Bulletin of Mathematical Biology (2026) 88:198
https://doi.org/10.1007/s11538-026-01754-9

Society for Mathematical Biology

METHODS



A High Throughput and User-Friendly Computational Tool for Quantifying and Visualizing Symmetry Variation in Image Datasets

Liam D. O'Brien[1] · **Punit Gandhi**[2] · **Maria-Veronica Ciocanel**[3,4] · **Adriana T. Dawes**[1,5]

Received: 11 May 2026 / Accepted: 1 September 2026
© The Author(s) 2026

Abstract

Symmetry is a fundamental property of biological systems, yet it is commonly reduced to discrete, subjective classifications that fail to capture the continuous and graded nature of symmetry in real organisms. We present a high-throughput, user-friendly computational framework for quantifying and visualizing biological symmetry as a continuous variable using *transformation information* ($\mathcal{TI}$), an information theoretic measure that quantifies the divergence between an object and its transformed self. We establish key mathematical properties of $\mathcal{TI}$, including invariance, equivariance, and Lipschitz continuity, that guarantee robustness to transformations such as translation, rotation, reflection, and rescaling, enabling reliable symmetry quantification across heterogeneous image datasets without manual landmarking or image alignment. Building on this theoretical foundation, we introduce a computational pipeline that automates symmetry center detection, $\mathcal{TI}$ curve computation, and visualization of the symmetry data using functional principal component analysis (fPCA). We demonstrate the utility of this framework on two biological datasets: a large collection of angiosperm flower images, where we investigate rotational and reflectional symme-

✉ Liam D. O'Brien
obrien.1093@osu.edu

✉ Adriana T. Dawes
dawes.33@math.ohio-state.edu

Punit Gandhi
gandhipr@vcu.edu

Maria-Veronica Ciocanel
veronica.ciocanel@duke.edu

1 Department of Mathematics, The Ohio State University, Columbus, OH, USA

2 Department of Mathematics and Statistics, Virginia Commonwealth University, Richmond, VA, USA

3 Department of Mathematics, Duke University, Durham, NC, USA

4 Department of Biology, Duke University, Durham, NC, USA

5 Department of Molecular Genetics, The Ohio State University, Columbus, OH, USA

Published online: 07 October 2026

tries, and a dataset of simulated fern images, where we investigate fractal self-similarity under scaling and translation. Our analysis of the floral dataset provides novel evidence of convergent evolution. In both datasets, fPCA of the resulting $\mathcal{TI}$ profiles reveals that symmetry is best characterized as a continuum rather than a set of discrete classes, and there is subtle morphological variation that is not captured by traditional categorical approaches. Together, these results suggest that $\mathcal{TI}$ provides a mathematically principled and biologically informative foundation for integrating symmetry into quantitative evolutionary, ecological, and developmental models.

Keywords Transformation information · Asymmetry quantification · Approximate symmetry · Symmetry classification · Image analysis · Floral symmetry

1 Introduction

Symmetry is a fundamental property of living systems, manifesting across multiple temporal and spatial scales, and driving biological functions such as fecundity and evolvability. In flowering plants, for example, symmetries in floral structures are tightly linked to their reproductive behavior and ecological interactions (Endress 1999; Luo et al. 1996; Neal et al. 1998; Sargent 1539; Fenster et al. 2004; Nikkeshi et al. 2015; Wang et al. 2023).

Uncovering the role of symmetry in biology relies on large-scale phenotyping; however, researchers often default to treating symmetry as a discrete, system-specific variable and classify specimens based on subjective visual criteria. Flowers, for instance, are typically grouped as actinomorphic (radial symmetry) or zygomorphic (bilateral symmetry); yet, many flowers fit neither category, making the classification inadequate. Such categorical approaches bias symmetry measurements, prevent meaningful comparison across heterogeneous populations, and impose an artificial structure on organisms that are inherently variable. By reducing a rich spectrum of symmetry to a set of subjective categorical labels, scientists risk discarding subtle information that may hold the key to understanding how symmetry affects biological fitness.

Mathematical descriptions of symmetry have traditionally relied on the rigorous framework of group theory, which defines symmetry as invariance under a specified transformation. While this formalism is powerful, its application to biological data is severely limited. Group theory classifies objects by the subgroup of transformations they are invariant under, partitioning them into discrete symmetry types such as C_n (cyclic) or D_n (dihedral). Since biological symmetries are rarely exact, most natural objects would not belong to any symmetry class under this classification. For example, a flower with nearly but not perfect five-fold rotational symmetry is assigned to the same class as a completely asymmetric object. Moreover, even among objects sharing a symmetry class, meaningful morphological differences are discarded. As a result, rigid group-theoretic definitions fail to capture the graded and continuous nature of symmetry in real biological data.

Methods for measuring approximate symmetry and morphology in biology have relied heavily on landmark-based morphometrics or outline analyses (Klingenberg 2002), which quantify deviations from ideal symmetry by comparing homologous

points or structures reflected across some axis. More recent work has aimed to automate and generalize symmetry measurements. For instance, Savriama and Klingenberg (2011) extended landmark-based morphometrics to consider general symmetries. These established methods still require expert input to select biologically-meaningful landmarks and often assume bilateral symmetry across a user-selected axis. To address these issues, Gandhi et al. (2021) proposed using an information-theoretic measure called *transformation information* ($\mathcal{TI}$) (Vstovsky 1997) for biological data. $\mathcal{TI}$ quantifies symmetry by comparing an object to its transformed self across a broad set of transformations. Rather than assigning each specimen to a discrete symmetry category, it represents symmetry with a function that varies continuously (in function space) and captures the full symmetry profile. $\mathcal{TI}$ does not require segmentation of the object or selection of key landmarks, avoiding two steps that are often difficult or subjective for biological imagery. However, this measure still requires that a meaningful set of symmetry transformations is selected, that the object of interest is isolated against a uniform black background, and that the center of the object is selected. This approach trades difficult or subjective steps for more tractable and unbiased ones, which can often be automated. Bermingham et al. (2023) also recently proposed the use of the *extended persistent homology transform* to automatically quantify approximate symmetries. This approach is well-suited for capturing topological changes in binary images or shapes under a symmetry transformation, whereas the $\mathcal{TI}$ approach captures changes in scalar-valued functions under a symmetry transformation (e.g. pixel intensity for an image).

Ciocanel et al. (2025) extended the $\mathcal{TI}$ framework by studying transitions in symmetry, proposing a formal analogy with bifurcation theory. This study used $\mathcal{TI}$ to characterize critical changes in symmetry across model-generated developmental patterns and demonstrated the sensitivity of $\mathcal{TI}$ to subtle changes in morphology that are not easily detected by eye.

In this work, we advance the $\mathcal{TI}$ paradigm by developing a robust, high-throughput, and user-friendly pipeline for symmetry analysis of image data. We first establish mathematically-desirable properties of $\mathcal{TI}$, ensuring that the method works for diverse data types with minimal preprocessing (Sect. 3). Building on this theoretical foundation, we present a general computational framework and accompanying software for measuring and comparing the symmetries of image data (GitHub 2025) (Sect. 4).

We then illustrate the versatility of the approach by applying our computational framework to two distinct systems: a large dataset of angiosperm flower images and a generated dataset of ferns (Sect. 5). In the floral dataset, our continuous, function-valued symmetry measure reveals something biologically meaningful: across flowering plants, distantly-related clades contain flowers that span the gamut of symmetry types, suggesting floral symmetries evolved independently across lineages and therefore carry real functional significance, likely tied to survival in specific ecological niches. The application to simulated ferns demonstrates the flexibility of our method, which can capture symmetry under arbitrary transformations. We investigate the fractal symmetry of ferns and show that fern symmetry also exists along a continuum. Together, these applications illustrate that our framework can detect subtle morphological features that may not be observable using traditional discrete classifications, allowing researchers to move towards developing more sophisticated statistical mod-

els that can correlate a continuous, function-valued symmetry measure to functional properties in organisms.

2 Introduction to Transformation Information ($\mathcal{TI}$)

Approximate symmetry can be measured by comparing an object to a transformed version of the object. When the transformed object is similar to the original, the object is approximately symmetric under that transformation. Objects of interest in biology may include concentrations of biochemicals, patterns of cell fates, or image datasets, which can all be represented with a distribution $\mu(x)$ that takes a spatial coordinate x as input and outputs a real number. Furthermore, we can transform $\mu(x)$ by composing it with an appropriate (invertible) transformation T_a. A distribution $\mu(x)$ transformed by T_a is given by the distribution $\mu(T_a^{-1}x)$, which we denote $T_a\mu$ (as in representation theory, where T_a can be thought of as a acting on μ (Golubitsky et al. (2012), Ch. 7). Transformation information (defined below) measures the divergence between μ and $T_a\mu$, yielding a value of zero if they are identical and a large value if they are very different. Thus, the value of $\mathcal{TI}$ provides a measure of deviation of the object from being (exactly) symmetric.

Definition 1 *(Transformation information)* Let $\mu : D \to \mathbb{R}_{>0}$ with $D \subset \mathbb{R}^n$ be a measurable and positive-valued function, and let $\{T_a : \mathbb{R}^n \to \mathbb{R}^n\}_{a\in\mathcal{A}}$ be a set of invertible transformations smoothly parameterized in a. Then the transformation information of μ under the family of transformations $\{T_a\}$ is given by

$$\mathcal{TI}[\mu, T_a] = \int_{\widetilde{D}} \mu(x) \log\left(\frac{\mu(x)}{T_a\mu(x)}\right) dx\,, \tag{1}$$

where $\widetilde{D} = D \cap T_a(D)$.

The measure is analogous to quantifying the relative entropy between μ and $T_a\mu$, which can be interpreted as the inefficiency in describing the system with the distribution $T_a\mu$ when the true distribution is μ. Relative entropy provides a conceptually-rich framework that connects symmetry with ideas from statistical mechanics and dynamical systems. Alternative definitions of $\mathcal{TI}$, which require μ to be normalized or alter the region of integration, have been considered and are discussed in Sect. A.1.

Remark 1 The assumption that μ is measurable guarantees that the integral is defined. In applications to data, this assumption is typically satisfied.

Remark 2 When data consists of images, a continuous signal is sampled on a discrete domain D. When D is discrete, $\widetilde{D}$ is empty for many transformations T_a, so $\mathcal{TI}[\mu, T_a]$ can vanish even for highly asymmetric data. We address this issue in Sect. 4.3.

Remark 3 To ensure that $\mathcal{TI}[\mu, T_a] < \infty$, we often require that D is compact (equivalently, closed and bounded) or that μ is bounded below by some positive value.

Remark 4 Although we primarily observe rotational and bilateral symmetry in biological systems, transformation information can measure the symmetry of an object

under any invertible transformation. We often assume that $\{T_a\}$ is a subset of the invertible affine transformations on $\mathbb{R}^n$, denoted $\mathrm{Aff}(n)$; however, we can certainly consider more general transformations such as the set of diffeomorphisms from $\mathbb{R}^n \to \mathbb{R}^n$.

Remark 5 An affine transformation on $\mathbb{R}^n$

$$T_a(x) = Ax + b$$

can be written in *homogeneous coordinates* with the $(n+1) \times (n+1)$ augmented matrix below.

$$\begin{pmatrix} A & b \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} T_a(x) \\ 1 \end{pmatrix}$$

In this notation, composition of affine transformations is equivalent to matrix multiplication. Furthermore, the determinant of an affine transformation is defined as the matrix determinant in homogeneous coordinates.

Definition 2 *(Local Minima and Maxima)* Let $f : X \to \mathbb{R}$ be a function on a topological space X. $x \in X$ is a local minimum (respectively, maximum) if there exists some neighborhood of x, $\mathcal{N}_x$ such that for every $y \in \mathcal{N}_x$, $f(x) \leq f(y)$ (respectively, $f(x) \geq f(y)$).

Definition 3 *(Approximate Symmetries and Maximal Asymmetries)* Let $\{T_a\}_{a \in \mathcal{A}}$ be transformations as introduced in Definition 1. Suppose that $\mathcal{A}$ is a topological space. Then T_s is an approximate symmetry (respectively, maximal asymmetry) if s is a local minimum (respectively, maximum) of $\mathcal{TI}[\mu, T_a]$ as a function of a (Fig. 1).

Remark 6 For isometries $T_a : D \to D$, approximate symmetries are generalizations of exact symmetries because $\mathcal{TI}[\mu, T_a] \geq 0$ and $\mathcal{TI}[\mu, T_a] = 0$ if and only if $\mu = T_a\mu$.

Remark 7 When thinking of $\mathcal{TI}[\mu, T_a]$ as a function of a, we will refer to the graph $\{(a, \mathcal{TI}[\mu, T_a])\} \subset \mathcal{A} \times \mathbb{R}_{\geq 0}$ as the $\mathcal{TI}$ *curve* (Fig. 1).

Remark 8 While approximate symmetries may themselves be a discrete set, the $\mathcal{TI}$ curve encodes more information. The full $\mathcal{TI}$ curve, rather than the approximate symmetries, is the primary object of our analysis, and the fPCA of Sect. 5 operates on these curves as elements of a function space.

The center of an object μ is important for considering reflections and rotations; we would like to rotate about the center and reflect across an axis that passes through the center. Heuristically, the center of an object is the point in the domain of μ from which we can observe the most symmetry. We define this formally below.

