

A structural method for assessing self-similarity in plants

Pascal Ferraro¹, Christophe Godin², and Przemyslaw Prusinkiewicz³

¹ LaBRI, Université de Bordeaux

² INRIA, UMR Cirad-INRA-CNRS AMAP, Montpellier

³ Department of Computer Science, University of Calgary

Introduction

The important role of architecture in the understanding of plants generates a need for investigational tools. Generic tools have already been developed to visualize plant architecture in three dimensions, to model the development of plant structure, to measure plant architecture, and to analyse and quantify relations between plant components. This paper introduces a new tool for the identification of self-similarity within plant architectures.

Reference

Pascal Ferraro, Christophe Godin, and Przemyslaw Prusinkiewicz. A structural method for assessing self-similarity in plants. In *Proceedings of the 4th International Workshop on Functional-Structural Plant Models*, pp. 56–60.

A structural method for assessing self-similarity in plants.

Pascal Ferraro¹, Christophe Godin², and Przemyslaw Prusinkiewicz³

¹ LaBRI - Université de Bordeaux 1 - France

² INRIA, UMR Cirad-INRA-CNRS AMAP - Montpellier France

³ Department of Computer Science - University Of Calgary - Calgary, Canada

1 Introduction

The important role of architecture in the understanding of plants [8, 12, 23] generates a need for investigational tools. Generic tools have already been developed to visualize plant architecture in three dimensions [20], to model the development of plant structure [6, 20], to measure plant architecture [24], and to analyse and quantify relations between plant components [11]. This paper introduces a new tool for the identification of self-similarity within plant architectures.

In 2001, Prusinkiewicz *et al.* [21] proposed a method for modeling plant branching structures that exploits their self-similarity. They showed that a lilac inflorescence can be modeled using only the measurements of the longest axis of each branch order (*cf.* Figure 17a in [21]). The applicability and accuracy of this technique depend on the degree of self-similarity present in the modeled structure. But similarity and self-similarity, like symmetry, are not easily quantifiable: they are usually considered as a property that is either present or absent, rather than a property that may be present to some degree [22]. In this paper, we propose a procedure for quantifying the degree of self-similarity in measured or simulated plants. This procedure is based on a method for comparing branching structures introduced in [7].

2 Biological motivation

In the growth processes of many living organisms, especially plants, regularly repeated appearances of closely related multicellular structures are readily noticeable. In the simplest case, one observes the same structure being periodically repeated along an axis, such as leaves along a stem. In more complicated cases, the entire structure of a previous stage is repeated as part of the organism at a later stage, resulting in “*compound* organisms or organs, [. . .], such as compound branching systems” [15].

This phenomenon of self-similarity [16] has been tentatively captured by several botanical notions. For example, Troll defined an inflorescence as a system consisting of the main florescence and of *paracladia* [8]; a paracladium is “a branch which bears the same sort of structures like those borne on the terminal portion of its mother branch” [14]. Self-similarity is also related to the definition of the *physiological age of meristems* in vegetative parts of plants [2, 9]. According to this concept, all the meristems of a plant undergo the same series of developmental stages (called physiological ages), and two meristems with the same physiological age produce essentially similar structures [1, 13].

These notions are reflected in several formal approaches to the description and modelling of plants. For example, Fritjers and Lindenmayer [8], defined *paracladial relationships* in branching patterns: a branching structure exhibits a paracladial relationship if “the behaviour of a branch is exactly the same as the behaviour of that part of the mother branch which has developed since the formation of the daughter branch”. Prusinkiewicz *et al.* introduced the concept of *branch mapping* [21], which holds if “given two branches of the same order, the shorter branch is identical [. . .] to the top portion of the longer branch”. This concept was used to reduce the complexity of three-dimensional models of branching structures, including compound inflorescences and tree structures. In the context of growth modeling, the notion of physiological age was exploited to model the development

of plants by de Reffye *et al.* [5], who introduced the notion of the *reference axis* as a list of all the differentiation stages of a tree meristem. By studying the different developmental stages of *Zelkova serrata* (Japanese elm), they showed that growth and branching processes are similar in any part of the tree and that it is possible to identify a theoretical axis that includes all the states of morphological differentiation of a growth unit. In this sense, the “reference axis” captures the behaviour of a typical meristem in a plant over time, *i.e.* the potential self-similar development of a plant.

