Applications of the Information Dimension in Detecting Border Perturbations

Soraia Bitencourt Carvalho[1*], Antônio Teófilo A. Nascimento[2]
[1]UNEB, PPGMSB; [2]UNEB, DCET; Salvador, Bahia; Alagoinhas, Bahia, Brazil

This study utilizes the concept of information dimension based on Shannon and Tsallis' entropy to analyze the contours of flat objects. Our objective is to employ the information dimension for detecting perturbations in borders. We create examples of squares with slight border perturbations, possessing the same mass and sharing identical box-counting dimensions yet exhibiting distinct information dimensions. This construction was devised with the understanding that entropy is responsive to the image frequency within each box. Consequently, the information dimension provides a more precise index of fractal shapes when compared to the box-counting dimension.
Keywords: Information Dimension. Box-Counting Dimension. Entropy.

Introduction

Numerous natural phenomena exhibit irregularities and complexity across various observation scales. The use of fractals to quantify the complexity of such phenomena or natural objects is increasingly prevalent, primarily due to the widespread identification of self-similarities in nature [1]. A fractal can be conceptualized as a complex geometric shape composed of smaller replicas of itself.

Fractals, being mathematical constructs, cannot be entirely described using traditional Euclidean geometry. However, they can be associated with a numerical value known as the fractal dimension, which provides insights into how the fractal's shape (topological nature) occupies its habitat (Euclidean space). Moreover, the fractal dimension quantifies how the "size" of a fractal set varies with different observation scales.

The fractal dimension characterizes a set as a whole or its boundary. In the former case, it indicates the density with which the set occupies its spatial domain, while in the latter, it denotes the irregularity of its perimeter. In both cases, determining the fractal dimension entails measuring the complexity of objects.

Several adaptations of the fractal dimension have been proposed, often based on the Hausdorff dimension for its mathematical rigor [2]. However, these adaptations, such as the area-perimeter relationship and the box-counting dimension, are predicated on measuring specific characteristics of the dataset and relating them to a length scale through a power law. The fractal dimension thus emerges as a function of the power law exponent, representing the slope of a straight line in log-log space following linear regression.

In this study, we aim to investigate the existence of an adapted metric capable of resolving ambiguities inherent in fractal dimensions, particularly those related to the box-counting dimension. We demonstrate that the information dimension, derived from Shannon entropy and for the first time enhanced with Tsallis entropy, can effectively detect subtle border perturbations that may elude detection by traditional box-counting methods.

The significance of this research lies in its experimental validation of the application of the information dimension to discern border perturbations from flat objects. Notably, this technique holds promise for ecological applications, including the detection of pathologies in cells and leaves and forest monitoring efforts.

To generate examples illustrating ambiguities concerning the box-counting dimension, we construct squares with consistent mass along their

Received on 16 November 2023; revised 20 December 2023.
Address for correspondence: Soraia Bitencourt Carvalho. BR 110, Km 03, Alagoinhas. Zipcode: 48.000.000. Alagoinhas, Bahia, Brazil. E-mail: soraiabitencourt75@gmail.com.

J Bioeng. Tech. Health 2023;6(Suppl 2):44-50

borders (pixels) and introduce perturbations while maintaining this mass. These perturbations result in varying pixel counts within each box, as determined by Shannon and Tsallis entropy, yielding different information dimensions for each perturbation.

The structure of this paper commences with a discussion of foundational concepts underpinning the method, including fractals, fractal dimension, box-counting dimension, and entropy. Subsequently, we characterize the dimensions and entropy variations specific to the information dimension and then describe the model and methodologies employed. We then present the results obtained and conclude with reflections on this research endeavor's limitations and future prospects.

Theoretical Framework

Fractals

The term 'fractal' was introduced by Mandelbrot in his seminal essay, derived from the Latin word 'fractus,' meaning broken, to describe objects that exhibit irregularities beyond traditional geometric configurations. Mandelbrot [3] defined fractals as conceptual entities possessing a similar structure across all spatial scales, characterized by self-similarity and scale-independence.

The fractal dimension, a fundamental concept in characterizing complex mathematical shapes known as fractals, quantifies the extent to which these shapes occupy space. Developed initially to quantify specific attributes of self-similar objects like fractal shapes, these measures find application in ecology, particularly in studying forest landscapes. This mathematical field offers captivating properties that can be appreciated for their beauty, intricacies, characteristics, and analogies.

