

ANALYZING SYMMETRY IN BIOLOGICAL SYSTEMS

David Milner, Hagit Hel-Or, Daniel Keren

Department of Computer Science
University of Haifa, Haifa, Israel 31905
hagit@cs.haifa.ac.il

Shmuel Raz, Eviatar Nevo

Institute of Evolution
University of Haifa, Haifa, Israel 31905

ABSTRACT

This paper suggests a new measure of symmetry for bifurcating structures, which relies not only on topology and ordering, but also on quantitative properties (e.g. length of branches). This measure is based on the amount of energy required to symmetrize the structure, hence is especially suitable to biological structures. It is tested by its ability to differentiate between two classes of leaves.

1. INTRODUCTION

Symmetry is a basic property of shapes and structures. In biological and physical systems, symmetry seems to imply stability and natural development [1]. Symmetry depends on shape contour, shape area and/or shape moments. We constrain our study to symmetry of a specific class, namely, mirror symmetry. A 2D object is mirror-symmetric if it is invariant under a reflection about a line (the mirror-symmetry axis). Mirror symmetry has been extensively studied, and numerous measures of symmetry have been developed to quantify the deviation from exact symmetry. These measures have been developed to deal with 2d shapes and polygons [2, 3], and 3D convex structures and polyhedra [4, 5]. However, these approaches are inadequate for dealing with many biological and chemical structures, namely bifurcating structures, e.g. molecules. Motivated by a biological/environmental study on the effects of stressful environments on the symmetry of plants and animals, we developed a symmetry measure for bifurcating structures such as veins of plant leaves (Figure 3a). Previous studies on symmetry of graph and tree structures [6] considered only topological aspects of the structures. However, given our motivation, the morphological/geometrical aspect of the structures is of crucial importance. In several studies, [7, 8, 9] measures of symmetry were developed for molecular structures which are 2D and 3D bifurcating structures. These approaches considered the geometry of the structures and the topological connectivity between molecular atoms, but assumed no other constraint on the structures. In this work we introduce a novel approach in which we constrain the topological

connectivity of the structures according to biological models. Our algorithm for measuring symmetry of bifurcating structures such as leaf veins, takes into account biological mechanism of development and growth [10] and relies on nature's laws rather than on arbitrary mathematical models.

In this study we present a measure of symmetry, and provide an algorithm for its computation. Additionally we demonstrate the effectiveness/performance of the measure in an ecological experiment. The measure of symmetry we propose is based on early work on Continuous Symmetry Measures (CSM) [2]. In these studies, symmetry is evaluated by estimating the energy (work) required to deform a non-symmetric configuration into a perfectly symmetric one. In [7, 8, 9] symmetry of molecules was measured using the CSM approach, by allowing deformations that preserve the topology of the original structures. In this paper we allow deformations that do not necessarily preserve topological connectivity but that are consistent with biological models of the source structures (in our case veins of leaves).

2. BACKGROUND

The motivation of this work originates from an ongoing ecological study on the effects of stressful environments on the growth of plants and animals. It has long been known that shape and form of plants and animals are highly affected by their environmental conditions. Stressful growing environments typically cause deviation from regularity of form and shape. Numerous measures of shape and form have been used to analyze the extremities of these effects, including area, width to length ratio, distance between specific features and more. One important measure of shape is that of symmetry which has already been shown to be a discriminating factor in biological systems [11]. However, attempts to measure symmetry of plant leaves, in support of these theories, were unsuccessful. We contend that this is due to the fact that the symmetry of leaf area or leaf boundary does not capture the distinctive factor of leaves. This is in accord with developmental models of leaves, where the effects of stressful environments strongly affect the stages of vein growth rather than the subtle variations in the leaf

boundary [10]. Thus the goal of this work is twofold: first we aim to develop a measure of symmetry of the bifurcating structure of leaf veins, and second, to quantitatively test these claims of ecological studies.

3. MEASURING SYMMETRY OF BIFURCATING STRUCTURES

Guided by the underlying biological processes, we define the measure of symmetry of a structure as the minimal amount of energy required to deform it into a symmetric structure. In our approach, deformations are restricted to those consistent with known biological mechanisms [10] (e.g. leaf growth). We concentrate on planar bifurcating structures representing veins of leaves. Figure 3a shows a specimen with the vein structure marked. These structures are assumed to consist of a central vein which is viewed as the symmetry axis, from which secondary veins bifurcate to both sides.