Definition 4 *(Symmetry Center)* Consider the continuous function $\mu : D \to \mathbb{R}_{>0}$. Suppose that $D \subset \mathbb{R}^n$ is compact and that $\mu > \delta$ for some $\delta > 0$. Let $\mathcal{TI}$ be defined as in Definition 1. Furthermore, assume that μ_c is defined as

$$\mu_c(x) = \begin{cases} \mu(x - c) & x \in D \cap (D + c) \\ \delta & x \in D \setminus (D + c)\,. \end{cases} \tag{2}$$

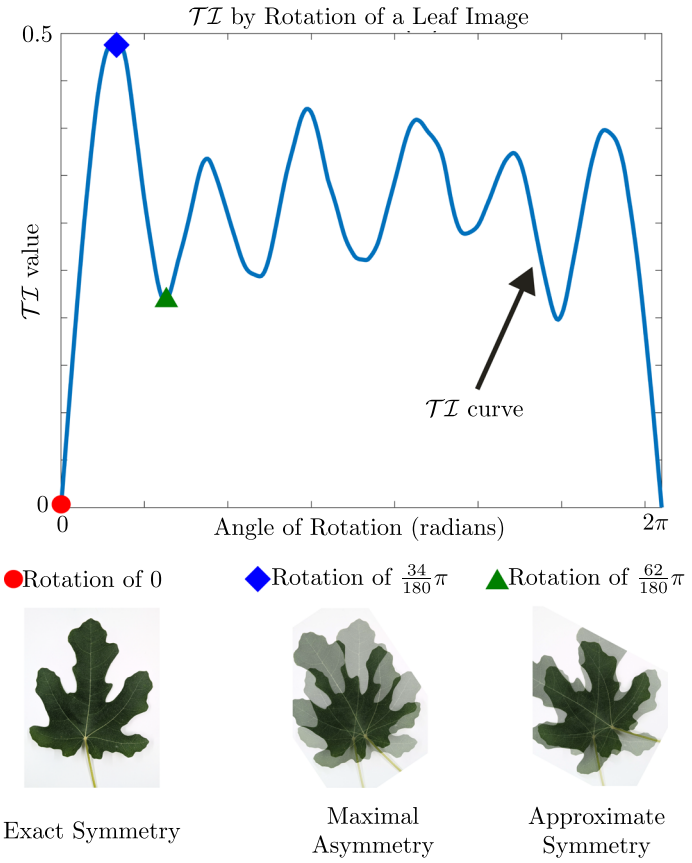


Fig. 1 We represent the leaf image by a positive-valued function μ and compute $\mathcal{TI}[\mu, R_\theta]$, where R_θ denotes rotation by angle θ (see Sect. 4). The resulting $\mathcal{TI}$ curve is shown above. Local minima and maxima of the curve identify approximate symmetries and maximal asymmetries, respectively. For example, $R_{(34\pi/180)}$ is the maximal asymmetry associated with the local maximum at $\theta = 34\pi/180$ (blue diamond), while $R_{(62\pi/180)}$ is the approximate symmetry associated with the local minimum at $\theta = 62\pi/180$ (green triangle). The corresponding rotated images (bottom) demonstrate low overlap for the maximal asymmetry and high overlap for the approximate symmetry

Now, define

$$M_c = \{a \in \mathcal{A} : a \text{ is a local minimum of } \mathcal{TI}[\mu_c, T_a]\},$$

and let

$$\overline{M}_c = M_c \setminus \{a \in \mathcal{A} : T_a = \mathrm{id}\}.$$

Then the symmetry center of μ is given by

$$\overline{x} := \mathrm{argmin}_{c \in D} \left\{ \inf_{a \in \overline{M}_c} \mathcal{TI}[\mu_c, T_a] \right\}. \tag{3}$$

In our computational pipeline, we use Definition 4 to identify the symmetry center of each image, finding the point from which we can observe the smallest $\mathcal{TI}$ value (see Fig. 2). However, there are various ways to define the "most symmetry." For example, one may consider the symmetry center to be the point with the largest number of approximate symmetries, with the lowest average $\mathcal{TI}$, or with the smallest min to max ratio. In our simulations, we find that each of these notions finds approximately the same symmetry center of the object.

We note one limitation of Definition 4 in practice: because $\mathcal{TI}$ is computed on $\widetilde{D} = D \cap T_a(D)$, candidate centers that produce little overlap between D and $T_a(D)$ under some transformation are evaluated on a smaller domain and can therefore attain spuriously small $\mathcal{TI}$ values, biasing the argmin in (3) towards such candidates. This is the same concern discussed in Sect. A.1, where we provide an alternative definition of $\mathcal{TI}$ that bypasses this issue but implies a slightly different notion of "symmetry." We show that the alternative definition of $\mathcal{TI}$ improves center estimation in Sect. A.3.

Another natural alternative is the intensity-weighted mean or centroid of the pixel coordinates

$$\overline{x} := \frac{\sum_{x \in D} x\mu(x)}{\sum_{x \in D} \mu(x)}, \tag{4}$$

which is far cheaper to compute. In our dataset it lies close to the symmetry centers obtained from Definition 4, and may be a reasonable substitute when computational cost is a concern (see Sect. A.3).

We nonetheless retain Definition 4 for our applications. The centroid measures the distribution of intensity rather than symmetry, so agreement is a property of these images rather than a general guarantee, and computing $\mathcal{TI}$ about the centroid yields strictly larger minima than computing it about the symmetry center. The centroid also yields a single point regardless of the symmetry transformation family. By contrast, the symmetry center locates the point from which the object exhibits the most symmetry and is adapted to the symmetry of interest.

Remark 9 We assume that μ is continuous and D is compact to ensure that the argmin in (3) exists. In our application to image data, D represents a finite set of pixels, meaning that the argmin will exist despite μ being discontinuous.

Remark 10 When searching for the symmetry center, we define $\overline{M}_c$ so that we only consider transformations T_a that are not equal to the identity map, id. If instead we considered M_c such that $T_a = \mathrm{id}$ for some $a \in M_c$, then for every c, we have that

$$\min_{a \in M_c} \mathcal{TI}[\mu_c, T_a] = \mathcal{TI}[\mu_c, \mathrm{id}] = 0.$$

implying that every $c \in D$ is a symmetry center, which is not useful for our analysis.

3 Mathematical Properties of $\mathcal{TI}$

A central requirement for any symmetry measure is that it should capture intrinsic structure rather than arbitrary choices in how an object or image is presented. In our

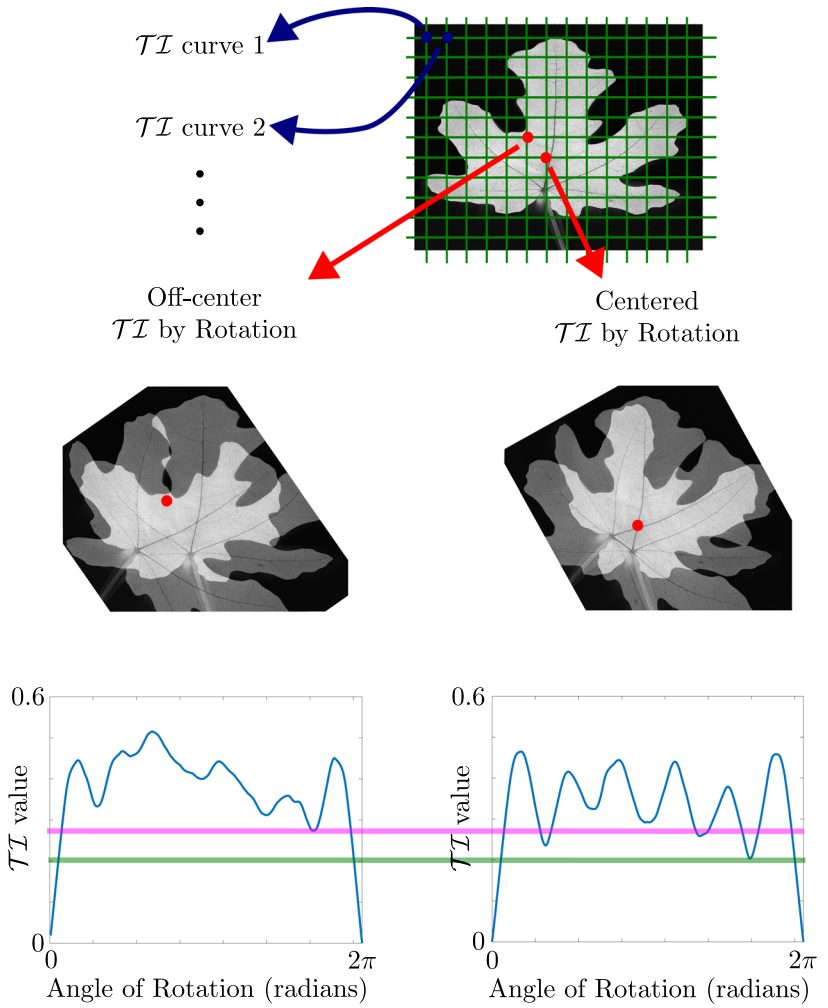


Fig. 2 To locate the symmetry center, we sample a grid of candidate points in the domain (shown in green). For each grid point, we compute the corresponding $\mathcal{TI}$ curve as if that point were the origin (see Sect. 4); two examples are shown below. We then record the smallest local minimum (excluding ones that correspond to the identity symmetry) for each potential symmetry center point. The point producing the smallest local minimum is then selected as the symmetry center of the image, shown by the red dot in the middle-right panel. In this case, the $\mathcal{TI}$ curve in the bottom right plot attains the smallest local minimum (green). For comparison, the $\mathcal{TI}$ curve obtained by rotating about the red point on the middle-left image has a higher minimum (pink)

context, this means that $\mathcal{TI}$ should be robust to affine transformations of the object of interest such as translation, rotation, reflection, and rescaling. This is especially important when analyzing large datasets, since inspecting each image individually is impractical. In this section, we show that two images representing the same object, but taken at different positions, orientations, or sizes produce $\mathcal{TI}$ curves that only differ in predictable, well-characterized ways. For example, when $\{T_a\}$ denotes rotations by

angle a, rotating an image does not change the $\mathcal{TI}$ curve despite the change in pixel arrangement. This property allows $\mathcal{TI}$ to reliably detect symmetries of images without requiring that they be perfectly aligned or scaled.

As before, to show that $\mathcal{TI}$ is robust, we assume that our object of interest μ is a measurable, real-valued function on $\mathbb{R}^n$, bounded below by some positive constant δ. The results, however, hold for objects or images defined on $D \subsetneq \mathbb{R}^n$ if there is a sufficiently large uniform background, as proven below (Proposition 1, Fig. 3). Furthermore, although we prove all results for functions on $\mathbb{R}^n$, similar results hold for functions on $\mathbb{Z}^n$, such as image data (see Sect. 4.3).

In Proposition 1 below, $g(x)$ defined on F represents the object of interest while the constant function m on $\mathbb{R}^n \setminus F$ represents the uniform background. If a domain D satisfies the specified properties, then $\mathcal{TI}$ is the same whether considering μ on the entire space or restricted to D.

Proposition 1 (Large background domain makes $\mathcal{TI}$ robust) *Let $\mu : \mathbb{R}^n \to \mathbb{R}$ be defined as follows:*

$$\mu(x) = \begin{cases} g(x) & x \in F \\ m & x \in \mathbb{R}^n \setminus F \end{cases}$$

for some $F \subset \mathbb{R}^n$ and some constant $m > 0$. Let $\{T_a : \mathbb{R}^n \to \mathbb{R}^n\}_{a \in \mathcal{A}}$ arbitrary family of invertible transformations. Furthermore, let $D \subset \mathbb{R}^n$ be such that $F \subset D$ and $T_a^{\pm}(F) \subset D$ for all $a \in \mathcal{A}$. Then

$$\mathcal{TI}[\mu, T_a] = \mathcal{TI}\left[\mu\big|_D, T_a\right] \text{ for all } a \in \mathcal{A}.$$

Proof Let $\widetilde{D} = D \cap T_a(D)$. Since $F \subset D$ and $T_a^{\pm}(F) \subset D$ for all a, $\mu\big|_{\mathbb{R}^n \setminus \widetilde{D}} \equiv m$ and $(T_a\mu)\big|_{\mathbb{R}^n \setminus \widetilde{D}} \equiv m$. Therefore, we have that

$$\begin{aligned} \mathcal{TI}[\mu, T_a] &= \int_{\mathbb{R}^n} \mu \log\left(\frac{\mu}{T_a\mu}\right) \\ &= \int_{\widetilde{D}} \mu \log\left(\frac{\mu}{T_a\mu}\right) + \int_{\mathbb{R}^n \setminus \widetilde{D}} \mu \log\left(\frac{\mu}{T_a\mu}\right) \\ &= \int_{\widetilde{D}} \mu \log\left(\frac{\mu}{T_a\mu}\right) + \int_{\mathbb{R}^n \setminus \widetilde{D}} m \log\left(\frac{m}{m}\right) \\ &= \int_{\widetilde{D}} \mu \log\left(\frac{\mu}{T_a\mu}\right) \\ &= \mathcal{TI}\left[\mu\big|_D, T_a\right]. \end{aligned}$$

□

Since an object with a sufficiently large uniform background has the same $\mathcal{TI}$ as an object defined on $\mathbb{R}^n$, we assume the object is defined on $\mathbb{R}^n$ for the rest of this section.

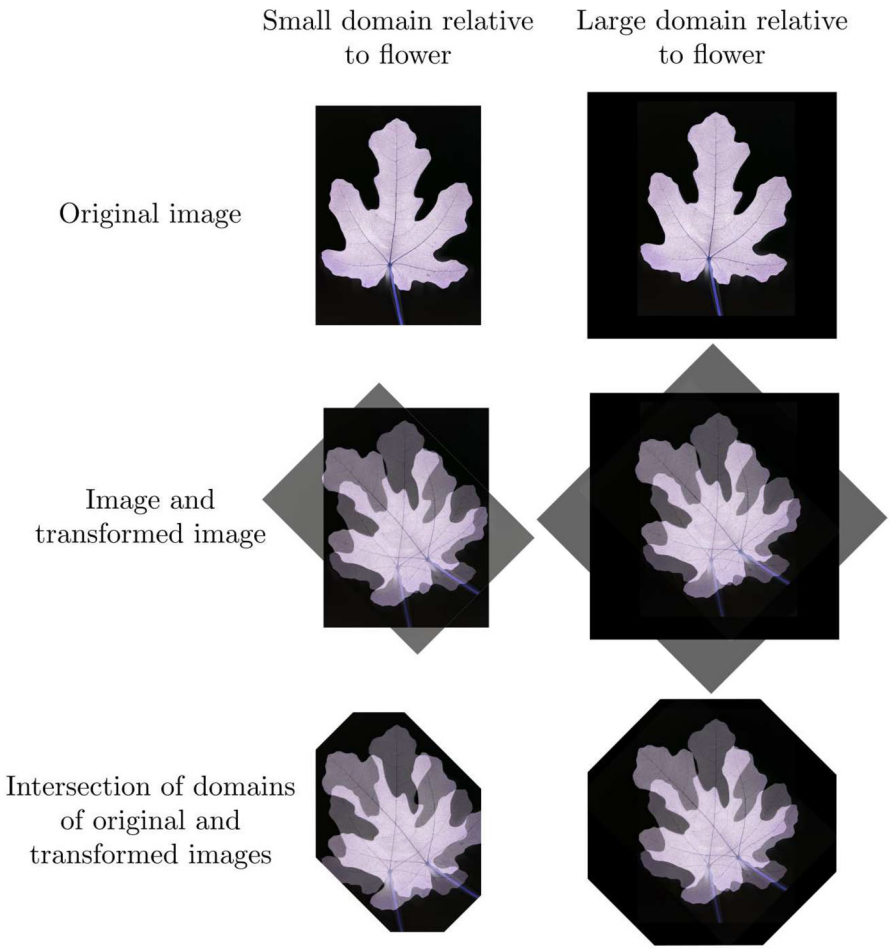


Fig. 3 Schematic illustrating Proposition 1. Left: when the domain is small relative to the leaf, the set $\widetilde{D}$ (bottom left) cuts off parts of the leaf, unlike in $\mathbb{R}^2$. Right: extending the uniform background ensures $\widetilde{D}$ contains both the original and transformed leaf; since the overlapping background contributes 0 to $\mathcal{TI}$, Proposition 1 shows that the $\mathcal{TI}$ results match those of a function defined on all of $\mathbb{R}^2$

3.1 Properties for General Transformations

To formally establish the robustness of $\mathcal{TI}$, we analyze its behavior when a transformation A is applied to the object μ. We consider the general case where A is an affine transformation and may represent a combination of translation, rotation, and scaling. We show that the $\mathcal{TI}$ curve of the transformed data $A\mu$ is not an arbitrary new function, but a predictable mapping of the original. This allows us to prove that the core information about the symmetry is preserved despite changes in the object's orientation or position.