The exploration of methods employing fractals and their geometric structures is pertinent, given the rising popularity of fractal geometry in recent years. It has emerged as an art form and a computationally driven tool for modeling various physical phenomena. Consequently, the significance of this knowledge becomes evident.

The box-counting dimension is one of the most widely used and computationally straightforward methods for computing the fractal dimension. However, as we will elucidate here, this method presents ambiguities stemming from its limitation to merely counting the number of boxes, or hypercubes, covering the analyzed object without considering the distribution of points within each covering box. This limitation underscores the potential value of extending the utility of the box-counting dimension through the information dimension.

Box-Counting Dimension

In the context of flat figures, this method consists of dividing the image into boxes, where the number of boxes containing some part of the figure, which represents the object under study, such as the fractal, are counted, and the dimension value can be calculated by following the Equation 1:

$$D_{Box} = \lim_{\varepsilon \to 0} \frac{\ln n(\varepsilon)}{ln\left(\frac{1}{\varepsilon}\right)} \tag{1}$$

Let ε represent the side length of each box, and $n(\varepsilon)$ denote the number of boxes containing any portion of the figure. As the scale decreases, the established dimension becomes more precise. The fundamental concept is to measure the figure while disregarding irregularities more minor than the scale ε by analyzing the behavior of the measurement as the scale approaches zero. Notably, the boxes are cubes for three-dimensional objects, whereas, in n-dimensional spaces, they are replaced by hypercubes.

Shannon Information Dimension

The concept of information dimension has yet to be widely adopted. Nevertheless, it offers a perspective on the complexity of information by considering the form and the distribution of information within the object under study. This dimension is defined by Shannon entropy, as outlined by Seuront [4].

Additionally, we propose the utilization of Tsallis entropy to emphasize boxes containing more information about the object.

The information dimension, employing entropy, enables understanding the diversity and intricacy of information within a dataset. It is a crucial metric for assessing the quantity of potentially valuable and significant information within a system. Entropy, in this context, measures the uncertainty or disorder inherent in a system and can be regarded as a measure of the amount and distribution of information within a dataset.

In our study, the dataset for entropy comprises the distribution of pixels within each box. In contrast, the information dimension with entropy pertains to the diversity and distinctiveness of information present in the dataset. Higher entropy signifies a greater variety and diversity of information within the dataset.

Historically, entropy emerged in information theory to quantify the average information required to encode a message within a given system. It is computed based on the probability of occurrence of different events or symbols within a dataset.Entropy highlights scenarios where a dataset contains only a few symbols, resulting in low entropy due to limited information diversity.Conversely, entropy would be high in a dataset with a uniform distribution of various symbols, indicating a broader range of information.

The fundamental concept of entropy was developed to measure the expected rarity or surprise of a random variable X within its distribution. In literature, entropy is commonly regarded as a measure of information, quantifying the average amount needed to describe a dataset based on the probability distribution of events.

The formula for Shannon entropy is expressed in the Equation 2 as following:

$$H(X) = -\sum p(x)\log p(x) \tag{2}$$

Where X represents the random variable associated with a particular experiment, and p(x) denotes the probability of event x occurring across all possible events.

To adapt the box-counting dimension, the number of size boxes ε is replaced by the entropy of the size boxes ε, where the probability distribution is determined by the frequency of the figure within each box ε. Specifically, the information dimension using Shannon entropy is defined by Equations (3) and (4):

$$D_S = \lim_{\varepsilon \to 0} \frac{\ln H(\varepsilon)}{ln\left(\frac{1}{\varepsilon}\right)} \tag{3}$$

H(ε) in Equation 4) is the Shannon entropy

$$H(\varepsilon) = -\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{p_i}(\varepsilon)\ln\left(\mathrm{p_i}(\varepsilon)\right) \tag{4}$$

Here, *pi(ε)* denotes the relative frequency of the object within the size box ε. In other words, given *N* as the total number of pixels in the image and *fi(ε)* as the number of pixels within size box ε, we have the Equation 5 as following:

$$p_i(\varepsilon) = \frac{f_i(\varepsilon)}{N} \tag{5}$$

Varyinh i, the number of size boxes ε changes, ensuring the following:

$$\sum_{i=1}^{n} p_i(\varepsilon) = 1$$

Tsallis Information Dimension

The contribution of statistical mechanics, pioneered by Tsallis, was to propose a potential generalization of the renowned entropies of Boltzmann, Gibbs, and Shannon, providing a framework for describing physical systems [1].