Given a leaf, we find the minimum energy deformation which transforms it into a symmetric structure. Deformations consist of a series of three permissible “elementary” deformations, motivated by the biological growth model of leaves [10]:

1. Insertion (deletion) – an additional secondary vein is added (deleted).
2. Translation – a secondary vein is translated along the main vein.
3. Elongation (contraction) – length of a secondary vein is changed.

Figure 1 shows examples of these “elementary” deformations. Every such action is associated with a cost function which represents the energy required to perform the action. The energy of deforming a structure accumulates the cost of performing each “elementary” deformation. The measure of symmetry of a structure is the cost of the deformation of minimum energy that transforms the structure into a symmetric structure.

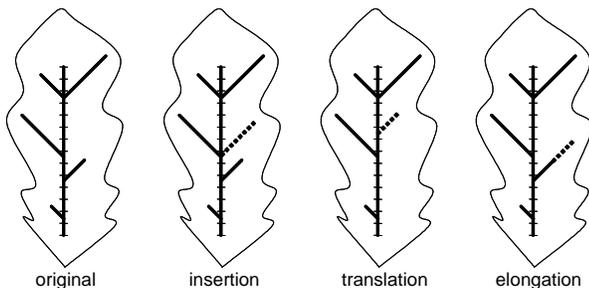


Fig. 1. Elementary deformations allowed in the symmetry measure evaluation.

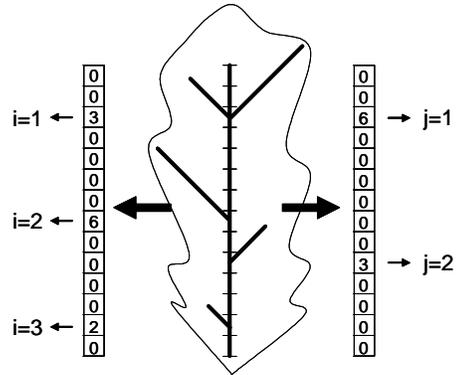


Fig. 2. The main leaf vein is segmented into sections and mapped into two vectors (see text).

To find the minimum energy deformation, we represent the problem as a string matching problem. For a given specimen, the central vein is segmented into a large finite number of sections. The segmented vein is mapped into two vectors, the length of which is equal to the number of segments in the vein. The vectors describe the left and the right bifurcation structure by assigning ‘0’ to an entry if there is no secondary vein bifurcating out of the vein segment corresponding to that entry. A positive real value is assigned to the entry if there is such a bifurcation and the assigned value equals the length of the secondary vein. An example is shown in Figure 2.

Finding the minimum energy deformation of the bifurcating structure is equivalent to finding the minimum cost of transforming one vector into the other, using “elementary” deformations. This problem is solved using dynamic programming.

Let $i = 1 \dots n$ enumerate the secondary veins on one side of the main vein and let $j = 1 \dots m$ be the enumeration of the secondary veins on the other side. Denote by $C_{i,j}$ the cost of the minimal energy deformation of the sub-structure consisting of the main vein and the first i secondary veins on the left and the first j secondary veins on the right.

The value $C_{i,j}$ is given by one of the following possibilities:

1. Pair the left vein i with the right vein j . Thus $C_{i,j}$ equals $C_{i-1,j-1}$ plus the cost of translating vein i to the position of vein j and the cost of changing the length of vein i to equal that of vein j .
2. Insert a new left vein as a pair for vein j . Thus $C_{i,j}$ equals $C_{i,j-1}$ plus the cost of inserting a new vein of length equal to vein j .
3. Insert a new right vein as a pair for vein i . Thus $C_{i,j}$ equals $C_{i-1,j}$ plus the cost of inserting a new vein of length equal to vein i .

It is easily shown that this recursive definition implies that $C_{n,m}$ is indeed an optimal solution which minimizes the cost function. Calculating $C_{i,j}$ requires calculating the costs of the sub-problems $C_{i-1,j-1}$, $C_{i-1,j}$ and $C_{i,j-1}$, thus, the dynamic programming approach solves for the optimal solution in a bottom-up approach initialized with $C_{0,0} = 0$.