The following identity characterizes how $\mathcal{TI}$ changes when an affine transformation acts on μ. Transforming the object of interest μ transforms the $\mathcal{TI}$ curve in a similar

way. Loosely speaking, the "shape" of the $\mathcal{TI}$ curve will stay the same when an affine transformation is applied to μ.

Remark 11 For an affine transformation $A \in \mathrm{Aff}(n)$, with determinant as defined in Remark 5, the quantity $|\det(A)|$ is the factor by which A scales volumes; for any measurable set $E \subset \mathbb{R}^n$ with image $AE := \{Ax : x \in E\}$, we have $\mathrm{vol}(AE) = |\det(A)|\,\mathrm{vol}(E)$. Writing χ_E for the indicator function of E and observing that $\chi_E(A^{-1}x) = \chi_{AE}(x)$, it follows that

$$\int_{\mathbb{R}^n} \chi_E(A^{-1}x)\,dx = |\det(A)| \int_{\mathbb{R}^n} \chi_E(x)\,dx.$$

Approximating by simple functions extends the identity to any integrable $f : \mathbb{R}^n \to \mathbb{R}$, so that

$$\int_{\mathbb{R}^n} f(A^{-1}x)\,dx = |\det(A)| \int_{\mathbb{R}^n} f(x)\,dx. \tag{5}$$

This is the change of variables $x \mapsto Ax$ applied in the proofs that follow, and it is the source of the factor $|\det(A)|$ in (6).

Proposition 2 ($\mathcal{TI}$ curve of transformed data is conjugate) *Let $\mu : \mathbb{R}^n \to \mathbb{R}_{>0}$ be a function bounded below by some positive constant δ. Furthermore, let $A, B \in \mathrm{Aff}(n)$. Then*

$$\mathcal{TI}[\mu, B] = \frac{1}{|\det A|}\mathcal{TI}[A\mu, ABA^{-1}]. \tag{6}$$

Proof

$$\begin{aligned}
\mathcal{TI}[A\mu, ABA^{-1}] &= \int A\mu \log\left(\frac{A\mu}{ABA^{-1}A\mu}\right) \\
&= \int \mu \circ A^{-1} \log\left(\frac{\mu \circ A^{-1}}{\mu \circ A^{-1} \circ A \circ B^{-1} \circ A^{-1}}\right) \\
&= \int \mu \circ A^{-1} \log\left(\frac{\mu \circ A^{-1}}{\mu \circ B^{-1} \circ A^{-1}}\right) \\
&= |\det(A)| \int \mu \log\left(\frac{\mu}{\mu \circ B^{-1}}\right) \qquad (x \mapsto Ax) \\
&= |\det(A)|\mathcal{TI}[\mu, B].
\end{aligned}$$

□

Since the "shape" of the $\mathcal{TI}$ curve is the same when an affine transformation acts on μ, some structure in the set of approximate symmetries of μ should also be preserved. We show below that transforming the data does not change the number of approximate symmetries. In fact, the set of approximate symmetries is conjugate.

Proposition 3 (Symmetries of transformed data are conjugate) *Let $\mu : \mathbb{R}^n \to \mathbb{R}_{>0}$ be a function bounded below by some positive constant δ. Suppose $\{T_a\} = \mathrm{Aff}(n)$ with the standard topology. If $\mathcal{S} \subset \mathrm{Aff}(n)$ is the set of approximate symmetries of μ, then for any $A \in \mathrm{Aff}(n)$, $A\mathcal{S}A^{-1}$ is the set of approximate symmetries of $A\mu$.*

Proof Let $M \in \mathrm{Aff}(n)$ be an approximate symmetry of μ (Definition 3). Then by definition, there exists some neighborhood of M, call it $\mathcal{N}_M$, such that for all $B \in \mathcal{N}_M$, $\mathcal{TI}[\mu, M] \leq \mathcal{TI}[\mu, B]$. Using (6), we can immediately see that for any $A \in \mathrm{Aff}(n)$ and any $B \in \mathcal{N}_M$, we also have that

$$\mathcal{TI}[A\mu, AMA^{-1}] \leq \mathcal{TI}[A\mu, ABA^{-1}]. \tag{7}$$

Consider the neighborhood of AMA^{-1} given by $\overline{\mathcal{N}} = A\mathcal{N}_M A^{-1}$. By (7), for every $C \in \overline{\mathcal{N}}$ we have that

$$\mathcal{TI}[A\mu, AMA^{-1}] \leq \mathcal{TI}[A\mu, C]. \tag{8}$$

Since there exists a neighborhood of AMA^{-1} given by $\overline{\mathcal{N}}$ such that for every element $C \in \overline{\mathcal{N}}$, (8) holds, AMA^{-1} is an approximate symmetry of $A\mu$ by definition. Since this holds $\forall M \in \mathrm{Aff}(n)$, we have shown that $A\mathcal{S}A^{-1}$ is a subset of the approximate symmetries of $A\mu$. We can prove the reverse inclusion in the same manner. □

We end this section by showing that scaling μ results in scaling the $\mathcal{TI}$ curve and thus does not change the set of approximate symmetries.

Proposition 4 (Scaling an object scales $\mathcal{TI}$) *Let $\mu : \mathbb{R}^n \to \mathbb{R}_{>0}$ be a function bounded below by some positive constant δ, and consider the scaling matrix $S = sI$ for $s \neq 0$, where I is the $n \times n$ identity matrix. Then for any $A \in \mathrm{Aff}(n)$,*

$$\mathcal{TI}[S\mu, A] = s^n \mathcal{TI}[\mu, A]$$

Proof

$$\begin{aligned}
\mathcal{TI}[S\mu, A] &= \int S\mu \log\left(\frac{S\mu}{AS\mu}\right) \\
&= \int \mu \circ S^{-1} \log\left(\frac{\mu \circ S^{-1}}{\mu \circ S^{-1} \circ A^{-1}}\right) \\
&= |\det(S)| \int \mu \log\left(\frac{\mu}{\mu \circ S^{-1} \circ A^{-1} \circ S}\right) && (x \mapsto Sx) \\
&= s^n \int \mu \log\left(\frac{\mu}{\mu \circ A^{-1}}\right) && (S \text{ commutes}) \\
&= s^n \mathcal{TI}[\mu, A].
\end{aligned}$$

□

Collectively, these results demonstrate that while affine transformations may shift or scale the data, the resulting $\mathcal{TI}$ curves and their approximate symmetries remain conjugate to the original. This mathematical consistency ensures that the intrinsic symmetry of an object, as measured by $\mathcal{TI}$, is preserved under deformation, providing a reliable foundation for empirical analysis. Building on this general framework, we now focus on rotational and reflectional symmetries, which are particularly important in biological systems.

3.2 Properties for Rotations and Reflections

In most biological applications, we are interested in rotational and reflectional symmetries given by the transformations $\{R_\theta\}$ for $\theta \in [0, 2\pi)$ and $\{\kappa_\theta\}$ for $\theta \in [0, \pi)$ (with the subspace topology inherited from $\mathbb{R}$). In homogeneous coordinates (Remark 5), reflections and rotations in $\mathbb{R}^2$ take the form

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{9}$$

$$\kappa_\theta = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) & 0 \\ \sin(2\theta) & -\cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{10}$$

Note that reflections only need to be defined for $\theta \in [0, \pi)$ because $\kappa_\theta = \kappa_{\theta+\pi}$.

Considering these transformations leads to special invariance and equivariance properties of the $\mathcal{TI}$ curve when we reorient the object μ by taking $R_\varphi\mu$ or $\kappa_\varphi\mu$. To describe these properties, we fix a function μ and refer to the $\mathcal{TI}$ curves of $\mathcal{TI}[\mu, R_\theta]$ and $\mathcal{TI}[\mu, \kappa_\theta]$ by $\mathcal{TI}_{\text{rot}}(\theta)$ and $\mathcal{TI}_{\text{ref}}(\theta)$, respectively. We also periodically extend $\mathcal{TI}_{\text{ref}}(\theta)$ so that both $\mathcal{TI}$ curves are defined on $[0, 2\pi)$.

Next, we define the operators

$$F_{\text{rot}} : \mu \mapsto \mathcal{TI}_{\text{rot}}(\theta), \tag{11}$$

$$F_{\text{ref}} : \mu \mapsto \mathcal{TI}_{\text{ref}}(\theta). \tag{12}$$

Lastly, define the following transformations on $[0, 2\pi)$:

$$\overline{R}_\varphi(\theta) = \theta + \varphi \mod 2\pi, \tag{13}$$

$$\overline{\kappa}_\varphi(\theta) = 2\varphi - \theta \mod 2\pi, \tag{14}$$

and as before denote the transformed versions of the $\mathcal{TI}$ curves as

$$\overline{R}_\varphi\mathcal{TI}(\theta) := \mathcal{TI}\left(\overline{R}_\varphi^{-1}(\theta)\right),$$

$$\overline{\kappa}_\varphi\mathcal{TI}(\theta) := \mathcal{TI}\left(\overline{\kappa}_\varphi^{-1}(\theta)\right).$$

See Fig. S1 for examples of these transformations. We then have the following result on how $\mathcal{TI}$ commutes with rotation and reflection transformations.

Proposition 5 (Equivariance of F_{rot} and F_{ref} under rotation and reflection of μ) *Let $\mu : \mathbb{R}^2 \to \mathbb{R}_{>0}$ be some function bounded below by a positive constant δ. Furthermore, let F_{rot}, F_{ref}, $\overline{R}_\varphi$, and $\overline{\kappa}_\varphi$ be defined as in (11)-(14). Then the following equations hold for every $\varphi \in [0, 2\pi)$.*

$$F_{\text{rot}}(R_\varphi\mu) = F_{\text{rot}}(\mu) \tag{15}$$

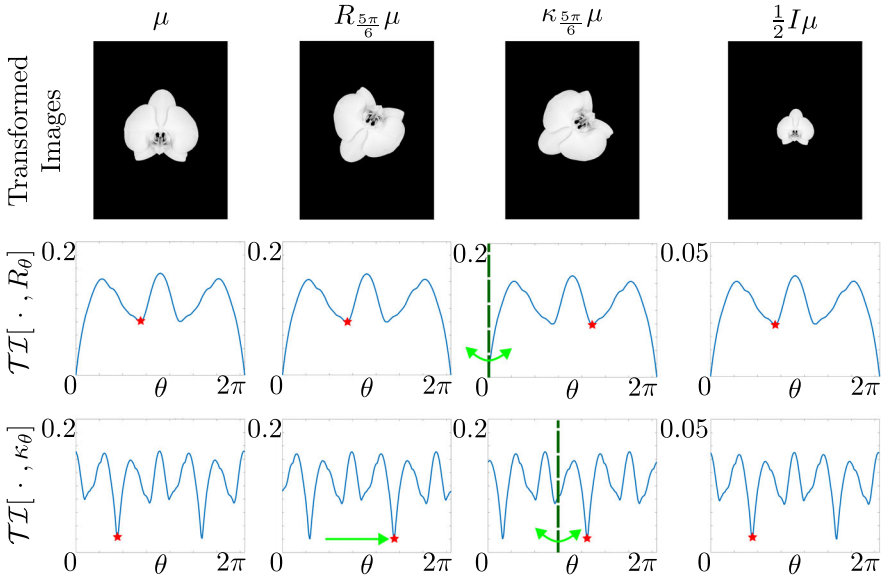


Fig. 4 Effect of image transformations on $\mathcal{TI}$. In row 2 (rotational symmetries), the $\mathcal{TI}$ curve is invariant under rotation (col. 2), reflected across $\theta = 0$ modulo 2π (green line) under image reflection (col. 3), and scaled by $(1/2)^2$ under image scaling (col. 4). In row 3 (reflection symmetries), rotation by $5\pi/6$ translates the $\mathcal{TI}$ curve by the same amount (col. 2), reflection across $\theta = 5\pi/6$ the curve across that axis modulo 2π (green line) (col. 3), and scaling again multiplies the $\mathcal{TI}$ curve by $(1/2)^2$ (col. 4). Red stars highlight points that help visualize the transformations. Small variations in $\mathcal{TI}$ curves arise from numerical error and boundary effects

$$F_{\text{rot}}(\kappa_\varphi \mu) = \overline{\kappa}_0 F_{\text{rot}}(\mu) \tag{16}$$

$$F_{\text{ref}}(R_\varphi \mu) = \overline{R}_\varphi F_{\text{ref}}(\mu) \tag{17}$$

$$F_{\text{ref}}(\kappa_\varphi \mu) = \overline{\kappa}_\varphi F_{\text{ref}}(\mu). \tag{18}$$

It follows that:

1. If $\mathcal{S}$ is the set of rotational approximate symmetries of μ, then $\mathcal{S}$ is the set of rotational approximate symmetries for $R_\varphi \mu$.
2. If $\mathcal{S}$ is the set of rotational approximate symmetries of $\kappa_\varphi \mu$, then $\overline{\kappa}_0 \mathcal{S}$ is the set of rotational approximate symmetries for $\kappa_\varphi \mu$.
3. If $\mathcal{S}$ is the set of reflectional approximate symmetries of μ, then $\overline{R}_\varphi \mathcal{S}$ is the set of reflectional approximate symmetries of $R_\varphi \mu$
4. If $\mathcal{S}$ is the set of reflectional approximate symmetries of μ, then $\overline{\kappa}_\varphi \mathcal{S}$ is the set of reflectional approximate symmetries of $\kappa_\varphi \mu$.