In reality, the self-similarities captured by paracladial relationships, branch mapping or physiological age are only approximate. While easy to observe in obvious cases, they may be hidden in more complex structures [4]. Here, we propose a computational method to quantify the similarity between different branching structures within a plant and thus quantitatively assess paracladial relationships. This method can be applied to plants represented as tree structures.

3 Evaluation of similarity between branching systems

A simple approach for comparing two branching systems consists of summarizing each individual by a small number of synthetic variables (*eg.* fruit production or crown size). The similarity between the systems is then reduced to the proximity of these synthetic variables. In contrast, our interest is in the similarity of structures. This requires a definition of distance that would take into account the topological and spatial organization of entities into branching structures. In the following, we will use the common description of plants as mathematical (graph theoretic) trees [10, 20].

A plant is considered as a connected assembly of botanical entities such as internodes or annual growth increments [3]. When describing plants at the architectural level, we assume that each such entity is physically attached to at most one parent entity. The resulting topological structure is a rooted tree graph $T = (V, E)$, *i.e.* a graph in which every vertex except one (the root r) has only one parent vertex [10]. The root has no parent vertex. In the following a tree graph rooted in r will be represented by $T[r]$, and the empty tree graph will be denoted by $\theta = (\emptyset, \emptyset)$. In order to identify the different axes on a given plant, two types of relations between entities are distinguished: an entity can either precede (symbol ‘<’) or bear (symbol ‘+’) another entity. We suppose that an entity can be attached to only one other entity by a connection ‘<’. A sequence of entities connected by ‘<’ forms an axis of a tree [20].

The evaluation of similarity between branching structures has been largely studied in computer science and is known as the *tree-to-tree comparison problem* [18]. The distance between two tree graphs is equal to the minimum cost of the sequence of edit operations needed to transform one tree graph into the other. We consider three kinds of elementary edit operations on a tree graph T [25]: substituting one vertex for another (note that this changes their labels), deleting a vertex, and inserting a vertex. A constraint is added to the definition of insertions and deletions: we only consider insertions of leaves or vertices between a parent vertex and its set of children; deletions are similarly constrained¹. A cost function is defined for each edit operation s which assigns a non-negative real number $\gamma(s)$ to s as follows: $\gamma(s) = d(v, w)$ if s is a substitution of v by w ; and $\gamma(s) = d_{\text{ind}}(v)$ if s is an insertion or a deletion of vertex v . Here d is a distance called the *elementary distance*, defined on the set of vertices of T . We assume, for any pair of vertices v and w , that $d(v, w) \leq d_{\text{ind}}(v) + d_{\text{ind}}(w)$.

Let S be a sequence of n edit operations (s_1, s_2, \dots, s_n) which transforms a tree graph T_1 into another tree graph T_2 . The cost $\Gamma(S)$ of S is defined by summing up the cost of the edit operations that compose S : $\Gamma(S) = \sum_{s_i \in S} \gamma(s_i)$. The dissimilarity measure $D(T_1, T_2)$ between a tree graph T_1 and a tree graph T_2 is then defined as the minimum cost of a sequence in S that transforms T_1 into T_2 .

¹ Zhang and Jiang [26] have shown that the computation of distance between tree graphs using such edit operations is a MAX SNP-hard problem unless constraints are added to the definition of edit operations.