Given that the Shannon information dimension accounts for the quantity of information within each box, it is natural to extend this concept by employing Tsallis entropy, defined by the Equation 6 as following:

$$H_{TS}(\varepsilon) = \frac{1}{q-1}\left(1 - \sum_{i=1}^{n} p_i(\varepsilon)^q\right) \tag{6}$$

It is worth noting that in the limit as $q \to 1$, the Tsallis entropy converges to the value of Shannon entropy. Consequently, we define the Tsallis information dimension as the Equation 7 below:

$$D_{TS} = \lim_{\varepsilon \to 0} \frac{\ln H_{TS}(\varepsilon)}{ln\left(\frac{1}{\varepsilon}\right)} \quad (7)$$

The parameter q modulates the information within each size bin ε, which can be harnessed to discern border perturbations more effectively.

Materials and Methods

We constructed examples of geometric shapes with identical box-counting dimensions yet differing information dimensions. This demonstration underscores that the information dimension effectively resolves ambiguities arising from the box-counting dimension, providing deeper insights into the complexity of particular flat objects. Such insights hold significant potential for applications in various fields, including detecting pathologies in ecology or material wear in engineering.

Our geometric models were crafted using GIMP software. The algorithms, graphs, and tables were developed using Python programming within the VSCode-integrated development environment. Specifically, we employed GIMP software to create squares (Figure 1), ensuring that their contours contained an equal number of pixels, thereby maintaining the same mass.

Notably, the vertices and dimensions of these squares were powers of two. This meticulous approach aimed to ensure that subdivisions of the boxes in the box-counting algorithm preserved the intersections of the previous divisions.

Additionally, we developed a software tool named FracDim in the Python programming language. This software facilitates the calculation of linear regressions for the box-counting dimensions and the Shannon and Tsallis information dimensions.

We implemented algorithms for computing Shannon and Tsallis entropy. Initially, the software was tested using regular geometric shapes. With each iteration, subtle modifications were introduced to the borders of the shapes to generate various flat figures while maintaining the same box-counting dimension. This iterative process led to the creation of modified squares with alterations on one, two, three, and four sides, as depicted in the figures. Remarkably, these squares share identical box-counting dimensions but exhibit different information dimensions.

Results and Discussion

Examples of geometric figures with minor border perturbations were generated to illustrate how the information dimension can resolve ambiguities inherent in the box-counting dimension. Both Shannon's and Tsallis' information dimensions were applied.

Figura 1. Square A with 1 perturbed side; square B with 2 perturbed sides; square C with 3 perturbed sides; square D with all sides perturbed.

Table 1 illustrates the ambiguity of the box-counting dimension, denoted as D_{Box}, where squares A, B, C, and D exhibit identical values. In the third column, the information dimension D_S derived from Shannon's entropy is presented.

In Figures 2-5, corresponding to the perturbed squares A, B, C, and D, it is noteworthy that the Shannon information dimension consistently registers values lower than the box-counting dimension. On the *y*-axis, these graphs depict the natural logarithm of the number of boxes intersecting the figure and the entropy of each box. In contrast, on the *x*-axis, we depict the natural logarithm of $1/\varepsilon$, where ε represents the length of the side of each box.

In Table 2, the same analysis was carried out by applying Tsallis entropy to the parameter q=0.5.

We observed that the entropy values were influenced by the parameter q, which can be advantageous in specific practical scenarios. This effect is illustrated in Figure 6, depicting the variation in Tsallis entropy within square A (Figure 1A).

Conclusion

This study delved into utilizing the information dimension, predicated on Shannon and Tsallis entropy, for detecting perturbations in the borders of objects representable by flat figures. Through targeted constructions, we showcased the efficacy of the information dimension in resolving ambiguities stemming from the box-counting dimension, thereby enhancing precision in analyzing border complexities in plane-modeled objects. The findings underscored that squares with subtle contour perturbations exhibited identical box-counting dimensions but divergent information

Table 1. Result of the box-counting dimension and Shannon.