Figure 4 shows an illustration of Proposition 5 for a flower image. In Sect. A.2, we provide the proofs for these statements.

4 Computational Pipeline

We now describe our MATLAB computational pipeline for calculating the transformation information of image data. The pipeline consists of the following steps, which we explain in further detail in this section:

1. Preprocessing images.
2. Representing an image with a real-valued function $\overline{\mu}$ on a finite domain.
3. Computing $\mathcal{TI}$ of image data.
4. Centering $\overline{\mu}$ to obtain a new function μ.

4.1 Preprocessing Images

$\mathcal{TI}$ is designed to reliably extract symmetry information with minimal user input. Common imaging techniques, such as fluorescence microscopy or slide imaging, provide the uniform backgrounds necessary for high-fidelity analyses, meaning $\mathcal{TI}$ can often be applied directly to raw data.

In our applications (Sect. 5), we ensure that the images have uniform black backgrounds through either controlled imaging conditions or straightforward digital background removal. In cases where images have light backgrounds or backgrounds with significant variations in pixel intensity, simple preprocessing steps such as taking the negative, smoothing, or thresholding can be employed to isolate the object of interest. These minor adjustments ensure that the resulting $\mathcal{TI}$ curves faithfully reflect the symmetry of the data. For further details on these implementations, see Sect. A.4 and Fig. S2.

4.2 Representing Each Image with A Real-Valued Function $\overline{\mu}$ on a Finite Domain

An image can be represented as a an array of three $\ell \times w$ matrices R, G, and B that represent the intensities of red, green, and blue light at each pixel on a scale from 0 to 255. We convert an image to grayscale by taking a weighted sum of R, G, and B. The result is a single grayscale matrix $M = (m_{ij})$ with $0 \leq m_{ij} \leq 255$ for any i, j. We define the representative function for the image as

$$\overline{\mu}(i, j) = m_{ij} + \delta \tag{19}$$

for some $\delta > 0$ (Fig. S2). We add a small δ value to each image to ensure that the function is positive on its domain, as is required to compute $\mathcal{TI}$. By Proposition 6 (Lipshitz continuity of $\mu \mapsto \mathcal{TI}[\mu, T_a]$), adding a small δ will only change the $\mathcal{TI}$ curve by a small amount. In our applications to flowers and ferns, we take $\delta = 1$. Note that the domain D of $\overline{\mu}$ is the finite set

$$D = \{i : 1 \leq i \leq \ell\} \times \{j : 1 \leq j \leq w\}. \tag{20}$$

Remark 12 The weighted sum of R,G, and B matrices can be adjusted to emphasize certain colors. For example, Gandhi et al. (2021) emphasize green by taking the

brightness of green minus the average pixel brightness ($M = (2G - R - B)/3$) when analyzing approximate symmetries of the unicellular freshwater desmid *Micrasteria*, which is a type of green algae. In Fig. S3, we demonstrate this method for an image of a leaf on dirt.

4.3 Computing $\mathcal{TI}$ of Image Data

In computing the $\mathcal{TI}$ of image data, we note that a grayscale image $\overline{\mu}$ (19) is defined on a finite domain D; therefore, the set $\widetilde{D} = D \cap T_a(D)$ will almost always be empty (see Fig. 5), and $\mathcal{TI}$ will thus be zero. To avoid this issue, we redefine $T_a\overline{\mu}$ on a subset of the lattice of integers by using linear interpolation of $\overline{\mu}$ between the points in $T_a(D)$. For the remainder of this manuscript, we will abuse notation and assume that we have already performed this process when considering the transformed image $T_a\overline{\mu}$, so that it is defined on a set that we denote as $\widetilde{T}_a(D) \subset \mathbb{Z}^2$. Then, for the application to images, we can rewrite $\mathcal{TI}$ as

$$\mathcal{TI}[\overline{\mu}, T_a] = \sum_{(i,j)\in\widetilde{D}} \overline{\mu}(i,j) \log\left(\frac{\overline{\mu}(i,j)}{T_a\overline{\mu}(i,j)}\right)$$

where $\widetilde{D} = D \cap \widetilde{T}_a(D)$.

4.4 Centering $\overline{\mu}$ to Obtain a New Function μ

To analyze an image's invariance under rotations and reflections, we must first establish its symmetry center (Definition 4). Given a grayscale image represented as a function $\overline{\mu} : D \to \mathbb{R}$ (Sect. 4.2) and a set of transformations $\{T_a\}$, the symmetry center can be found using Definition 4.

Numerically, we begin by selecting a grid of candidate symmetry centers $c \in D$. Candidate centers are restricted to the central 25% of the image since candidates near the boundary produce trivial symmetries (Sect. 2, Fig. 11). For each candidate, we translate the image so that c is mapped to the origin, obtaining the centered image $\overline{\mu}_c$. Fixing the center, we compute transformation information $\mathcal{TI}[\overline{\mu}_c, T_a]$ and find the approximate symmetries, excluding the identity symmetry. The strongest approximate symmetry, given by the smallest local minimum of the $\mathcal{TI}$ curve, is stored (Fig. 2). Repeating this procedure over all candidate symmetry centers c produces a scalar value for each candidate. The true symmetry center is then defined as the point c at which this value achieves its minimum. We then set $\mu := \overline{\mu}_c$ to be the final function for our analysis. Details of the implementation, including discretization choices, and peak-detection procedures are provided in the accompanying code repository (GitHub 2025).

In our applications, we compute the symmetry center for two distinct transformation groups: reflections $\{\kappa_\theta\}$ across an axis at angle θ, and rotations $\{R_\theta\}$ by angle θ (9),(10). We denote the resulting reflectional symmetry center as $\overline{x}^\kappa$ and the rotational symmetry center as $\overline{x}^R$. Our pipeline produces four specific $\mathcal{TI}$ curves:

1. Rotation about $\overline{x}^R$: $\mathcal{TI}[\overline{\mu}_{\overline{x}^R}, R_\theta]$.

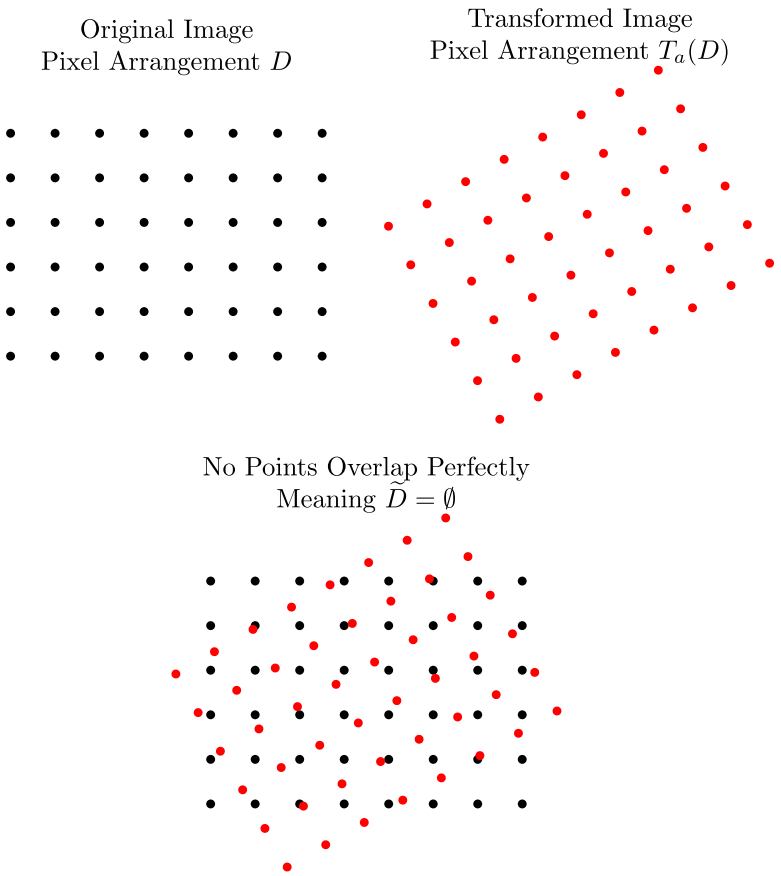


Fig. 5 An image $\overline{\mu}$ as in Sect. 4.2 is supported on a lattice of pixels D (top left). Thus, the transformed image $T_a\overline{\mu}$ is supported on a transformed domain $T_a(D)$ (top right). Taking the intersection of these domains (bottom), no pixels align perfectly, so $\widetilde{D} = D \cap T_a(D) = \emptyset$

2. Reflection about $\overline{x}^R$: $\mathcal{TI}[\overline{\mu}_{\overline{x}^R}, \kappa_\theta]$.
3. Rotation about $\overline{x}^\kappa$: $\mathcal{TI}[\overline{\mu}_{\overline{x}^\kappa}, R_\theta]$.
4. Reflection about $\overline{x}^\kappa$: $\mathcal{TI}[\overline{\mu}_{\overline{x}^\kappa}, \kappa_\theta]$.

Remark 13 As a consequence of Proposition 7, if an object has perfect reflection symmetry, then the $\mathcal{TI}$ curves for rotations and reflections contain identical information.

5 Applications

We apply the computational framework outlined in Sect. 4 to two datasets of flower and respectively fern images. The resulting collection of $\mathcal{TI}$ curves characterizes the symmetries present in each image, enabling the use of functional data analysis to explore correlations between symmetry and specific functional properties of the specimens. As an initial step, we project each image's $\mathcal{TI}$ curve onto the first two

functional principal components (i.e. *functional principal component analysis* (fPCA)) using the PACE toolbox (Yao et al. 2003) for the flower $\mathcal{TI}$ curves and Shi et al. (2022)'s code for the fern $\mathcal{TI}$ curves. We use two methods as each is optimized for functions on $\mathbb{R}$ and $\mathbb{R}^2$, respectively.

In the floral dataset, the fPCA plots (Fig. 6) reveal that $\mathcal{TI}$ curves overlap across evolutionarily distinct clades, suggesting that symmetry types evolved independently and are therefore functionally significant. In the fern dataset, we investigate fractal symmetry, demonstrating that $\mathcal{TI}$ can represent arbitrary symmetry transformations. In both applications, the visualizations demonstrate that the symmetries are distributed continuously in the function space of $\mathcal{TI}$ curves, rather than discrete clusters based on idealized symmetry categories.

5.1 Floral Symmetries

We begin by investigating the symmetries of flowers from the *Plant Gateway* poster (Byng et al. 2018). Using the computational pipeline described in Sect. 4, we generate $\mathcal{TI}$ curves by rotation and reflection for each image. We then perform fPCA and present the results for rotational $\mathcal{TI}$ in Fig. 6. The analogous results for reflectional $\mathcal{TI}$ are included in Sect. A.5. For our dataset, the remaining two $\mathcal{TI}$ curves ($\mathcal{TI}[\overline{\mu}_{\overline{x}^R}, \kappa_\theta]$ and $\mathcal{TI}[\overline{\mu}_{\overline{x}^\kappa}, R_\theta]$, Sect. 4.4) are redundant since the rotational center $\overline{x}^R$ and reflectional center $\overline{x}^\kappa$ agree closely (Sect. A.3). The rotational fPCA results show that floral symmetry exists along a continuum rather than within the traditional discrete categories of bilateral (zygomorphic) and radial (actinomorphic) symmetries. Notably, we see in the bottom panel of Fig. 6 that points with different colors, representing flowers from different clades, are intermixed. This shows that the range of exhibited symmetries are represented within each clade, providing evidence that the symmetries evolved independently and are not merely incidental but adaptive. We come to the same conclusion by analyzing reflectional symmetry (Sect. A.5).

The fPCA also effectively captures intuitive morphological differences; for example, the fPCA of rosids presents two diverging branches. Flowers towards the tip of the upper branch have 5-fold rotational symmetry, while those at the tip of the lower branch have 4-fold rotational symmetry. Towards the left of the plot, the flowers are characterized by indistinguishable petals and small variations in symmetry. These trends, however, are not visually apparent from the embeddings of reflectional $\mathcal{TI}$ curves (Sect. A.5).

