

Article

Point Divergence Gain and Multidimensional Data Sequences Analysis

Renata Rychtáriková ^{1,*}, Jan Korbel ^{2,3,4}, Petr Macháček ¹ and Dalibor Štys ¹

¹ Institute of Complex Systems, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Kompetenzzentrum MechanoBiologie in Regenerativer Medizin, Faculty of Fisheries and Protection of Waters, University of South Bohemia in České Budějovice, Zámek 136, 373 33 Nové Hradky, Czech Republic; info@imagecode.eu (P.M.); stys@frov.jcu.cz (D.Š.)

² Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria; korbeja2@fjfi.cvut.cz

³ Complexity Science Hub Vienna, Josefstädter Strasse 39, 1080 Vienna, Austria

⁴ Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19 Prague, Czech Republic

* Correspondence: rrychtarikova@frov.jcu.cz; Tel.: +420-387-773-833

Received: 27 December 2017; Accepted: 30 January 2018; Published: 3 February 2018

Abstract: We introduce novel information-entropic variables—a Point Divergence Gain ($\Omega_{\alpha}^{(l \rightarrow m)}$), a Point Divergence Gain Entropy (I_{α}), and a Point Divergence Gain Entropy Density (P_{α})—which are derived from the Rényi entropy and describe spatio-temporal changes between two consecutive discrete multidimensional distributions. The behavior of $\Omega_{\alpha}^{(l \rightarrow m)}$ is simulated for typical distributions and, together with I_{α} and P_{α} , applied in analysis and characterization of series of multidimensional datasets of computer-based and real images.

Keywords: point divergence gain (PDG); Rényi entropy; data processing

1. Introduction

Extracting the information from raw data obtained from, e.g., a set of experiments, is a challenging task. Quantifying the information gained by a single point of a time series, a pixel in an image, or a single measurement is important in understanding which points bring the most information about the underlying system. This task is especially delicate in case of time-series and image processing because the information is not only stored in the elements, but also in the interactions between successive points in a time series. Similar, when extracting information from an image, not all pixels have the same information content. This type of information is sometimes called local information because the information depends not only on the frequency of the phenomenon but also on the position of the element in the structure. The most important task is to identify the sources of information and to quantify them. Naturally, it is possible to use standard data-processing techniques based on quantities from information theory like, e.g., Kullback–Leibler divergence. On the other hand, the mathematical rigorousness is typically compensated by an increased computational complexity. For this end, a simple quantity called Point Information Gain and its relative macroscopic variables—a Point Information Gain Entropy and a Point Information Gain Entropy Density—were introduced in [1]. In [2], mathematical properties of the Point Information Gain were extensively discussed and applications to real-image data processing were pointed out. From the mathematical point of view, the Point Information Gain represents a change of information after removing an element of a particular phenomena from a distribution. The method is based on the Rényi entropy, which has been already extensively used in multifractal analysis and data processing (see e.g., Refs. [2–5] and references therein).

In this article, we introduce an analogous variable to the Point Information Gain. This new variable locally determines an information change after an exchange of a given element in a discrete set. We use a simple concept of entropy difference between the original set and the set with the exchanged element. The resulting value is called Point Divergence Gain $\Omega_\alpha^{(l \rightarrow m)}$ [6,7]. The main idea is to describe the importance of changes in the series of images (typically representing a video record from an experiment) and extract the most important information from it. Similar to the Point Information Gain Entropy and the Point Information Gain Entropy Density, the macroscopic variables called a Point Divergence Gain Entropy I_α and a Point Divergence Gain Entropy Density P_α are defined to characterize subsequent changes in a multidimensional discrete distribution by one number. The goal of this article is to examine and demonstrate some properties of these variables and use them for examination of time-spatial changes of information in sets of discrete multidimensional data, namely series of images in image processing and analysis, after the exchange of a pixel of a particular intensity for a pixel at the same position in the consecutive image. The main reason for choosing the Point Divergence Gain as the relevant quantity for the analysis of spatio-temporal changes is the fact that it represents an information gain of each pixel change. One can also consider model-based approaches based on the theory of random-fields, which can be more predictive in some cases. On the other hand, the model-free approach based on entropy gives us typically more relevant information for real data, where it is typically difficult to find an appropriate model. For the overview of model-based approaches in the random field theory, one can consult, e.g., Refs. [8–10].

The paper is organized as follows: in Section 2, we define the main quantity of the paper, i.e., the Point Divergence Gain and the related quantities and discuss its theoretical properties. In Section 3, we show applications of the Point Divergence Gain to image processing for both computer-based and real sequences of images. We show that the Point Divergence Gain can be used as a measure of difference for clustering methods and detects the most prominent behaviour of a system. In Section 4, we explain the presented methods and finer technical details necessary for the analysis including algorithms. Section 5 is dedicated to conclusions. All image data, scripts for histogram processing, and Image Info Extractor Professional software for image processing are available via [sftp://160.217.215.193:13332/pdg](ftp://160.217.215.193:13332/pdg) (user: anonymous; password: anonymous.).