This dissimilarity measure coincides with the definition of a distance between unordered² tree graphs proposed by Zhang [25]. He proposed a polynomial algorithm to recursively calculate the cost of the optimal sequence of edit operations, which was applied to compare plant architectures by Ferraro and Godin [7]. In that case, the computed distance strongly depended on the size of the compared tree graphs. In order to make the comparison results less dependent on this size, we normalize them by dividing the distance by the total number of vertices in compared tree graphs:

$$\hat{D}(T_1, T_2) = \frac{D(T_1, T_2)}{(|T_1| + |T_2|)}.$$

This dissimilarity measure is a non-negative real number less than 1.

4 Quantification of tree self-similarity

In this section we extend the previous algorithm to identify and quantify self-similarity of plant structures.

Let us consider a plant \mathcal{P} represented by a tree graph $T[r]$. The set of all sub-tree graphs of $T[r]$ is denoted by $\mathcal{T}(r)$. Let v be a vertex of $T[r]$ and let $T[v]$ be a sub-tree graph rooted in v of $T[r]$; then the *axis* $A[v] = (V, E)$ of $T[v]$ is the sub-graph of $T[v]$ such that $V = \{v_1, v_2, \dots, v_n\}$, $v_1 = v$, and for any i in $\{1 \dots n - 1\}$, (v_i, v_{i+1}) is an edge of type ' $<$ ' (Figure 1). $A(r)$ is called the *main axis* of \mathcal{P} (i.e., the trunk).

If $T[v]$ is a sub-tree of \mathcal{P} and $A[v]$ is the axis of $T[v]$ starting at v , we denote by $\mathcal{A}(v)$ the set of sub-trees of \mathcal{P} rooted in a vertex of $A[v]$:

$$\mathcal{A}(v) = \{T[v_1], T[v_2], T[v_3], \dots, T[v_n]\}.$$

Note that, in this definition, v_{i+1} is necessarily a child of v_i . $\mathcal{A}(v)$ is thus a set of embedded tree graphs (Figure 1). Using the above notions, we quantify the similarity of a branch to the main axis as follows:

Definition 1 For any sub-tree graph $T[v]$ in $\mathcal{T}(r)$, we call the *paracladial coefficient* of $T[v]$ the quantity

$$pr(v) = \min_{T[r_i] \in \mathcal{A}(r)} \{\hat{D}(T[v], T[r_i])\}.$$

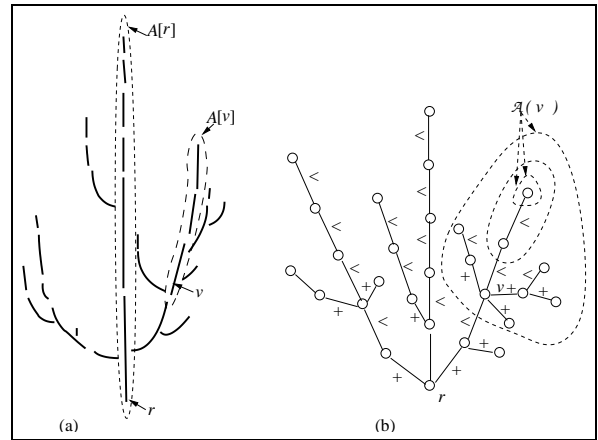


Fig. 1. Axis definition in a branching system (a) and in its representation as a tree graph (b).

² An unordered tree graph is one in which there is no ordering distinction among the children of any vertex.

The paracladial coefficient is thus a measure of similarity between a given branch and the top portion of the main axis that is the closest (in the sense of distance \hat{D}) to that branch. In order to evaluate self-similarity of an entire structure, we compute statistical parameters such as the mean value of the distance and its variance:

$$\overline{PR} = \frac{1}{|\mathcal{T}(r)|} \times \sum_{v \in \mathcal{T}} pr(v),$$

$$var(PR) = \frac{1}{|\mathcal{T}(r)|} \times \sum_{v \in \mathcal{T}} (pr(v) - \overline{PR})^2.$$

Here $|\mathcal{T}(r)|$ represents the number of sub-tree graphs in $\mathcal{T}(r)$. Low values of \overline{PR} indicate high self-similarity of the described branching structure.