The continuum of symmetry can be represented using a variety of visualization techniques. We choose fPCA here because it is a classical and well-established approach for comparing curves; however, there may be benefits to other visualizations depending on the application of interest. In Fig. 7, we present two such alternatives. Both methods begin by vectorizing the $\mathcal{TI}$ curves as $(\mathcal{TI}[\mu, R_{\theta_1}], ..., \mathcal{TI}[\mu, R_{\theta_N}])$ for some $0 \leq \theta_1 < ... < \theta_N \leq 2\pi$. The top panel applies standard *principal component analysis* (PCA) to these vectors and projects them onto the first two principal components. The bottom panel computes the GW_τ distance (Kravtsova et al. 2023) between $\mathcal{TI}$ vectors and embeds them in the plane using *multidimensional scaling* (MDS). GW_τ is an optimal transport metric derived from the Gromov-Wasserstein

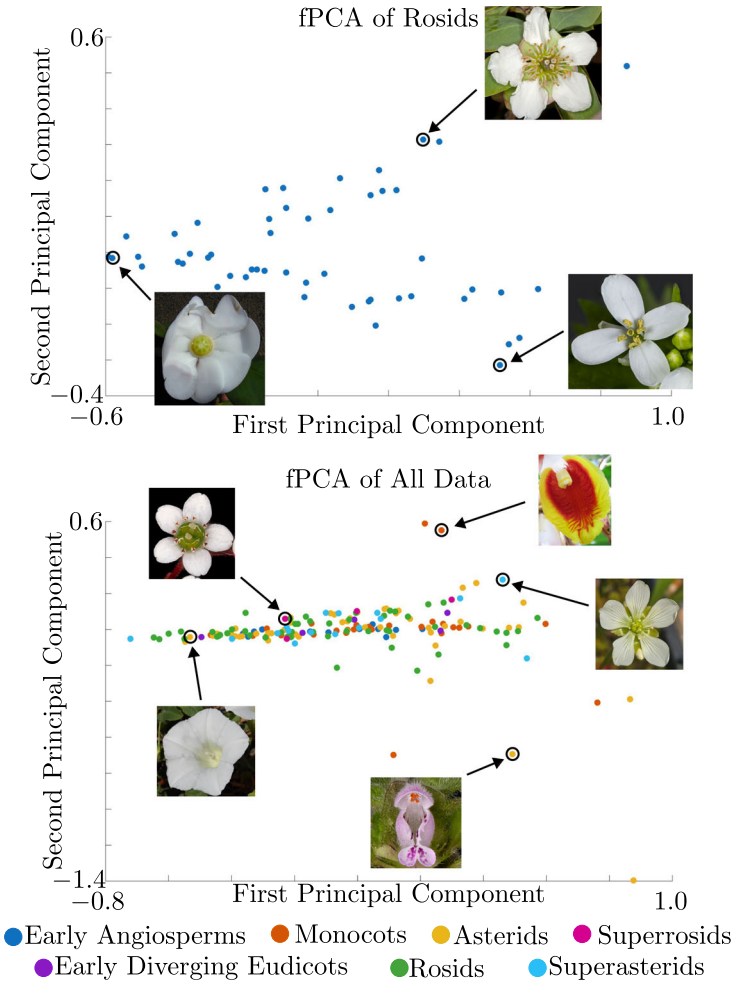


Fig. 6 fPCA of $\mathcal{TI}$ curves by rotation about the rotational symmetry center ($\mathcal{TI}[\mu_{\overline{x}R}, R_\theta]$) reveals a continuous landscape of symmetries. Top: rosid curves projected onto the first two principal components, capturing 85.9% of the variance. Two branches emerge corresponding to 5-fold and 4-fold symmetries. Where the branches converge, flowers have indistinct petals and small variations in their $\mathcal{TI}$ curves. Bottom: all projected $\mathcal{TI}$ curves, color-coded by clade, capturing 71.9% of the variance. Flowers with radial symmetry cluster along a central axis, transitioning from indistinct to well-defined petals from left to right, while flowers off the axis exhibit bilateral symmetry. The flower images presented on the graphs are stock images that closely match the image data used for our analysis (Descouens 2022; Game 2017; Starr and Starr 2026; North Queensland Plants 2026; Reese 2024; falcon_311 2025; Carr 2005; Edison and Ford Winter Estates 2010; Smith 2024)

distance. GW_τ calculates the shape difference between two ordered data sets by representing data sets as metric-measure spaces. This measure consistently outperforms standard measures including Euclidean and Dynamic Time Warping (DTW) while being computationally efficient.

A systematic comparison of the strengths and limitations of these visualization approaches is left to future work.

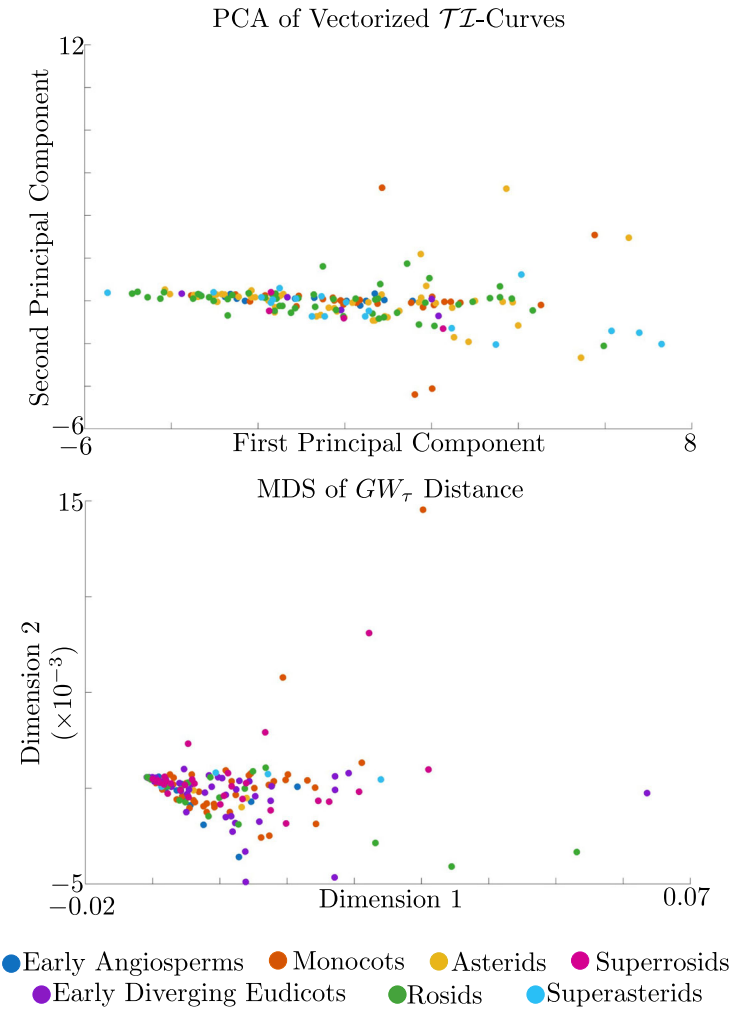


Fig. 7 Alternative visualizations of symmetry landscapes. After vectorization, $\mathcal{TI}$ curves are visualized using two approaches: PCA projection onto the first two principal components (top) and MDS embedding based on GW_τ distances (bottom). Both representations capture the overlap in symmetry types across clades and the continuous variations in symmetry that were present in the fPCA

5.2 Fern Symmetries

The $\mathcal{TI}$ framework accommodates various forms of symmetry. To demonstrate this, we analyzed the self-similarity (fractal symmetry) of simulated fern images under translation and scaling. Each simulated fern is the attractor of an iterated function system, where we vary a parameter to get different simulations (see Sect. A.6 for details).

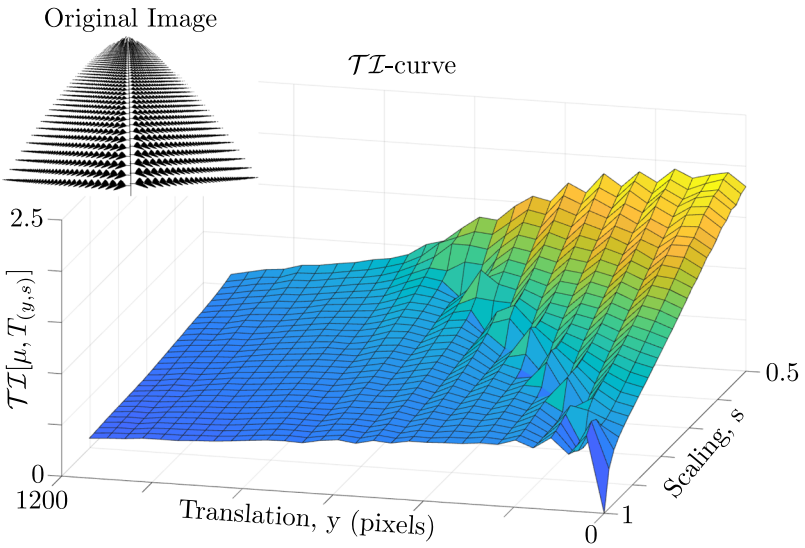


Fig. 8 The $\mathcal{TI}$ curve for the fern (top left) under vertical translation and scaling. Since the transformation $T_{(s,t)}$ relies on two parameters, the $\mathcal{TI}$ curve is a 2-dimensional surface

The following set of transformations, expressed in homogeneous coordinates, represents the composition of scaling and translation along the positive y-axis:

$$T_{(y,s)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ for } (y, s) \in [0, L] \times (0, 1] .$$

Given that the transformation relies on two parameters, the resulting $\mathcal{TI}$ "curve" is a mapping from $\mathbb{R}^2 \to \mathbb{R}$, as shown in Fig. 8.

We applied fPCA to these surfaces using the algorithm from Shi et al. (2022); the resulting projection onto the first two principal components is shown in Fig. 9. Complementing our observations in flowers, the ferns exhibit a symmetry continuum. There is a clear progression of symmetries, ranging from Barnsley-type ferns in the top right to *Cyclosorus*-type ferns in the bottom left, suggesting that future correlational studies should treat symmetry as a gradient rather than a discrete variable.

We note that the fPCA plot in Fig. 9 shows a linear, ordered curve. It is not obvious that this should be the case, i.e., that the linear family of IFS's outlined in Sect. A.6 will trace out a linear curve in the space of $\mathcal{TI}$ curves, since the map from any function μ to the associated $\mathcal{TI}$ curve is nonlinear. We also may not expect a one-parameter family of IFS's to trace a one-dimensional curve in the space of $\mathcal{TI}$ curves, since the standard continuity argument does not necessarily apply here. In particular, despite the attractor A_p of the IFS varying continuously in p with respect to the Hausdorff metric, the function $p \mapsto \chi_{A_p}$ (where χ_{A_p} is the indicator function of A_p) is not necessarily a continuous map into L^1, meaning we cannot directly compose the operations and apply Proposition 6. It is therefore an open question whether the map from the parameter p

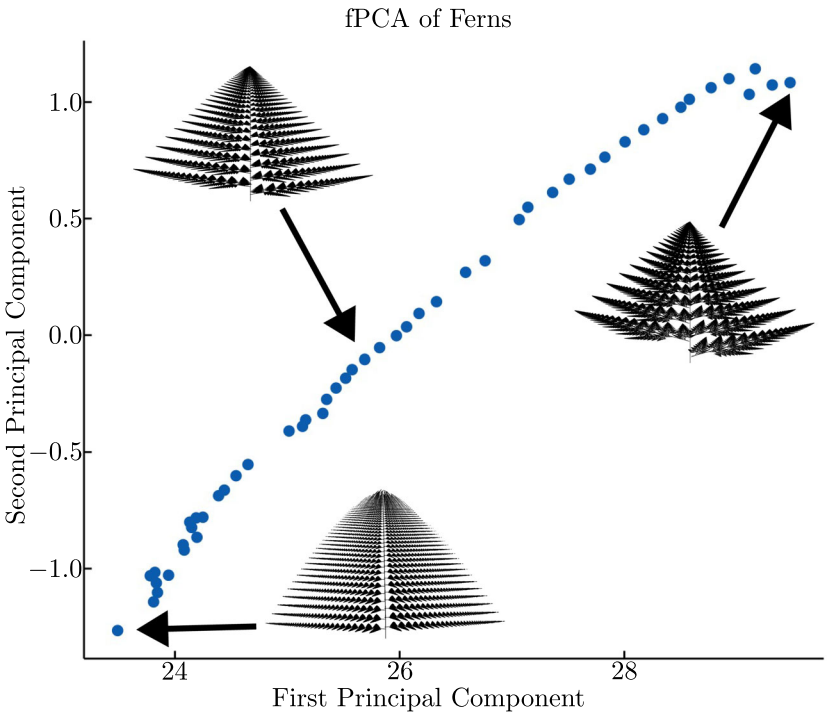


Fig. 9 Functional principal component analysis of fern $\mathcal{TI}$ curves. The first two principal components show a clear progression of symmetries, ranging from Barnsley-type ferns in the top right to *Cyclosorus* in the bottom left

to the associated $\mathcal{TI}$ curve is continuous in this case. We thus interpret the smooth linear progression in Fig. 9 as an empirical observation about this family of shapes rather than a consequence of the fern construction.

We also note that the set of fractals chosen here is relatively simple compared to real ferns, allowing us to analyze a two-parameter family of transformations to capture the fern symmetry. In more general settings, the analysis would require the exploration of symmetries in the full six-parameter family of affine transformations, which is possible with our approach, albeit computationally expensive.

6 Discussion

We used $\mathcal{TI}$ to represent symmetry as a broad, continuous spectrum of values and developed a user-friendly computational tool to measure and visualize the symmetries of image data. Applying this tool to angiosperm flower images recovered a spectrum of symmetries spanning evolutionarily-distinct clades. This revealed structure that traditional discrete classifications, such as radial versus bilateral symmetry, would have obscured and thus enabled biologically-meaningful conclusions about convergent evolution.

Traditional methods for measuring symmetry rely on user-selected landmarks. Our method still requires the user to select a family of symmetry transformations of interest and to ensure that the object is isolated against a uniform background. Despite this, the method replaces the difficult, subjective steps of landmark identification with more tractable and unbiased ones. No landmarks or homologous points are required, and the symmetry center, which would otherwise be a user-selected quantity, is detected automatically using computation (Definition 4). We demonstrate the fidelity of the method by investigating $\mathcal{TI}$'s mathematical properties such as invariance, equivariance, and continuity. Equivariance and invariance ensure that the symmetry measure is agnostic to image orientation, while continuity ensures resilience to noise and measurement error. Together, these properties provide evidence that our method does not require image alignment or landmarking, and that the preprocessing steps required can be automated with standard tools (Sect. A.4).

Building on these theoretical guarantees, the computational pipeline measures and visualizes continuous symmetry across an image dataset: for each image, it locates the symmetry center, computes the $\mathcal{TI}$ curve, and represents the symmetry as a two-dimensional data point via functional principal component analysis (fPCA). Representing symmetry as a function rather than a scalar or a categorical descriptor preserves more geometric information about each specimen, and the resulting "symmetry space" supports quantitative comparison across an entire dataset.

For example, in the fPCA plot of the floral images (Fig. 6), we observe diverse symmetry types, yet there is substantial overlap of symmetry types between clades. This overlap suggests that similar symmetries arose independently in different evolutionary lineages. While the overlap alone does not fully establish selective mechanisms, it provides strong evidence that similar symmetries may be favored in certain environmental conditions and motivates further investigation into the role of natural selection of symmetry. It would not be possible to see this evidence without a continuous measure of symmetry.