This approach is detailed in the following: We first define the three elementary deformations:

$$\begin{aligned} COI(l) &= I_0 * l \\ COT(l, d) &= T_0 * l * d \\ COE(l, new_l) &= E_0 * |l - l_{new}| \end{aligned}$$

COI is the Cost Of Insertion of a vein of length l . Note, the cost is invariant to the position of vein insertion along the main vein and to the side (left/right) on which insertion is performed. Cost of insertion is equal to cost of deletion, thus, the latter is not used in the algorithm.

COT is the Cost Of Translating a vein of length l a distance of d units. Direction of translation is irrelevant to cost.

COE is the cost of elongating (shortening) a vein of original length l to new length l_{new} . Cost of length change is invariant to scaling direction (elongation or shortening).

I_0 , T_0 and E_0 are scalar weights determining the relative affects of each elementary deformation. Their values were set according to the biological growth model [10], thus, $I_0 \gg T_0 \gg E_0$. Weight Values were set to: $I_0 = 5$, $T_0 = 2$ and $E_0 = 1$.

Using these cost functions for elementary deformations, we define the following cost functions used in the algorithm:

$$\begin{aligned} COP_{i,j} &= COT(\min(\text{length}(i), \text{length}(j)), \\ &\quad \text{abs}(\text{pos}(i) - \text{pos}(j))) + \\ &\quad COE(\text{length}(i), \text{length}(j)) \\ COI_i &= COI(\text{length}(i)) \end{aligned}$$

COP is the cost of pairing left vein i with right vein j . Cost consists of translating and changing length of one vein to equal the position and length of the other. Since COT is dependent on the length of the vein, the translation is performed on the shorter of the two veins. Since COE is symmetric in direction of length change, w.l.o.g. the length of the left vein is changed to equal that of the right.

COI_i is the cost of inserting a new vein to be paired with vein i .

We now define the recursive cost function $C_{i,j}$ as follows:

$$\begin{aligned} C_{i,j} &= \min(C_{i-1,j-1} + COP_{i,j}, \\ &\quad C_{i,j-1} + COI_j, \\ &\quad C_{i-1,j} + COI_i) \\ C_{0,0} &= 0 \end{aligned}$$

The dynamic programming process which computes the values $C_{i,j}$, outputs the optimal solution $C_{n,m}$, which is defined as the measure of symmetry of the original bifurcating structure. In addition to the symmetry value, the sequence of elementary deformations, used to obtain the optimal solution is reconstructed and a symmetrized version of the original structure is obtained.

4. AN EXAMPLE

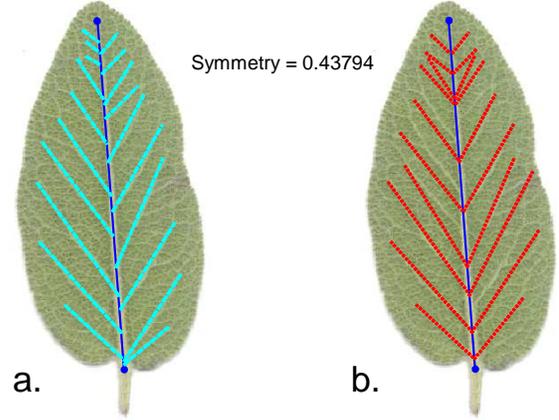


Fig. 3. a) Veins of the leaf are marked. b) The resulting symmetric structure.
Note: Figure is in Color.

Figure 3a shows an example of a leaf with marked veins. The resulting symmetric vein structure and the symmetry value are shown in Figure 3b.

5. EXPERIMENTS

The symmetry measure developed for bifurcating structures was tested in an ecological study on the effects of stressful environments on the growth of plants and animals. Stress effects have been successfully demonstrated in the "Evolution Canyon" model system in lower Nahal Oren, Mount Carmel, Israel [12]. The microclimatic conditions on the two slopes vary dramatically, affecting the evolution of organisms at all levels. The south-facing slope (SFS) receives more solar radiation, higher temperature and stronger illumination than the north-facing slope (NFS). Consequently, the SFS is considered a more stressful environment for growth and survival of plants and animals than the NFS. It has been shown that the area of a leaf is an important characteristic which is affected by stress conditions. It is also widely believed that leaf symmetry is affected by stress conditions during growth [11].