5 Example

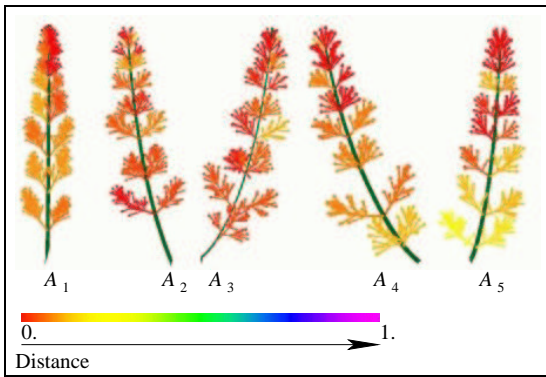


Fig. 2. Color-coded representation of paracladial coefficients of branches of order 2

In this section we apply our method to assess self-similarity in a sample of five *Syringa vulgaris* (common lilac) inflorescences. The inflorescences were collected in Calgary in the spring of 2001. The length and the diameter of each internode, the spatial orientation of each branch, and the length and the diameter of each flower have been measured using a digital calliper connected to a computerized data collection system. The data were analyzed using AMAPmod [11], a software dedicated to plant architecture analysis, in which the algorithms of comparison between tree-graph structures have been implemented. The results were visualized using the modeling program cpfg [17].

	\overline{PR}	$var(PR)$	nb
A_1	0.0164	0.00298	70.83%
A_2	0.0196	0.00247	78.26%
A_3	0.0212	0.00315	68.18%
A_4	0.0227	0.00223	79.16%
A_5	0.0192	0.00336	72.72%

Table 1. The mean value \overline{PR} and variance $var(PR)$ of the paracladial index $pr(v)$ for five lilac inflorescences, and (c) percentage of branches of order 2 perfectly similar to top portions of their supporting branches.

To evaluate self-similarity, we first need to define a local distance between elementary entities of the branching systems. In our example, we used a topological distance, with the elementary distance d defined as $d(v, w) = 0$ and the distance d_{indel} defined as $d_{indel}(v) = 1$ for any pair (v, w) of vertices (botanical internodes). These definitions capture *topological* self-similarity [19] of the architecture (*i.e.*, the topological organization of entities in the branching system), without taking into account any geometrical attributes.

The paracladial coefficients $pr(v)$ of all the order-2 branches are visualized in Figure 2. Generally, these coefficients are very small: the inflorescences show only a few “errors” in the mapping. This high degree of self-similarity is reflected in the statistics collected in Table 1.

6 Conclusion

This paper addressed the problem of quantifying the degree of self-similarity in branching plant structures. To this end, we introduced the notion of paracladial index and its statistics. The index characterizes the similarity between an individual branch and the main stem of the structure, whereas the statistics provide a global measure of self-similarity of the entire structure.

The presented results can be applied to improve the methodologies of measurement and architectural modeling of plants. For example, when measuring plants, it may be unnecessary to consider all branches, because of their likely similarity to the main branch. When modeling plants, some of the data characterizing the main branch may be re-used in the lateral branches, thus simplifying the models. The paracladial index and its statistics make it possible to quantify the resulting errors.

Acknowledgment

The authors thank Lars Mündermann for sharing with us his database of lilac inflorescence measurements, and for contributing to the model of these inflorescences.

