identification required 10 h on our samples). Moreover, the certainty index allows for selective exploitation of the results for statistical studies.

CONCLUSIONS

The automated analysis of anatomical wood sections will contribute greatly to understanding secondary plant growth and development. Herein, we propose a fully automatic method for the recognition of operational cell files in mass processing.

We have also established reliability coefficients to characterize the files, their components and their measurement. These reliability coefficients are mainly used to filter results for statistical studies conducted by botanists and aim to draw their attention to specific biological configurations.

We have shown that the method is equally applicable to digital images of stained histological wood sections and images of sanded ligneous planes.

File identification is based on detection of cell alignments showing gradual geometric variations. The cells in the image are automatically cropped by watershed lines calculated from a set of judiciously filtered local minima, thus reducing the risk of exaggerated segmentation. Cell files are defined by the longest straight vertices that are 'geometrically similar' to the adjacency graph created from the watershed catchment basins. These alignments are constructed one after the other by successive aggregations of vertices in progressively enlarged neighbourhoods. Unlike traditional approaches, cell typing here is not based on the creation of reinforcement training, but on a decision tree established with wood anatomists.

In addition to file detection, the cells are characterized by a set of values (wall thickness, lumen surface, cell circularity, etc.) that are necessary for the statistical analysis of files and their variability. To ensure that the measures are relevant, reliability indices have been created, based on existing methods and biological reality (file length and continuity). These indices are used to select the most important files or highlight particular biological disturbances or configurations.

The percentage of unusable files is offset by the quantity of images that can be processed. Mass processing allows us to study spinal files up to the cambium over several wood growth units.

These initial results are already very interesting and confirm that the method has great potential and could be suitable for statistical studies and the detection of invariants or trends in large observation samples. Comparative studies undertaken with images annotated by experts have shown that the proposed method is time-efficient and generally insensitive to the optical blur potentially present in micrographs. A blur indicator has been introduced to study the stability of the method and in the short term could be used to correct the measurements obtained. Finally, a reliability indicator to characterize files, in terms of both significant expression and geometric behaviour, is given either to draw

the attention of experts to specific biological configurations or to remove potentially erroneous data from statistical studies.

ACKNOWLEDGEMENTS

We thank the NUMEV Laboratory of Excellence (University of Montpellier 2) and the SIBAGHE Graduate School of the University of Montpellier 2 for their support. We extend special thanks to Christine Heinz and Christophe Godin for their availability and the relevant and constructive observations they provided during different interviews and meetings. Additionally, a big thank you to François Pailler and Merlin Ramel for their patience and help in preparing the different samples and optimizing technical protocols. Special thanks also to Michaël Guéroult who took the time to provide assistance and share his skills and knowledge in the use of optical microscopy.

LITERATURE CITED

- Andrès E. 1994. Discrete circles, rings and spheres. *Computers and Graphics* **18**: 695–706.
- Baggett D, Nakaya M, McAuliffe M, Yamaguchi TP, Lockett S. 2005. Whole cell segmentation in solid tissue sections. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology* **67**: 137–143.
- Barthélémy D, Caraglio Y. 2007. Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany* **99**: 375–407.
- Beucher S. 2012. *Towards a unification of waterfalls, standard and P algorithms*. CMM/Mines Paristech publication.
- Breiman L, Friedman J, Stone CJ, Olshen RA. 1984. *Classification and regression trees*. London: Chapman and Hall/CRC.
- Brunel G, Borianne P, Subsol G, Jaeger M, Caraglio Y. 2012. Automatic characterization of the cell organization in light microscopic images of wood: Application to the identification of the cell file. In: Guo Y, Kang MZ, Dumont Y, eds. *Plant growth modeling, simulation, visualization and applications. The Fourth International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications*, Shanghai, China, 58–65.
- Brunel G, Borianne P, Subsol G, Jaeger M. 2013. Simple-graphs fusion in image mosaic: application to automated cell files identification in wood slices. In: Kämäräinen J-K, Koskela M, eds. *Lecture notes in computer science. Image analysis*. Berlin: Springer, 34–43.
- Busin L, Shi J, Vandenbroucke N, Macaire L. 2009. Color space selection for color image segmentation by spectral clustering. *IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 262–267. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=5478603&queryText%3DColor+space+selection+for+color+image+segmentation+by+spectral+clustering>.
- Clair B, Gril J, Di Renzo F, Yamamoto H, Quignard F. 2007. Characterization of a gel in the cell wall to elucidate the paradoxical shrinkage of tension wood. *Biomacromolecules* **9**: 494–498.
- Cormen TH, Cazin X, Kocher G-L. 2001. *AD. Introduction to algorithms*. Cambridge, MA: MIT Press.
- Forgy E. 1965. Cluster analysis of multivariate data: efficiency vs interpretability of classifications. *Biometrics* **21**: 768–769.
- Fourcaud T, Zhang X, Stokes A, Lambers H, Körner C. 2008. Plant growth modelling and applications: the increasing importance of plant architecture in growth models. *Annals of Botany* **101**: 1053–1063.
- Gilmore P, Kelley C.T. 1995. An implicit filtering algorithm for optimization of functions with many local minima. *SIAM Journal* **5**: 269–285.
- Gindl W. 2001. Cell-wall lignin content related to tracheid dimensions in drought-sensitive Austrian pine (*Pinus nigra*). *Iawa Journal* **22**: 113–120.
- Guédon Y, Taugourdeau O, Caraglio Y, Sabatier S. 2013. Characterizing the balance between ontogeny and environmental constraints in forest tree development using growth phase duration distributions. In: Nygren P, Lintunen A, Godin C, Nikinmaa E, Sievänen R, eds. *Proceedings of the 7th International Conference on Functional Structural Plant Models*, Saariselka, Finland, 237–239.