The $\mathcal{TI}$ representation of symmetry also provides a foundation for testing correlations between symmetry and other quantitative traits. We anticipate that the continuous symmetry metric given by $\mathcal{TI}$ will offer greater predictive power than traditional categorical classifications. The integration of $\mathcal{TI}$ into predictive models has already shown promise in other contexts. For instance, Rostami et al. (2023) demonstrated that machine learning classifiers perform significantly better when incorporating information from $\mathcal{TI}$. They demonstrated that "coarse-graining" types of actin networks based on shared symmetry led to much higher classification accuracy than classification based on the actin networks alone. In future work, we plan to use the framework proposed here to perform similar symmetry-informed classification in broader contexts. One promising application is to regular patterns of vegetation growth in dryland ecosystems (see, e.g., Gandhi et al. (2019)), where $\mathcal{TI}$ may help inform the classification of spatial patterns and quantify the deviation from spatial periodicity (Kästner et al. 2024).

Several limitations of this work also provide directions for future research. Our current computational framework is optimized for two-dimensional images with uniform backgrounds. Automated background removal techniques could improve the applications to general images. Furthermore, extending the framework to three-dimensional

data, such as point clouds or Z-stacks (in confocal microscopy), will be necessary to capture the full complexity of many biological structures. More generally, the $\mathcal{TI}$ framework can be applied beyond spatial symmetries, including higher-dimensional or nonphysical datasets where symmetry is not visually apparent yet still meaningful. Additionally, while we have established the descriptive strength of $\mathcal{TI}$, the development of statistical models linking these curves to functional or mechanistic properties in biology remains an open challenge.

Such models would benefit from understanding how properties of a family of shapes are represented in the space of $\mathcal{TI}$ curves. The fern family is instructive since the family of IFS's has a linear parameterization in p, yet our framework does not predict that the fPCA plot of $\mathcal{TI}$ curves would also be linear (as observed). Identifying conditions under which such properties of the symmetry can be inferred from the underlying model will provide a foundation for understanding how a model of a shape influences functional or mechanistic properties through symmetry.

Overall, our findings suggest that symmetry, which is traditionally treated as a qualitative, categorical trait, can be successfully transformed into a quantitative, continuous variable. By developing a high-throughput method that treats symmetry as a function rather than a scalar or categorical value, we enable better integration of symmetry into statistical and dynamical systems modeling and provide a more nuanced approach to investigating the prominent role of symmetry in biology.

A Appendix

A.1 Alternative Definitions of $\mathcal{TI}$

In addition to the hypotheses in Definition 1 for transformation information, Ciocanel et al. (2025) require $\mu : D \to \mathbb{R}_{>0}$ to be normalized. That is, it must satisfy $\int_D \mu(x)\, dx = 1$. From the definition of $\mathcal{TI}$, we can easily see that normalizing μ amounts to scaling the $\mathcal{TI}$ curve. Indeed, if a distribution $\overline{\mu}$ is normalized by a scaling factor $s \in \mathbb{R}$ so that $\mu := s\overline{\mu}$, then for any transformation $A \in \mathrm{Aff}(n)$,

$$\begin{aligned}
\mathcal{TI}[\mu, A] &= \mathcal{TI}[s\overline{\mu}, A] \\
&= \int_{\widetilde{D}} s\overline{\mu} \log\left(\frac{s\overline{\mu}}{A(s\overline{\mu})}\right) \\
&= s\int_{\widetilde{D}} \overline{\mu} \log\left(\frac{s\overline{\mu}}{sA\overline{\mu}}\right) \\
&= s\int_{\widetilde{D}} \overline{\mu} \log\left(\frac{\overline{\mu}}{A\overline{\mu}}\right) \\
&= s\mathcal{TI}[\overline{\mu}, A].
\end{aligned}$$

The decision to use normalization thus depends on the analytical goal. We recommend applying normalization if the objective is to compare an object's shape independent of its size. Conversely, normalization should be omitted when size disparity between

objects is a meaningful component of the data analysis, particularly when comparing $\mathcal{TI}$ curves using data visualization methods that compare $\mathcal{TI}$ curves component-wise. It is important to note, however, that the approximate symmetries and maximal asymmetries are identical regardless of whether normalization is applied.

A potential issue with the $\mathcal{TI}$ definition is that integrating over the domain $\widetilde{D}$ can "cut off" parts of the transformed function $A\mu$ for some $A \in \mathrm{Aff}(n)$, creating artificial symmetry (see Fig. 3). We avoid this issue by embedding the objects of interest within a sufficiently large, uniform background (Proposition 1). An alternative approach is to extend $\overline{\mu} : D \subset \mathbb{R}^n \to \mathbb{R}_{>0}$ to a function μ defined on all of $\mathbb{R}^n$. Assume that $\overline{\mu}(x) > \delta$ for some $\delta > 0$ for all $x \in D$. Then let:

$$\mu(x) = \begin{cases} \overline{\mu}(x) & x \in D \\ \delta & x \in \mathbb{R}^n \setminus D. \end{cases} \tag{21}$$

With this definition, $\widetilde{D} = \mathbb{R}^n$ for any transformation $A \in \mathrm{Aff}(n)$, and thus $\mathcal{TI}[\mu, A]$ is calculated by integrating over $\mathbb{R}^n$.

A.2 Additional Mathematical Properties of $\mathcal{TI}$

Proof of equation (15)

$$\begin{aligned} F_{\mathrm{rot}}(R_\varphi \mu) &= \int (\mu \circ R_\varphi^{-1}) \log \left(\frac{\mu \circ R_\varphi^{-1}}{\mu \circ R_\varphi^{-1} \circ R_\theta^{-1}} \right) \\ &= \int \mu \log \left(\frac{\mu}{\mu \circ R_\varphi^{-1} \circ R_\theta^{-1} \circ R_\varphi} \right) && (x \mapsto R_\varphi x) \\ &= \int \mu \log \left(\frac{\mu}{\mu \circ R_\theta^{-1}} \right) && (\text{rotations commute}) \\ &= \mathcal{TI}_{\mathrm{rot}}(\theta) \\ &= F_{rot}(\mu). \end{aligned}$$

□

Proof of equation (16)

$$\begin{aligned} F_{\mathrm{rot}}(\kappa_\varphi \mu) &= \int (\mu \circ \kappa_\varphi^{-1}) \log \left(\frac{\mu \circ \kappa_\varphi^{-1}}{\mu \circ \kappa_\varphi^{-1} \circ R_\theta^{-1}} \right) \\ &= \int \mu \log \left(\frac{\mu}{\mu \circ \kappa_\varphi^{-1} \circ R_\theta^{-1} \circ \kappa_\varphi} \right). && (x \mapsto \kappa_\varphi x) \end{aligned}$$

Now, we will show that $\kappa_\varphi^{-1} \circ R_\theta^{-1} \circ \kappa_\varphi = R_{-\theta}^{-1}$. Taking $z = (x, y) \in \mathbb{C}$, rotations R_θ and reflections κ_θ on $\mathbb{R}^2$ are equivalent to the following transformations on $\mathbb{C}$

$$R_\theta z = e^{i\theta} z \tag{22}$$
$$\kappa_\theta z = e^{2i\theta} \overline{z}. \tag{23}$$

Furthermore, it is easy to see that $\kappa_\theta^{-1} = \kappa_\theta$ and $R_\theta^{-1} = R_{-\theta}$. Therefore,

$$\begin{aligned} \kappa_\varphi^{-1} \circ R_\theta^{-1} \circ \kappa_\varphi z &= \kappa_\varphi^{-1} \circ R_\theta^{-1} e^{2i\varphi} \overline{z} \\ &= \kappa_\varphi^{-1} \circ e^{-i\theta} e^{2i\varphi} \overline{z} \\ &= e^{2i\varphi} e^{i\theta} e^{-2i\varphi} z \\ &= e^{i\theta} z \\ &= R_\theta z \\ &= R_{-\theta}^{-1} z. \end{aligned}$$

So,

$$\begin{aligned} F_{\text{rot}}(\kappa_\varphi \mu) &= \int \mu \log \left(\frac{\mu}{\mu \circ \kappa_\varphi^{-1} \circ R_\theta^{-1} \kappa_\theta} \right) \\ &= \int \mu \log \left(\frac{\mu}{\mu \circ R_{-\theta}^{-1}} \right) \\ &= \mathcal{TI}_{\text{rot}}(-\theta) \\ &= \overline{\kappa}_0 F_{\text{rot}}(\mu). \end{aligned}$$

□

Proof of equation (17)

$$\begin{aligned} F_{\text{ref}}(R_\varphi \mu) &= \int (\mu \circ R_\varphi^{-1}) \log \left(\frac{\mu \circ R_\varphi^{-1}}{\mu \circ R_\varphi^{-1} \circ \kappa_\theta^{-1}} \right) \\ &= \int \mu \log \left(\frac{\mu}{\mu \circ R_\varphi^{-1} \circ \kappa_\theta^{-1} \circ R_\varphi} \right). \qquad (x \mapsto R_\varphi x) \end{aligned}$$

Using complex coordinates,

$$\begin{aligned} R_\varphi^{-1} \circ \kappa_\theta^{-1} \circ R_\varphi z &= R_\varphi^{-1} \circ \kappa_\theta^{-1} e^{i\varphi} z \\ &= R_\varphi^{-1} e^{2i\theta} e^{-i\varphi} \overline{z} \\ &= e^{-i\varphi} e^{2i\theta} e^{-i\varphi} \overline{z} \\ &= e^{2i(\theta-\varphi)} \overline{z} \end{aligned}$$

$$= \kappa_{\theta-\varphi}^{-1} z.$$

Thus,

$$\begin{aligned} F_{\text{ref}}(R_\varphi \mu) &= \int \mu \log\left(\frac{\mu}{\mu \circ R_\varphi^{-1} \circ \kappa_\theta^{-1} \circ R_\varphi}\right) \\ &= \int \mu \log\left(\frac{\mu}{\mu \circ \kappa_{\theta-\varphi}^{-1}}\right) \\ &= \mathcal{TI}_{\text{ref}}(\theta - \varphi) \\ &= \mathcal{TI}_{\text{ref}}(\overline{R}_\varphi^{-1}(\theta)) \\ &= \overline{R}_\varphi F_{\text{ref}}(\mu). \end{aligned}$$

□

Proof of equation (18)

$$\begin{aligned} F_{\text{ref}}(\kappa_\varphi \mu) &= \int (\mu \circ \kappa_\varphi^{-1}) \log\left(\frac{\mu \circ \kappa_\varphi^{-1}}{\mu \circ \kappa_\varphi^{-1} \circ \kappa_\theta^{-1}}\right) \\ &= \int \mu \log\left(\frac{\mu}{\mu \circ \kappa_\varphi^{-1} \circ \kappa_\theta^{-1} \circ \kappa_\varphi}\right) \qquad (x \mapsto \kappa_\varphi x). \end{aligned}$$

Using complex coordinates,

$$\begin{aligned} \kappa_\varphi^{-1} \circ \kappa_\theta^{-1} \circ \kappa_\varphi z & \\ &= \kappa_\varphi^{-1} \kappa_\theta^{-1} (e^{2i\varphi} \overline{z}) \\ &= \kappa_\varphi^{-1} (e^{2i(\theta-\varphi)} z) \\ &= e^{2i(2\varphi-\theta)} \overline{z} \\ &= \kappa_{2\varphi-\theta} z, \end{aligned}$$

so

$$\begin{aligned} F_{\text{ref}}(\kappa_\varphi \mu) &= \mathcal{TI}_{\text{ref}}(2\varphi - \theta) \\ &= \mathcal{TI}_{\text{ref}}(\overline{\kappa}_\varphi^{-1}(\theta)) \\ &= \kappa_\varphi F_{\text{ref}}(\mu). \end{aligned}$$

□

Proposition 6 ($\mathcal{TI}$ is Lipshitz continuous) *Let let* $\{T_a\}_{a \in \mathcal{A}} \subset \text{Aff}(n)$ *be such that*

$$\sup_{a \in \mathcal{A}} \{|\det(T_a)|\} < \infty,$$

fix $D \subset \mathbb{R}^n$ such that $|D| < \infty$ (Lebesgue measure of D), and define

$$\mathcal{P}_{\delta,M} = \{f \in L^1(D) : 0 < \delta < f(x) < M \ a.e.\}$$
$$\mathcal{Y} = \{f : \mathcal{A} \to \mathbb{R}\}$$

where $\mathcal{P}_{\delta,M}$ has the metric induced by the L^1 norm, and $\mathcal{Y}$ has the sup *norm. Lastly, define $F : \mathcal{P}_{\delta,M} \to \mathcal{Y}$ as $F : \mu \mapsto \mathcal{TI}(a)$ where $\mathcal{TI}(a)$ is the $\mathcal{TI}$ curve of $\mathcal{TI}[\mu, T_a]$ as a function of $\mathcal{A}$. Then F is Lipshitz continuous, that is:*

$$|F(\mu) - F(\eta)|_{\sup(a)} \leq L\|\mu - \eta\|_{L^1(D)}$$

for some constant L.

Proof Take some arbitrary T_a, and let $\mu, \eta \in \mathcal{P}_{\delta,M}$. We can write

$$\mathcal{TI}[\mu, T_a] = \int_{\widetilde{D}} \Phi(\mu, T_a\mu)$$

where

$$\Phi(u, v) = u \log\left(\frac{u}{v}\right).$$

We will first show that the function $\Phi(u, v) : \mathbb{R}^2 \to \mathbb{R}$ is Lipshitz continuous on $S := (\delta, M) \times (\delta, M)$. Indeed, the partial derivatives

$$\Phi_u = \log\left(\frac{u}{v}\right) + 1 \tag{24}$$
$$\Phi_v = -\frac{u}{v} \tag{25}$$

are bounded on the convex set S. By the mean value theorem. For any $(u, v), (u', v') \in S$, there exists a point $c \in S$ such that

$$\Phi(u, v) - \Phi(u', v') = [(u, v) - (u', v')] \cdot \nabla\Phi(c) \tag{26}$$
$$= \Phi_u(c)(u - u') + \Phi_v(c)(v - v') \tag{27}$$

where $\cdot$ denotes the standard Euclidean inner product (or dot product), and $\nabla\Phi(c)$ denotes the gradient of Φ evaluated at c. Let

$$K_1 = \sup_{c \in S} |\Phi_u(c)|$$
$$K_2 = \sup_{c \in S} |\Phi_v(c)|,$$

and take $K = \max\{K_1, K_2\}$. Now, taking the absolute value on both sides of (26), (27), we have

$$|\Phi(u, v) - \Phi(u', v')| \leq K(|u - u'| + |v - v'|) \tag{28}$$

Therefore, Φ is Lipshitz when considering the domain $\mathbb{R}^2$ with the *Manhattan norm*.