References

1. A. Arber. *Natural philosophy of plant form*. Cambridge University Press, 1950.
2. D. Barthélémy. Levels of organization and repetition phenomena in seed plants. *Acta Biotheoretica*, 39:309–323, 1991.
3. A. D. Bell. *Plant Form, An illustrated Guide to Flowering Morphology*. Oxford University Press, 1993.
4. E. Costes and Y. Guédon. Modelling the sylleptic branching on one-year-old trunks of apple cultivars. *Journal of American Society for Horticultural Science*, 1(122):53–62, 1997.
5. P. de Reffye, P. Dinouard, and D. Barthélémy. Architecture et modélisation de l'orme du japon *zelkova serrata* (thunb.) makino (ulmaceae): la notion d'axe de référence. In *De la forêt cultivée à l'industrie de demain: 3e colloque Sciences et Industries du Bois, Arbora*, pages 351–352, Bordeaux, France, 1990.
6. P. deReffye, T. Fourcaud, F. Blaise, D. Barthélémy, and F. Houllier. A functional model of tree growth and tree architecture. *Silva Fennica*, 31(3):297–311, 1997.
7. P. Ferraro and C. Godin. A distance measure between plant architectures. *Annals of Forest Science*, 57:445–461, 2000.
8. D. Frijters and A. Lindenmayer. Developmental descriptions of branching patterns with paracladial relationships. *Automata, Langages, Development*, 1976.
9. L. Gatsuk, O. Smirnova, L. Zaugolnova, and L. Zhukova. Age states of plants of various growth forms: a review. *Journal of Ecophysiology*, 68:675–696, 1980.
10. C. Godin and Y. Caraglio. A multiscale model of plant topological structures. *Journal of theoretical biology*, 191:1–46, 1998.
11. C. Godin, Y. Guédon, E. Costes, and Y. Caraglio. Measuring and analyzing plants with the amapmod software. In M. Michalewicz, editor, *Advances in computational life sciences, Vol I : Plants to ecosystems*, volume January, pages 63–94. Csiro, Australia, 1997. chapitre 4.
12. F. Hallé, R. Oldeman, and P. Tomlinson. *Tropical trees and forests. An architectural analysis*. Springer Verlag, New York, 1978.
13. H. Honda. Description of the form of trees by the parameters of tree-like body. *Journal of theoretical biology*, 1971.
14. A. Lindenmayer. Paracladial relationships in leaves. *Ber. Deutsch. Bot. Ges. Bd*, 90:287–301, 1977.
15. G. M. Herman, A. Lindenmayer, and G. Rozenberg. Description of developmental languages using recurrence system. *Mathematical systems theory*, 8(4):316–341, 1975.
16. B. Mandelbrot. *The fractal geometry of nature*. W.N. Freeman, New York, USA, 1983.
17. R. Měch. CPMG version 3.4 user's manual. Department of Computer Science, University of Calgary, May 1998.
18. A. Ohmori and E. Tanaka. A unified view on tree metrics. *Syntactic and Structural Pattern Recognition*, 45:85–100, 1988.
19. P. Prusinkiewicz. Self-similarity in plants: Integrating mathematical and biological perspectives. In M. Novak, editor, *Proceedings of Fractal 2004*, Vancouver, BC, April 2004 (to appear). World Scientific Singapore.
20. P. Prusinkiewicz and A. Lindenmayer. *The algorithmic beauty of plants*. Springer Verlag, New York, 1990.
21. P. Prusinkiewicz, L. Mündermann, R. Karwowski, and B. Lane. The use of positional information in the modeling of plants. In A. SIGGRAPH, editor, *Computer Graphics Proceedings, Annual Conference Series, 2001*, pages 289 – 300, Los Angeles, CA, August 2001.
22. J. Rosen. *Symmetry discovered*. Dover Publications, inc., Mineola, New York, 1998.
23. J. Ross. *The radiation regim and the architecture of plant stands*. Junk W. Pubs., The Hague, The Netherlands, 1981.
24. H. Sinoquet, P. Rivet, and C. Godin. Assessment of the three-dimensional architecture of walnut trees using digitizing. *Silva Fennica*, 3:265–273, 1997.
25. K. Zhang. A constrained edit distance between unordered labeled trees. *Algorithmica*, 15:205–222, 1996.
26. K. Zhang and T. Jiang. Some max snp-hard results concerning unordered labeled trees. *Information Processing Letters*, 49:249–254, 1994.