Defining reliability coefficients in an automated method of identification and characterization of radial files in microscopic images of gymnosperms

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

Proceedings of the 7th International Conference on Functional-Structural Plant Models, Saariselkä, Finland, 9 - 14 June 2013. Eds. Risto Sievänen, Eero Nikinmaa, Christophe Godin, Anna Lintunen & Pekka Nygren. http://www.metal.fi/fspm2013/proceedings. ISBN 978-951-651-408-9.

Defining reliability coefficients in an automated radial file identification and characterization method in microscopic images of gymnosperms

Guilhem Brunel[1,2], **Philippe Borianne**[2], **Gérard Subsol**[3], **Marc Jaeger**[2] and **Yves Caraglio**[2]
[1]*CIRAD - UMR AMAP, France.* [2]*Université Montpellier 2, France.* [3]*CNRS – LIRMM, France.*
guilhem.brunel@cirad.fr

Highlights: The analysis of wood anatomical sections is of great interest for understanding the growth and development of plants. We propose a novel method for automatically identifying and characterizing radial files in wood microscopic images of gymnosperms. A key point is to be able to assign *a priori* reliability coefficient to the results, particularly for statistical processing in large-scale analyses. We describe in this paper the principle used to establish reliability coefficients to evaluate the radial file identification process and the geometrical measurements of cells and their components.

Keywords: image processing, wood microscopic images, radial file identification, *a priori* reliability.

INTRODUCTION

Wood structure reflects the physiological and molecular regulation of cambium activity (Cato et al. 2006) and also registers environmental conditions (Barlow et al. 2005). Understanding cambium growth mechanisms calls for a study of cell pattern regularity, and of cell disruption or modification in space and time (Liang et al. 1997). Two types of organizations are considered: the growth ring representing cell production at a given time, and the radial file representing the activity of an initial cell over time (Rozenberg et al. 2004); radial files help in understanding the development, differentiation and temporal changes of cells.

Morphological fluctuations of successive cells along a radial file and the influence of external factors have been investigated for a long time (Ford et al. 1978), but those investigations were greatly restricted by a limited number of radial file comparisons and a restrictive number of successive cell descriptions. Those studies, concerning secondary growth and its relationships with primary growth, were based on fragmentary studies, due to high acquisition and processing costs. Large-scale studies for understanding wood growth and disturbance are available nowadays thanks to automated radial file identification and cell characterization (Brunel et al. 2012, Kennel et al. 2011). In fact, wood continuously records changes in the development of the tree. Such new methodologies and approaches offer real new opportunities for more in-depth investigations of tree biology and climate change (Fonti et al. 2010). The size and shape fluctuations observed along radial files are usually considered to be the result of external constraints (Frankenstein et al. 2005). This is a restrictive and simplified viewpoint: these fluctuations are probably due to a combination of external and internal factors. We aim to promote an objective quantification of these measurements, on the basis of statistics based on numerous image datasets.

This paper focuses on the reliability of such data from statistical studies, i.e. produced by automated radial file identification (Brunel et al. 2012). For large-scale data analyses, it is necessary to ensure *a posteriori* that the measurements are valid in order to find invariants or validate hypotheses. This is a classic line taken in statistics (Bruton et al. 2000) and experimental approaches.

MATERIALS

We processed histological sections of several gymnosperm species: *Pinus caramanica*, *Pinus nigra* and *Abies alba*. Wood cross-sections with a thickness of 20 µm were produced with a vibratome. The sections were stained to increase the contrast between the lumen and cell walls. They were then digitized with an Olympus DP71 LCD camera on an Olympus BX51 microscope. The square color images produced had a resolution of 4 million pixels.

OVERVIEW OF THE METHOD

Radial files are alignments of substantially similar cells in terms of color dynamics, shapes and sizes. Intuitively, the notion of cell alignment implies the existence of cellular organization based on the neighborhood relationships between cells. Those properties are used to identify radial files.

Our approach, described in (Brunel et al. 2012) was divided into three steps: cell identification which specified single cells, cell organization which detected radial files, and cell classification which gave the biological and qualitative typing of cells and radial files.

The incremental construction of radial files was based on two assumptions: (i) such a file is independent from image orientation; (ii) two consecutive cells of a file are very similar. From a methodological point of view, cell individualization was derived from a watershed algorithm (Vincent et al. 1991), and the cell pattern was described by an adjacency graph created from the watershed crest lines. Radial files were then built by finding a reversible path in the graph under spatial constraints (maintaining a specific direction) and similarity constraints (two consecutive cells should show close geometric characteristics). Our method produced several layers of results, respectively corresponding to different observation levels: (i) radial files were classified according to their length, their fractionation and their cell self-similarity, (ii) cells were characterized by geometric parameters (such as size, diameter, shape, etc.) and topological parameters (number of neighbors), (iii) cell component elements (wall, lumen) were characterized by geometric parameters (size, thickness, diameter, etc.).