We can then write

$$
\begin{aligned}
|\mathcal{TI}[\mu, T_a] - \mathcal{TI}[\eta, T_a]| &= \left| \int_{\widetilde{D}} \Phi(\mu, T_a\mu) - \int \Phi(\eta, T_a\eta) \right| \\
&= \left| \int_{\widetilde{D}} \Phi(\mu, T_a\mu) - \Phi(\eta, T_a\eta) \right| \\
&\leq \int_{\widetilde{D}} |\Phi(\mu, T_a\mu) - \Phi(\eta, T_a\eta)| \\
&\leq K \int_{\widetilde{D}} (|\mu - \eta| + |T_a\mu - T_a\eta|) \qquad \text{equation (28)} \\
&= K \left(\int_{\widetilde{D}} |\mu - \eta| + \int_{\widetilde{D}} |\mu \circ T_a^{-1} - \eta \circ T_a^{-1}| \right) \\
&\leq K(1 + |\det(T_a)|) \int_{\widetilde{D}} |\mu - \eta| \\
&\leq K(1 + |\det(T_a)|) \|\mu - \eta\|_{L^1(D)}
\end{aligned}
$$

Let $L := K(1 + \sup_{a \in \mathcal{A}}\{|\det(T_a)|\})$, where the supremum is finite by hypothesis. Then for every a,

$$|\mathcal{TI}[\mu, T_a] - \mathcal{TI}[\eta, T_a]| \leq L\|\mu - \eta\|_{L^1(D)},$$

or equivalently

$$|F(\mu) - F(\eta)|_{\sup(a)} \leq L\|\mu - \eta\|_{L^1(D)}.$$

□

Proposition 7 (Conditions for equivalent reflections and rotations) *Let $\mu : \mathbb{R}^2 \to \mathbb{R}$. If there exists some $\varphi \in [0, 2\pi)$ such that $\kappa_\varphi \mu = \mu$, then for all $\theta \in [0, 2\pi)$,*

$$\kappa_\theta \mu = R_{2(\theta-\varphi)}\mu.$$

Proof As in (22), (23), we will use the fact that $\mathbb{R} \cong \mathbb{C}$ (i.e. they are isomorphic as vector spaces). Assume that there exists some $\varphi \in [0, 2\pi)$ such that $\kappa_\varphi \mu \equiv \mu$. Then for any $z \in \mathbb{C}$

$$
\begin{aligned}
\kappa_\theta \mu(z) &= \mu(e^{2i\theta}\overline{z}) \\
&= \kappa_\varphi \mu(e^{2i\theta}\overline{z}) \qquad \text{(by assumption)} \\
&= \mu(e^{2i\varphi} e^{-2i\theta} z) \\
&= \mu(e^{2i(\varphi-\theta)} z) \\
&= \mu(R_{2(\varphi-\theta)} z) \\
&= R_{2(\theta-\varphi)}\mu(z).
\end{aligned}
$$

□

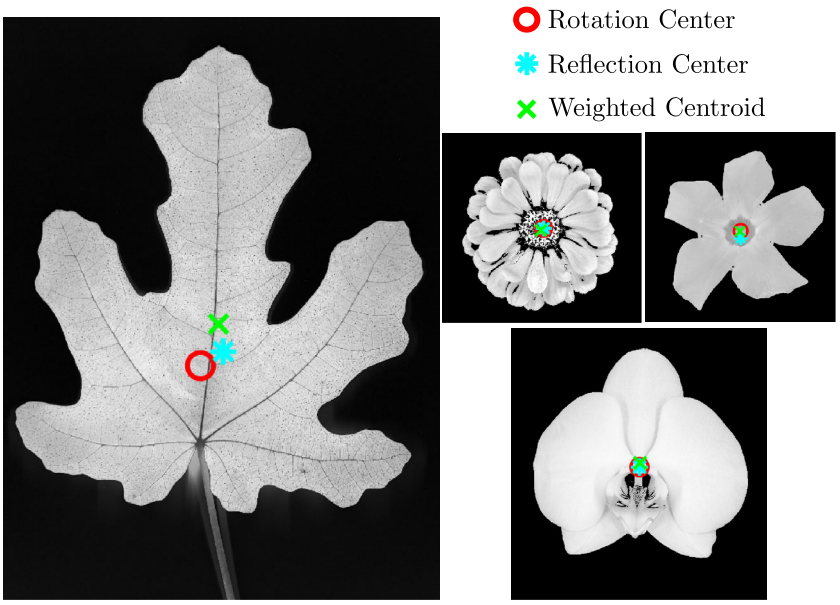


Fig. 10 In four sample images, the symmetry center by rotation $\overline{x}^R$, reflection $\overline{x}^\kappa$, and the intensity-weighted centroid agree closely

A.3 Comparing Approaches for Center Detection

Definition 4 locates the symmetry center by maximizing the observed symmetry over candidate points, which is considerably more computationally expensive than computing the intensity-weighted centroid (4). Here we compare the rotational and reflectional symmetry centers to the intensity-weighted centroid, and to each other, across images in the floral dataset. All distances are reported as a percentage of the image's diagonal length. When computing the symmetry centers, we use the definition of $\mathcal{TI}$ given by (21) to avoid artificial symmetries and observe improvements as shown in Fig. 11. Rather than integrating over $\widetilde{D}$, the intersection of the domains of the original and transformed image, we instead integrate over the union of the images' domains, trivially extending each image into the domain where it is not defined by setting the pixel intensity to a preselected small value δ. For more details on symmetry center detection, see Sect. 4.4 (Fig. 10).

Figure 12 shows the distribution of distances across the floral dataset for all three centers. Table 1 provides summary statistics for each distribution. The rotational symmetry center $\overline{x}^R$ agrees most closely with the centroid, with a median displacement of 1.43% and 93.6% of images below 5% displacement. The reflectional center $\overline{x}^\kappa$ is displaced somewhat further from the centroid (median 2.75%) but with a tighter distribution as the distance is less than 10% in all images. The two symmetry centers differ from one another more than either differs from the centroid (median 3.57%).

Despite close agreement of the centers across most of the dataset, in 3.5% of images (6 of 171) the distance between $\overline{x}^R$ and the centroid is over 10%, and in 4.7% of images

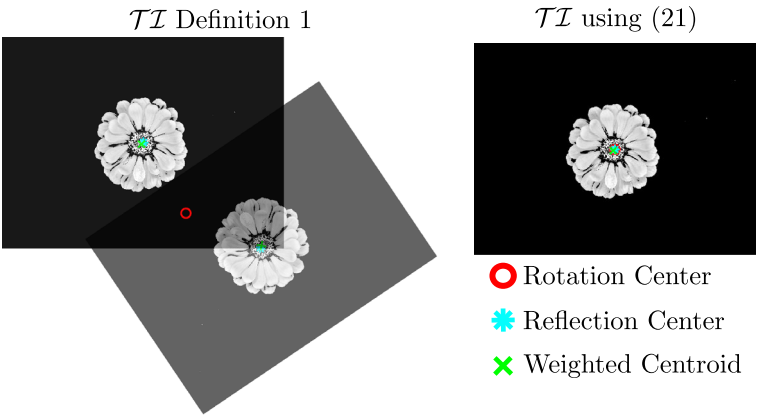


Fig. 11 As discussed in Sect. A.1, using Equation (21) helps avoid artificial symmetries that may arise with Definition 1. For some images, this provides an improvement in the center detection. Left: the rotational center detected with Definition 1 is far from the flower. Rotating the image by $214\pi/180$ about that center places the rotated flower entirely on the background of the original image, yielding an artificially small $\mathcal{TI}$ value. Right: using Equation (21) for $\mathcal{TI}$, the rotational center falls in the middle of the flower

(8 of 171) the distance between symmetry centers is more that 10% apart. These cases show that there are cases where the choice of center selection method, and hence the choice among the four $\mathcal{TI}$ curves of Sect. 4, may be significant. Further investigation is required to understand why the centers differ in some images but not in others.

A.4 Details of Preprocessing

If a grayscale image is represented by the matrix $M = (m_{ij})$, its negative is simply given by $N = (255 - m_{ij})$. As before, we define the representative function of the image as

$$\overline{\mu}(i, j) = (255 - m_{ij}) + \delta\,,$$

where $\delta > 0$ is a small value added to avoid zero pixel values, thus ensuring $\mathcal{TI}$ is defined (see Fig. S4).

In cases where the background of an image has significant variations in light intensity and color, the image requires additional preprocessing so that the $\mathcal{TI}$ measure does not detect properties of the background. If every point in the background has lower light intensity than the points in the object, then we can use thresholding to remove the background noise. Mathematically, consider a grayscale image represented by the intensity matrix $M = (m_{ij})$, and suppose the object of interest is represented with a set of pixels $F \subset D$. If for every $(i, j) \in F$ and $(\ell, k) \in D \setminus F$, we have

$$m_{\ell k} < m_{ij}, \tag{29}$$

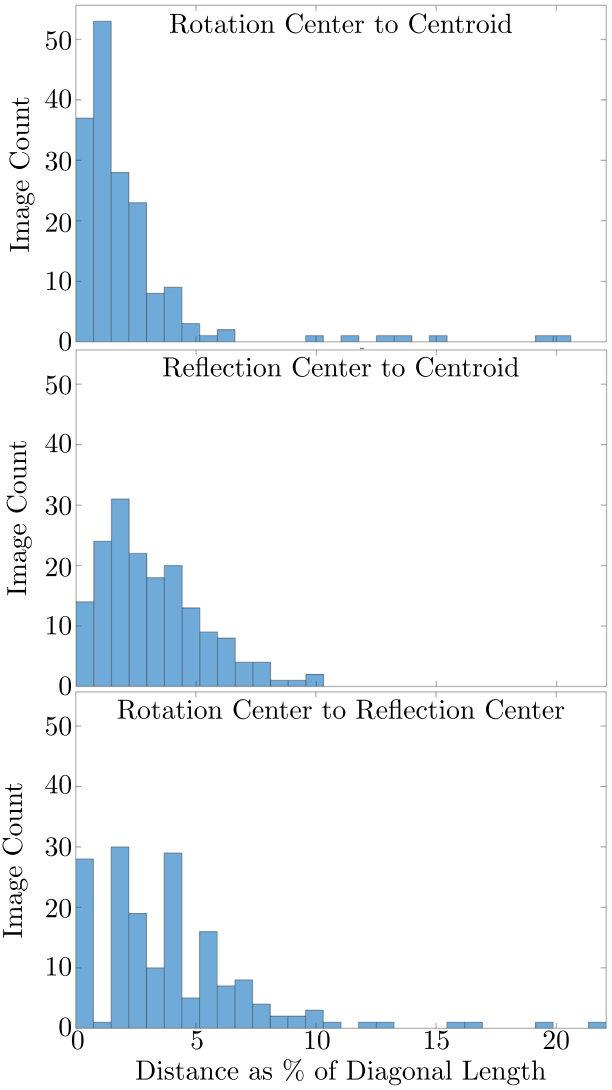


Fig. 12 Across all flowers in the dataset from Byng et al. (2018), we compute the pairwise distances between the symmetry center by rotation, reflection, and the intensity-weighted centroid. The histograms shows the distance between each center as a percentage of the image's diagonal length

then we can choose a threshold $T := \min_{(i,j)\in F}\{m_{ij}\}$ and set

$$\overline{\mu}(i,j) = \begin{cases} 1 & m_{ij} < T \\ m_{ij} & m_{ij} \geq T. \end{cases} \tag{30}$$

Then $\overline{\mu}$ retains the grayscale image's pixel intensity on the object F and achieves the minimal pixel intensity on the complement $D \setminus F$. Alternatively, if we wish to

Table 1 Summary statistics of the pairwise distances between the rotational symmetry center $\overline{x}^R$, the reflectional symmetry center $\overline{x}^\kappa$, and the intensity-weighted centroid, across 171 floral images. Distances are expressed as a percentage of image diagonal length

Pair	Mean	Median	p_{90}	$< 5\%$	$< 10\%$
$\overline{x}^R$ to centroid	2.24	1.43	3.98	93.6%	96.5%
$\overline{x}^\kappa$ to centroid	3.18	2.75	6.09	81.9%	100%
$\overline{x}^R$ to $\overline{x}^\kappa$	3.91	3.57	7.41	69.6%	95.3%

investigate symmetry in the object's global boundary geometry rather than in textures or internal surface features, we can consider the binarized image

$$\overline{\mu}(i,j) = \begin{cases} 1 & m_{ij} < T \\ 255 & m_{ij} \geq T. \end{cases} \tag{31}$$

The difference between thresholding (30) and binarization (31) can be observed in Fig. S2.

In general, the threshold T can be difficult to find and varies between images within a dataset. Fortunately, there are well-established methods for automated preprocessing. In our computational pipeline (GitHub 2025), we use Otsu's method (available via the MATLAB function `otsuthresh`), which finds a threshold T that maximizes the variance in pixel intensity between the groups created by the threshold.

Sometimes, thresholding is insufficient to distinguish between the background and the object since (29) is not satisfied. To resolve this issue, we propose taking the largest connected component of the binarized image or filtering an image before thresholding. To take the largest connected component, one can use the MATLAB function `bwareafilt`. Alternatively, filters smooth an image by averaging pixel values over a neighborhood. For example, we can apply a Gaussian filter with the MATLAB function `imgaussfilt`, which makes subsequent thresholding more effective. In fig. S4, when the standard deviation of the Gaussian filter is $\sigma = 0$ (i.e. no change to the image), there are light pixels near the bottom of the leaf, which could bias the $\mathcal{TI}$ curve. Increasing the standard deviation "smooths out" the image and removes background noise.

A.5 Reflectional $\mathcal{TI}$ for Floral Dataset

The analysis in Sect. 5.1 considers $\mathcal{TI}$ by rotation about the rotational symmetry center. Here we present the corresponding analysis for the $\mathcal{TI}$ curve by reflection about the reflectional symmetry center ($\mathcal{TI}[\overline{\mu}_{\overline{x}^\kappa}, \kappa_\theta]$, Sect. 4.4), with fPCA plots for the $\mathcal{TI}$ curves shown in Fig. 13. Note that the reflectional symmetry centers were computed taking the original definition of $\mathcal{TI}$ rather than integrating over the union the original and transformed images' domains as described by (21) and in Sect. A.3.