- Haddad RA, Akansu AN. 1991.** A class of fast Gaussian binomial filters for speech and image processing. *IEEE Transactions on Signal Processing* **39**: 723–727.
- Heinrich L. 2007.** Reaction wood varieties caused by different experimental treatments. TRACE, Tree Rings in Archaeology. *Climatology and Ecology* **5**: 224–232.
- Hitz OM, Gärtner H, Heinrich I, Monbaron M. 2008.** Wood anatomical changes in roots of European ash (*Fraxinus excelsior* L.) after exposure. *Dendrochronologia* **25**: 145–152.
- Jeacocke MB, Lovell BC. 1994.** A multi-resolution algorithm for cytological image segmentation. *Proceedings of the 1994 Second Australian and New Zealand Conference on Intelligent Information Systems, 1994.* IEEE, 322–326.
- Jones R, Bischof L. 1996.** A graph-based segmentation of wood micrographs. Computing science and statistics 28. *Proceedings in Computing Science and Statistics* **28**: 12–20.
- Kennel P, Subsol G, Guérout M, Borianne P. 2010.** Automatic identification of cell files in light microscopic images of conifer wood. In: Djemal K, Deriche M, eds. *2nd International Conference on Image Processing Theory, Tools and Applications.* IEEE, 98–103.
- Krause C, Rossi S, Thibeault-Martel M, Plourde P-Y. 2010.** Relationships of climate and cell features in stems and roots of black spruce and balsam fir. *Annals of Forest Science* **67**: 402–402.
- Lachaud S, Catesson A-M, Bonnemain J-L. 1999.** Structure and functions of the vascular cambium. *Comptes Rendus de l'Académie des Sciences – Series III. Sciences de la Vie* **322**: 633–650.
- Ladjal S. 2006.** Estimation du flou dans les images naturelles. In: *Quinzième congrès francophone Reconnaissance des formes et intelligence artificielle (RFIA).* François-Rabelais University Press (in French).
- Marcuzzo M, Quelhas P, Campilho A, Mendonça AM, Campilho A. 2009.** Automated Arabidopsis plant root cell segmentation based on SVM classification and region merging. *Computers in Biology and Medicine* **39**: 785–793.
- McInerney T, Terzopoulos D. 1999.** T-Snakes: topology adaptive snakes. *Medical Image Analysis* **4**: 840–845.
- Moëll MK, Donaldson LA. 2001.** Comparison of segmentation method for digital image analysis of confocal microscope images to measure tracheid cell dimensions. *IAWA Journal* **22**: 267–288.
- Nicolini E, Caraglio Y, Pélissier R, Leroy C, Roggy J. 2003.** Epicormic branches: a growth indicator for the tropical forest tree, *Dicorynia guianensis* Amshoff (Caesalpiniaceae). *Annals of Botany* **92**: 97–105.
- Park J, Keller JM. 2001.** Snakes on the watershed. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**: 1201–1205.
- Quelhas P, Nieuwland J, Dewitte W, Mendonça AM, Murray J, Campilho A. 2011.** *Arabidopsis thaliana* automatic cell file detection and cell length estimation. In: Kamel M, Campilho A, eds. *Image analysis and recognition.* Berlin: Springer, 1–11.
- Rathgeber CBK, Rossi S, Bontemps J-D. 2011.** Cambial activity related to tree size in a mature silver-fir plantation. *Annals of Botany* **108**: 429–438.
- Rossi S, Deslauriers A, Anfodillo T, et al. 2006.** Conifers in cold environments synchronize maximum growth rate of tree-ring formation with day length. *New Phytologist* **170**: 301–310.
- Rossi S, Simard S, Rathgeber CBK, Deslauriers A, Zan CD. 2009.** Effects of a 20-day-long dry period on cambial and apical meristem growth in *Abies balsamea* seedlings. *Trees* **23**: 85–93.
- Rozenberg P, Schüte G, Ivkovich M, Bastien C, Bastien J-C. 2004.** Clonal variation of indirect cambium reaction to within-growing season temperature changes in Douglas-fir. *Forestry* **77**: 257–268.
- Schneider CA, Rasbaand WS, Eliceiri KW. 2012.** NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**: 671–675.
- Thibeault-Martel M, Krause C, Morin H, Rossi S. 2008.** Cambial activity and intra-annual xylem formation in roots and stems of *Abies balsamea* and *Picea mariana*. *Annals of Botany* **102**: 667–674.
- Travis AJ, Hirst DJ, Chesson A. 1996.** Automatic classification of plant cells according to tissue type using anatomical features obtained by the distance transform. *Annals of Botany* **78**: 325–331.
- Vincent L, Soille P. 1991.** Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**: 583–598.
- Wang M, Zhou X, Li F, Huckins J, King RW, Wong STC. 2008.** Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. *Bioinformatics* **24**: 94–101.
- Zunic J, Hirota K. 2008.** Measuring shape circularity. *Progress in Pattern Recognition, Image Analysis and Applications Lecture Notes. Computer Science* **5197**: 94–101.