DEFINITION OF *A PRIORI* RELIABILITY COEFFICIENTS

Quality assessment of these results depends on many factors. It concerns the data set itself (image noise, biological configuration complexity, etc.), the process (algorithm approximations, computation costs, stability, etc.), and the geometrical scale (related to the observation level). Each result should be qualified using different indicators, notably to keep the significant outcome while processing high volume data. It is thus relevant to assign an estimated reliability coefficient to each result. We illustrate cases on file and cell scales.

Radial file construction reliability. We propose defining this radial file reliability coefficient R_f from three criteria: length, fractionation and similarity. The first criterion derives from the file length, L_f, compared to the threshold value T. T is set by 2means and defines the minimum significant length. Fractionation is evaluated from the file construction algorithmic cost. It involves the total number of cells L_{fm} stacked to build the file, the length L_f of the file, and the number of sections N_f. The similarity criterion is built from the product of the *n-1* consecutive cell surface variations. The final reliability coefficient R_f of the file *f* is then defined by a product of normalized terms ranging between 0 and 1. When this coefficient tends to 1, the reliability of the radial file increases.

$$R_f = \left(1 - max\left(\frac{T - L_f}{T}, 0\right)\right)\left(1 - \frac{L_{fm} - L_f}{N_f * L_{fm}}\right)\prod_{j=0}^{n-1}\left(1 - \frac{|Sf_j - Sf_{j+1}|}{Sf_j + Sf_{j+1}}\right)$$

Lumen area estimation reliability. The reliability coefficient of the lumen surface is defined on a similar principle, by a product of two normalized terms, based on the cell size and an image local blur estimation. In fact, the watershed algorithm is known as a robust method for extracting cell boundaries, since it is based on intensity discontinuities which are less sensitive to blurring, whereas the lumen area shows high sensitivity to local blur. The cell area definition involves a two-class clustering split, corresponding to the lower and the higher intensity classes. The local blur estimator is derived from (Ladjal 2006), based on the relationship between local dynamics and intensity differences.

EXPERIMENTS, CONLUSION AND WORK IN PROGRESS

Experiments were performed on ten colored sections of different species of gymnosperms, representative of biological variability. The color quantifies the reliability results, both on a global (Fig. 1) and local (Fig. 2) scale. On a global scale, we can see in Fig. 1 that the reliability estimator properly qualifies the cell files: green colors are assigned to straight and regular cell lines whereas red colors correspond to incomplete cells on the side of the image or incrusted cells. Such a color map alerts the expert to potential unsure classifications and emphasizes the ambiguous biological status of a file or cells. On a local scale, for lumen area computation, the estimator also retrieves obvious cases. The reliability factor also ensures areas showing lower intensity variations (areas in green in Fig. 2, right).

Several points were studied. (i) in angiosperms, the vessel morphology is different and calls for a different evaluation of cell similarity *which defines the third term of the file reliability coefficient*, by taking into account the larger diameter instead of its surface. (ii) other blur estimators were studied, such as estimators based on multi-scale filtering in order to assess the independence of the reliability coefficient with

respect to the blur estimation method. (iii) how the local reliability coefficient could be used to correct the measured values according to the local blur estimator relationship (Fig 2. Left).

To conclude, we introduced some reliability coefficients for image processing related to cell organization identification and illustrated the cases of radial file detection and of lumen area computation. The reliability coefficients were essentielly used to filter the input data of statistical studies conducted by botanists or to draw attention to special biological configurations.

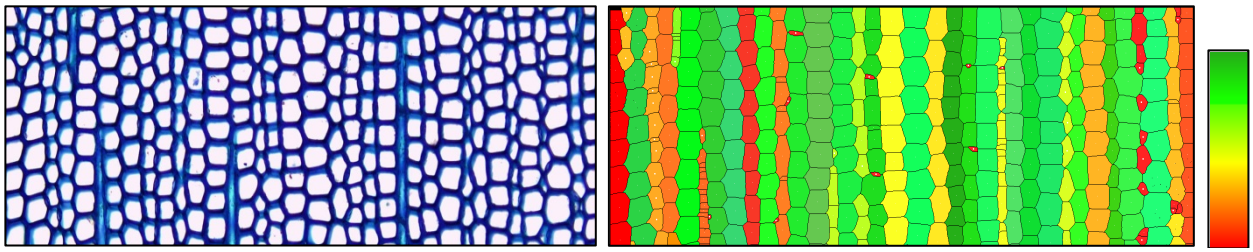

Fig. 1. Left: cross-section of *Abies alba*. Right: automated identification of radial files. Color value qualifies the reliability coefficient: weakly reliable files appear in red tones and highly reliable files are in green tones.