The two analyses are not independent. By Proposition 7, an object with exact reflection symmetry has rotational and reflectional $\mathcal{TI}$ curves that determine each

other, so in a dataset of exactly mirror-symmetric flowers, the $\mathcal{TI}$ curves for rotation and reflection would contain the same information. Real flowers deviate from exact mirror symmetry, and the difference can be seen when comparing Fig. 6 and Fig. 13. We hypothesize that part of the difference is also because rotational $\mathcal{TI}$ is orientation-agnostic while reflectional $\mathcal{TI}$ is not (Proposition 5).

When analyzing the fPCA plots for the reflectional $\mathcal{TI}$, the qualitative conclusions of Sect. 5 are unchanged. First, symmetry appears as a continuum rather than being grouped into distinct categories. Second, the $\mathcal{TI}$ curves in the fPCA plot of all flowers (Fig. 13, bottom) overlap across distinct clades, suggesting convergent evolution of floral symmetry. However, the structure that was visually interpretable in the rotational analysis is less apparent here. In particular, the rosid fPCA plot does not separate into branches corresponding to the 4-fold and 5-fold rotational symmetry, and the full fPCA plot does not show separation into bilateral and radial symmetry. Thus, $\mathcal{TI}$ by reflection is capturing information that may not be accessible via conventional approaches, like visual inspection of the first two components in PCA.

A.6 Details of Fern Simulations

To obtain simulated images of ferns, we numerically compute the attractor of the following iterated function system (IFS) (Barnsley 1993), where $p \in \mathbb{R}$ is a parameter.

$$
\begin{aligned}
f_1(x, y) &= \begin{pmatrix} 0 & 0 \\ 0 & 0.16 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ -0.4p \end{pmatrix} \\
f_2(x, y) &= \begin{pmatrix} 0.85(1-p) + 0.95p & 0 \\ -0.04(1-p) - 0.005p & 0.85(1-p) + 0.93p \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -0.002p \\ 1.6(1-p) + 0.5p \end{pmatrix} \\
f_3(x, y) &= \begin{pmatrix} 0.2(1-p) + 0.035p & -0.26(1-p) - 0.2p \\ 0.23(1-p) + 0.16p & 0.22(1-p) + 0.04p \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -0.09p \\ 1.6(1-p) + 0.02p \end{pmatrix} \\
f_4(x, y) &= \begin{pmatrix} -0.15(1-p) - 0.04p & 0.28(1-p) + 0.2p \\ 0.26(1-p) + 0.16p & 0.24(1-p) + 0.04p \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0.083p \\ 0.44(1-p) + 0.12p \end{pmatrix}
\end{aligned}
$$

When $p = 0$, the attractor is a Barnsley-type fern as shown in Fig. 9 (top right) and when $p = 1$, the attractor is a *Cyclosorus*-type fern as in Fig. 9 (bottom left).

In our analysis, we include 51 images, which are given by numerically computing the attractor of the IFS with $p = 0,\ 0.02,\ 0.04,\ \ldots,\ 1.00$.

To simulate an attractor, we fix p and consider a random walk given by the IFS. Specifically, we start at $(x, y) = (0, 0)$. We select one of the transformations f_1, f_2, f_3, f_4 based on the probability distribution $\boldsymbol{\pi}$ given by

$$
\begin{aligned}
P(f_1 \text{ is selected}) &= 0.01(1-p) + 0.02p \\
P(f_2 \text{ is selected}) &= 0.85(1-p) + 0.84p \\
P(f_3 \text{ is selected}) &= 0.07 \\
P(f_4 \text{ is selected}) &= 0.07.
\end{aligned}
$$

Once f_i is selected, we update (x, y) by applying f_i and repeat the process. This is made explicit with the following pseudocode.

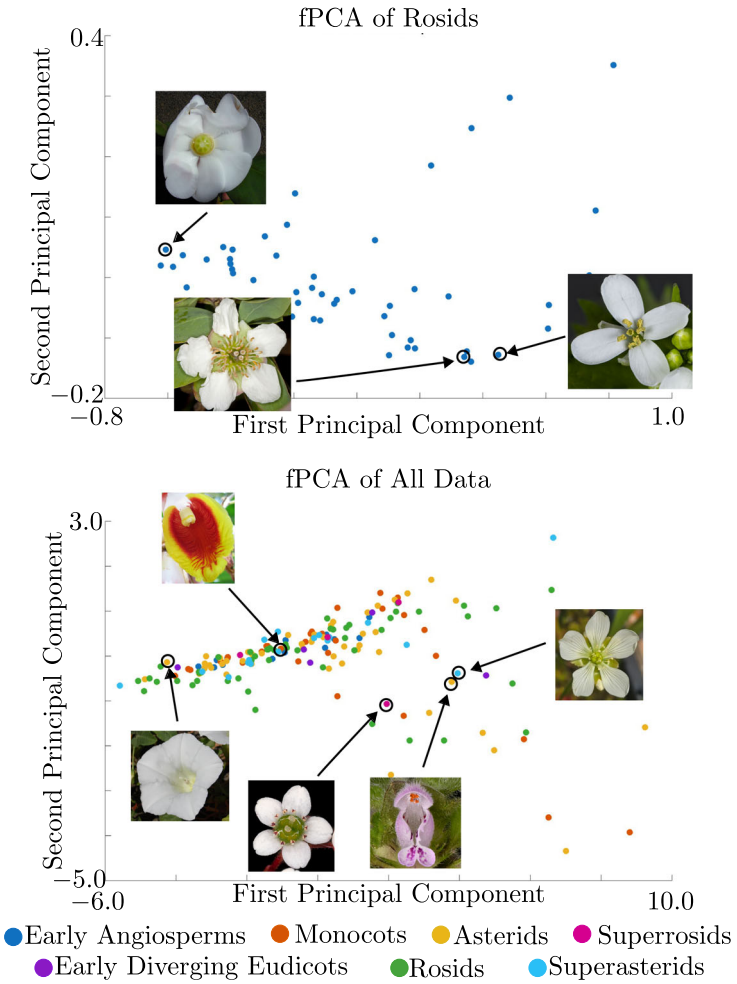


Fig. 13 fPCA of $\mathcal{TI}$ curves by reflection about the reflectional symmetry center ($\mathcal{TI}[\mu_{\overline{x}^\kappa}, \kappa_\theta]$) reveals a continuous landscape of symmetries, similar to the rotational $\mathcal{TI}$ about the rotational symmetry center (Fig. 6). Top: rosid $\mathcal{TI}$ curves projected onto the first two principal component symmetries, capturing 85.517% of the variance. Bottom: all projected $\mathcal{TI}$ curves, color-coded by clade, capturing 85.538% of the variance. The flower images presented on the graphs are stock images that closely match the image data used for our analysis (Descouens 2022; Game 2017; Starr and Starr 2026; North Queensland Plants 2026; Reese 2024; falcon_311 2025; Carr 2005; Edison and Ford Winter Estates 2010; Smith 2024). Clades remain intermixed as in Fig. 6; however, visual separation into 4-fold and 5-fold symmetry is not apparent in the rosids, and separation into bilateral and rotational symmetry is not apparent in the full dataset

We consider a scatter plots of the points $\{(x_i, y_i)\}_{i=1}^N$ and use the MATLAB function `exportgraphics` to convert the scatter plot to an image file.

Algorithm 1: Iterated Function System (IFS) Attractor

Input: Transformations f_1, f_2, f_3, f_4, probability vector $\boldsymbol{\pi}$, number of iterations N
Output: Point sequence $\{(x_i, y_i)\}_{i=1}^{N}$
Initialize $(x, y) \leftarrow (0, 0)$;
for $i = 1, 2, \ldots, N$ **do**
  Draw $k \sim \boldsymbol{\pi}$, so that $k \in \{1, 2, 3, 4\}$;
  $(x, y) \leftarrow f_k(x, y)$;
  Store $(x_i, y_i) \leftarrow (x, y)$;
return $\{(x_i, y_i)\}_{i=1}^{N}$

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11538-026-01754-9.

Data availability The code referenced in this paper is available at GitHub (2025). It contains the computational pipeline used to create all $\mathcal{TI}$ curves and fPCA plots presented. All figures were created using MATLAB and Inkscape.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Barnsley MF (1993) Fractals Everywhere, 2nd edn., pp. 86–87. Academic Press, Boston, MA. IFS code and random iteration algorithm for the Barnsley fern

Bermingham N, Robins V, Turner K (2023) Planar symmetry detection and quantification using the extended persistent homology transform. 2023 Topological Data Anal Visualization (TopoInVis). IEEE, pp 1–9

Byng JW, Smets EF, Vugt R, Bidault E, Davidson C, Kenicer G, Chase MW, Christenhusz MJ (2018) The phylogeny of angiosperms poster: a visual summary of apg iv family relationships and floral diversity. The Global Flora 4(7):1–4

Carr GD (2005) Photograph of Saxifragaceae plant. University of Hawaii Botany Department Vascular Plant Families. http://www.botany.hawaii.edu/faculty/carr/saxifrag.htm

Ciocanel M-V, Gandhi P, Niklas K, Dawes AT (2025) Characterizing symmetry transitions in systems with dynamic morphology. Math Biosci 384:109431

Descouens D (2022) Photograph of Plant Specimen. Wikimedia Commons. Accessed: 2026-05-03. https://commons.wikimedia.org/w/index.php?curid=116433055

Edison and Ford Winter Estates (2010) Shell Ginger closed blooms. https://www.edisonfordwinterestates.org/name-that-plant-vi-answer-shell-ginger/

Endress PK (1999) Symmetry in flowers: diversity and evolution. Int J Plant Sci 160(S6):3–23

falcon_311 (2025) Droseraceae flowers. Reddit (r/SavageGarden). Accessed: 2026-05-03. https://www.reddit.com/r/SavageGarden/comments/1l5vor3/droseraceae_flowers/

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD (2004) Pollination syndromes and floral specialization. Annu Rev Ecol Evol Syst 35(1):375–403

Game J (2017) Plant Photograph. Wikimedia Commons. Accessed: 2026-05-03. https://commons.wikimedia.org/w/index.php?curid=57693388

Gandhi P, Ciocanel MV, Niklas K, Dawes AT (2021) Identification of approximate symmetries in biological development. Philosophical Transactions Royal Soc A Math Phys Eng Sci 379: https://doi.org/10.1098/rsta.2020.0273

Gandhi P, Iams S, Bonetti S, Silber M (2019) In: D'Odorico P, Porporato A, Wilkinson Runyan C (eds.) Vegetation Pattern Formation in Drylands, pp. 469–509. Springer, Cham. https://doi.org/10.1007/978-3-030-23269-6_18

GitHub (2025) Accompanying MATLAB software for measuring and comparing the symmetries of image data. GitHub. Available at: https://github.com/ldobrien1234/TransInfo

Golubitsky M, Stewart I, Schaeffer DG (2012) Singularities and Groups in Bifurcation Theory: Volume II, vol 2. Springer, New York, NY

Klingenberg CP (2002) Morphometrics and the role of the phenotype in studies of the evolution of developmental mechanisms. Gene 287(1–2):3–10

Kravtsova N, McGee RL II, Dawes AT (2023) Scalable gromov-wasserstein based comparison of biological time series. Bull Math Biol 85(8):77

Kästner K, van de Vijsel RC, Caviedes-Voullième D, Frechen NT, Hinz C (2024) Unravelling the spatial structure of regular dryland vegetation patterns. CATENA 247:108442. https://doi.org/10.1016/j.catena.2024.108442

Luo D, Carpenter R, Vincent C, Copsey L, Coen E (1996) Origin of floral asymmetry in antirrhinum. Nature 383(6603):794–799

Neal PR, Dafni A, Giurfa M (1998) Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. Annu Rev Ecol Syst 29(1):345–373

Nikkeshi A, Kurimoto D, Ushimaru A (2015) Low flower-size variation in bilaterally symmetrical flowers: support for the pollination precision hypothesis. Am J Bot 102(12):2032–2040

North Queensland Plants (2026) Australian Carnivorous Plants - Droseraceae: Drosera nana. Accessed: 2026-01-09 http://www.northqueenslandplants.com/Australian%20Plant%20Families%20A-F/Droseraceae/Drosera/Drosera%20nana.html

Rostami MW, Bannish BE, Gasior K, Pinals RL, Copos C, Dawes AT (2023) Inferring local molecular dynamics from the global actin network structure: a case study of 2d synthetic branching actin networks. J Theor Biol 575:111613

Reese RN (2024) Convolvulaceae: Convolvulus arvensis (Native Plant Research). South Dakota State University Open PRAIRIE. Accessed: 2026-05-03. https://openprairie.sdstate.edu/nativeplant/230/

Sargent RD (2004) Floral symmetry affects speciation rates in angiosperms. In: Proceedings of the Royal Society of London. Series B: Biological Sciences 271(1539), 603–608

Savriama Y, Klingenberg CP (2011) Beyond bilateral symmetry: geometric morphometric methods for any type of symmetry. BMC Evol Biol 11(1):280

Smith DG (2024) Lamium purpureum (purple deadnettle). Delaware Wildflowers; [accessed 2026 Sep 16]. https://delawarewildflowers.org/plant.php?id=1113

Starr F, Starr K (2026) Close-up of a flower (Clusia rosea). Useful Tropical Plants. Accessed: 2026-05-03 https://tropical.theferns.info/image.php?id=Clusia+rosea

Shi H, Yang Y, Wang L, Ma D, Beg MF, Pei J, Cao J (2022) Two-dimensional functional principal component analysis for image feature extraction. J Comput Graph Stat 31(4):1127–1140

Vstovsky G (1997) Transform information: a symmetry breaking measure. Found Phys 27(10):1413–1444

Wang Y, Luo A, Lyu T, Dimitrov D, Liu Y, Li Y, Xu X, Freckleton RP, Hao Z, Wang Z (2023) Global distribution and evolutionary transitions of floral symmetry in angiosperms. Sci Adv 9(43):2555

Yao F, Müller H-G, Clifford AJ, Dueker SR, Follett J, Lin Y, Buchholz B, Vogel JS (2003) Shrinkage estimation for functional principal component scores, with application to the population kinetics of plasma folate. Biometrics 59(3):676–685

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