In this experiment we show that the symmetry measure of bifurcating structures applied to the veins of the leaves, indeed measures the deviation in perfect symmetry as affected by the stressful environment. We show this correlation using classification. We build a classifier that discriminates between samples of the canyon's SFS and the NFS, and show that classifying a novel specimen based on vein symmetry improves classification over classification based on the leaf area.

The classifier used the Support Vector Machine (SVM) technique [13]. We had 282 samples available: 138 from the NFS and 144 from the SFS. Due to the small number of examples we adopted the V-Fold Cross Validation method for testing the quality of classification. Thus, subsets of the SFS and NFS samples were used to for training the SVM. Samples not included in these subsets were then used as test cases to evaluate classification. This process was repeated 111 times with different subsets. The classifier was trained and then tested using vein symmetry, using leaf area and using both as the measured features. The percentage of correct classification was determined in each case. Average results over all runs are shown in Table 1.

	Vein Symmetry	Area	Vein + Area
Average	64.8649	74.9868	79.9682

Table 1. Correct classification results using Vein Symmetry, leaf area and both as measured features.

Results show that leaf area allows %75 correct classification. Results also show that classification can be performed using vein symmetry as well. However, given the popularity of using the area as a discriminating feature it is important to see that using vein symmetry together with leaf area as features, improves recognition by %5.

6. CONCLUSION

We proposed a measure of symmetry for bifurcating structures, namely leaf veins, which is based on biological growth models. The measure is based on geometrical as well as topological properties of the structures. The measure quantifies the amount of energy required to deform the structure into a perfectly symmetric one. Deformation is restricted to actions which are in accord with biological growth models. We introduce a computational scheme to calculate the symmetry measure, and showed the effectiveness of the symmetry value in an ecological experimentation.

7. REFERENCES

- [1] V. Valen, "A study of fluctuating asymmetry," *Evolution*, vol. 16, pp. 125–142, 1962.
- [2] H.Zabrodsky, S.Peleg, and D.Avnir, "Symmetry as a continuous feature," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, no. 12, pp. 1154–1166, 1995.
- [3] A.V. Tuzikov and S.A. Sheynin, "Symmetry measure computation for convex polyhedra," *Journal of Mathematical Imaging and Vision*, vol. 16, no. 1, pp. 41–56, January 2002.
- [4] M. Kazhdan, B. Chazelle, D. Dobkin, A. Finkelstein, and T. Funkhouser, "A reflective symmetry descriptor," in *European Conference on Computer Vision ECCV*, May 2002, pp. 642–656.
- [5] I. Shimshoni, Y. Moses, and M. Lindenbaum, "Shape reconstruction of 3d bilaterally symmetric surfaces," *International Journal Computer Vision*, vol. 39, no. 2, pp. 97–110, 2000.
- [6] L. Babai, "Isomorphism testing and symmetry of graphs," *Ann. Discrete Mathematics*, vol. 8, pp. 101–109, 1980.
- [7] H.Zabrodsky, S.Peleg, and D.Avnir, "Continuous symmetry measures," *Journal of the American Chemical Society - JACS*, vol. 114, pp. 7843–7851, 1992.
- [8] H.Zabrodsky, S.Peleg, and D.Avnir, "Continuous symmetry measures ii: Symmetry groups and the tetrahedron," *Journal of the American Chemical Society - JACS*, vol. 115, pp. 8278–8289, 1993.
- [9] H.Zabrodsky and D.Avnir, "Continuous symmetry measures, iv: Chirality," *Journal of the American Chemical Society - JACS*, vol. 117, pp. 462–473, 1995.
- [10] R. Aloni, "Foliar and axial aspects of vascular differentiation-hypotheses and evidence," *J. Plant Growth Regul.*, vol. 20, pp. 22–34, 2001.
- [11] A.Moller and J.Swaddle, *Asymmetry, Developmental Stability, and Evolution*, Oxford University Press, 1997.
- [12] E. Nevo, "Evolution in action across phylogeny caused by microclimatic stresses at "evolution canyon"," *Theoretical Population Biology*, vol. 52, pp. 231–243, 1997.
- [13] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.