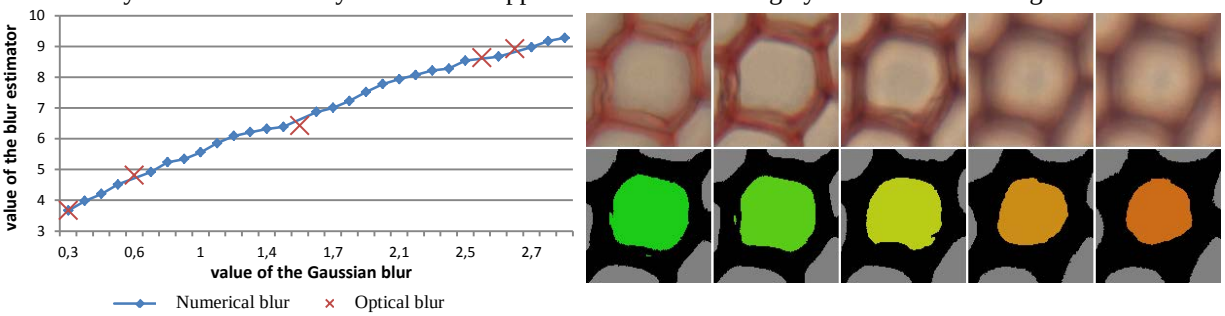

Fig. 2. Left: relationship between a Gaussian blur applied to a cell image and the local blur estimator. The estimator emphasizes the blur tendency; red crosses stand for five microscopic optical blurs, matched from their area variance. Right: the five corresponding lumen areas from the optical blurs: the color qualifies the reliability coefficient, from green for the real lumen area to orange for the reduced lumen area.

The authors acknowledge NUMEV Labex and SIBAGHE Graduate School of the University of Montpellier 2 for their support.

LITERATURE CITED

Barlow PW, Powers SJ. 2005. Predicting the environmental thresholds for cambial and secondary vascular tissue development in stems of hybrid aspen**.** *Annals of Forest Science* **65**: 565-573.

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 files, *Plant Growth Modeling, Simulation, Visualization and Applications, IEEE press, ISBN 978-1-4673-0070-4*, 58-65.

Bruton A, Conway JH, Holgate ST. 2000. Reliability: What is it, and how is it measured?, *Physiotherap* **86**:94-100

Cato S, McMillan L, Donaldson L, Richardson T, Echt C, Gardner R. 2006. Wood formation from the base to the crown in *Pinus radiata*: gradients of tracheid wall thickness, wood density, radial growth rate and gene expression. *Plant Molecular Biology* **60**:565–581.

Fonti P, von Arx G, Garcia-Gonzalez I, Eilmann B, Gärtner H, Eckstein D. 2010. Studying global change through investigation of the plastic responses of xylem anatomy in tree rings. *New Phytologist* **185**: 42–53.

Ford ED, Robards A, Piney MD. 1978. Influence of environmental factors on cell production and differentiation in the early wood of *Picea sitchensis*. *Annals of Botany* **42** *(3): 683-692.*

Frankenstein C, Eckstein D, Schmitt U. 2005. The onset of cambium activity – A matter of agreement? Dendrochronologia **23**: 57-62.

Kennel P, Subsol G, Guéroult M, Borianne P. 2010. Automatic identification of cell files in light microscopic images of conifer wood. *2nd International Conference on Image Processing Theory Tools and Applications* 98–103.

Ladjal S. 2006. Blur estimation in Natural Images, *15^{e} congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle* 112-124.

Liang C, Filion L, Cournoyer L. 1997. Wood structure of biotically and climatically induced light rings in eastern larch (*Larix laricina*). *Canadian Journal of Forest Research* **27**: 1538.1547.

Rozenberg P, Schüte G, Ivkovitch M, Bastien C, Bastien JC. 2004. Clonal variation of indirect cambium reaction to within-growing season temperature changes in Douglas-fir. *Forestry* **77**:257-268.

Vincent L, Soille P. 1991. Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations, *IEE Transactions on Pattern Analysis and Machine Intelligence* **13-6**:583–598.