Squar	D_{Box}	D_S
A	1.2796710276704335	0.33592615925964914
B	1.2796710276704335	0.33781554676065440
C	1.2796710276704335	0.33965496788217200
D	1.2796710276704335	0.34131330650656094

Figure 2. Information dimension and box-counting for square A (Figure 1A).

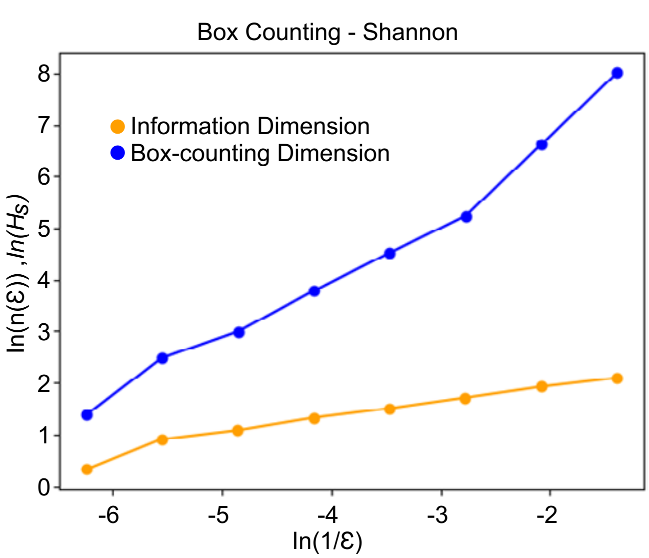


Figure 3. Information dimension and box-counting for square B (Figure 1B).

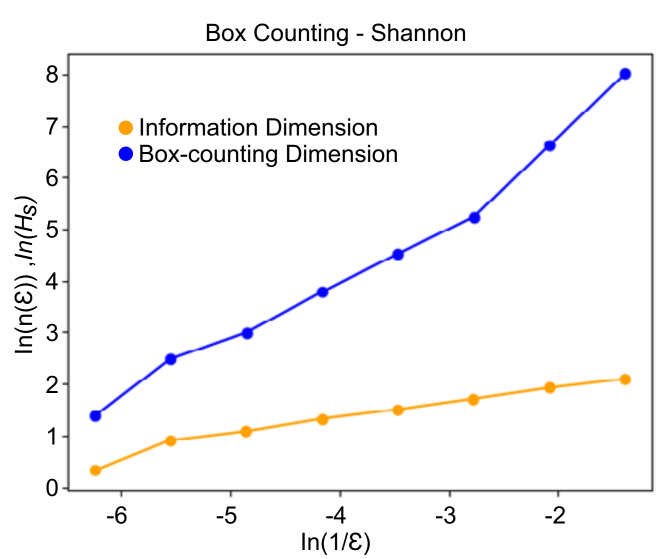


Figure 4. Information dimension and box-counting for square C (Figure 1C).

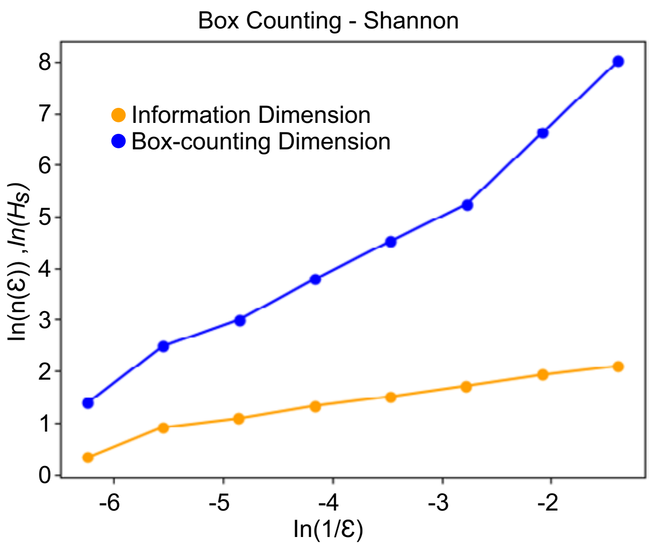


Figure 5. Information dimension and box-counting for square D (Figure 1D).

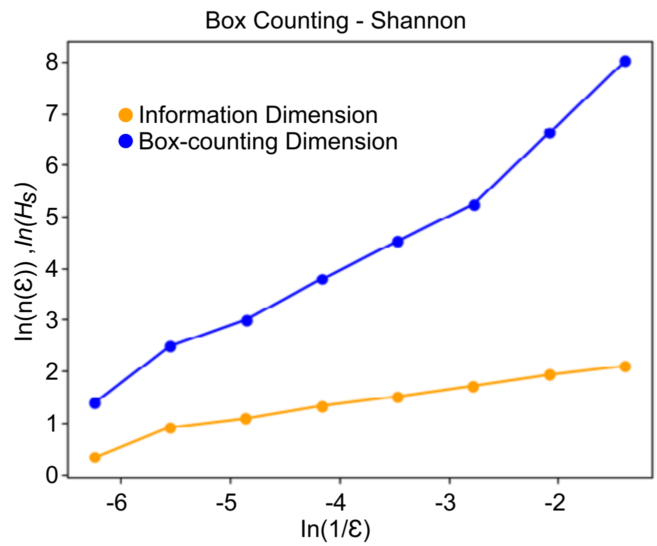
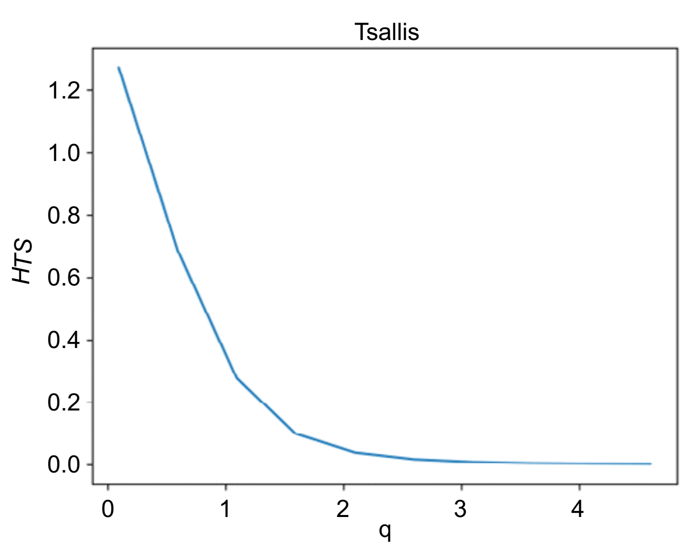


Table 2. Result of the box-counting dimension and Tsallis.

Square	**D**$_{Box}$	**D**$_{S}$
A	1.2796710276704335	0.7899382233835297
B	1.2796710276704335	0.7976331670012815
C	1.2796710276704335	0.8054932828767232
D	1.2796710276704335	0.8126195937450961

Figure 6. Dimension of information with Tsallis entropy of Figure 1A.



dimensions. This underscores entropy's sensitivity to the distribution of information within each box, rendering it a potent tool for detecting irregularities.

Moreover, the introduction of Tsallis entropy in this study facilitated a heightened emphasis on the information within each box, courtesy of its entropic parameter q. This parameter modulated information sensitivity, proving beneficial in various practical applications. The significance of this research lies in empirically demonstrating the efficacy of the information dimension in discerning border perturbations in flat objects. This method holds promise across diverse domains, including ecology, pathology detection in cells and leaves, and forest monitoring. The results underscore the information dimension's utility as a valuable tool for analyzing border perturbations and, consequently, object complexity. This opens avenues for further advancements and future applications.

References

1. Bekenstein JD. Some comments on Boltzmann-Gibbs statistical mechanics. Phys Rev 1973;7.
2. Falconer K. Fractal Geometry – Mathematical Foundations and Applications. Second Edition, John Wiley & Filhos, Ltda, 2003.
3. Mandelbrot BB. The fractal geometry of nature. [S.l.]: WH freeman New York, 173, 1983.
4. Seuront L. Fractals and multifractals in ecology and aquatic science. Taylor e Francis Group LLC, 2010.