2. Basic Properties of Point Divergence Gain and Derived Quantities

2.1. Point Divergence Gain

Recently, a quantity called Point Information Gain (PIG, $\Gamma_\alpha^{(i)}$) [6,7] and its generalization based on the Rényi entropy [2] have been introduced. We show how to apply the concept of PIG to sequence of multidimensional data frames.

Let us assume a set of variables with k possible outcomes (e.g., possible colours of each pixel). The $\Gamma_\alpha^{(i)}$ is a simple variable based on entropy difference and enables us to quantify an information gain of each phenomenon. It is simply defined as a difference between entropy of an original discrete distribution

$$P = \{p_j\}_{j=1}^k = \left\{ \frac{n_1}{n}, \dots, \frac{n_k}{n} \right\}, \quad (1)$$

which typically describes a frequency histogram of possible outcomes. Let us also define a distribution, where one occurrence of the i -th phenomenon is omitted, i.e.,

$$P^{(i)} = \left\{ p_j^{(i)} \right\}_{j=1}^k = \left\{ \frac{n_1}{n-1}, \dots, \frac{n_i-1}{n-1}, \dots, \frac{n_k}{n-1} \right\}. \quad (2)$$

Thus, the Point Information Gain is defined as

$$\Gamma_\alpha^{(i)} \equiv \Gamma_\alpha^{(i)}(P) = \mathcal{H}_\alpha(P^{(i)}) - \mathcal{H}_\alpha(P), \quad (3)$$

where \mathcal{H}_α is the Rényi entropy (Despite all computer implementations being calculated as \log_2 , the following derivations are written in natural logarithm, i.e., \ln .)

$$\mathcal{H}_\alpha(P) = \frac{1}{\alpha - 1} \ln \sum_i p_i^\alpha. \tag{4}$$

The Rényi entropy represents a one-parametric class of information quantities tightly related to multifractal dynamics and enables us to focus on certain parts of the distribution [11]. Unlike the typically used Rényi’s relative entropy [3,4,11–17], the Point Information Gain $\Gamma_\alpha^{(i)}$ is a simple, computationally tractable quantity. Its mathematical properties have been extensively discussed in [2]. On the same basis, we can define a Point Divergence Gain (PDG, $\Omega_\alpha^{(l \rightarrow m)}$), where a discrete distribution $P^{(i)}$ is replaced by a distribution

$$P^{(l \rightarrow m)} = \left\{ p_j^{(l \rightarrow m)} \right\}_{j=1}^k = \left\{ \frac{n_1}{n}, \dots, \frac{n_l - 1}{n}, \dots, \frac{n_m + 1}{n}, \dots, \frac{n_k}{n} \right\}, \tag{5}$$

which can be obtained from the original distribution P , where the occurrence of the examined l -th phenomenon ($n_l \in \mathbb{N}^+$) is removed and supplied by a point of the occurrence of the m -th phenomenon ($n_m \in \mathbb{N}_0$). The main idea behind the definition is to quantify the information change in the subsequent image, if only one point is changed. Analogous to the Point Information Gain $\Gamma_\alpha^{(i)}$, the Point Divergence Gain can be defined as

$$\Omega_\alpha^{(l \rightarrow m)} \equiv \Omega_\alpha^{(l \rightarrow m)}(P) = \mathcal{H}_\alpha \left(P^{(l \rightarrow m)} \right) - \mathcal{H}_\alpha(P). \tag{6}$$

Let us first show its connection to the Point Information Gain $\Gamma_\alpha^{(i)}$. Since $P^{(l)} = P^{(l \rightarrow m, m)}$, it is possible to express the Point Divergence Gain as

$$\Omega_\alpha^{(l \rightarrow m)}(P) = \mathcal{H}_\alpha \left(P^{(l \rightarrow m)} \right) - \mathcal{H}_\alpha \left(P^{(l \rightarrow m, m)} \right) + \mathcal{H}_\alpha \left(P^{(l)} \right) - \mathcal{H}_\alpha(P) = \Gamma_\alpha^{(l)}(P) - \Gamma_\alpha^{(m)}(P^{(l \rightarrow m)}). \tag{7}$$

Let us investigate mathematical properties of the PDG. The $\Omega_\alpha^{(l \rightarrow m)}$ can be rewritten as

$$\begin{aligned} \Omega_\alpha^{(l \rightarrow m)} &= \mathcal{H}_\alpha \left(P^{(l \rightarrow m)} \right) - \mathcal{H}_\alpha(P) \\ &= \frac{1}{1 - \alpha} \ln \left(\sum_{j=1}^k \left(p_j^{(l \rightarrow m)} \right)^\alpha \right) - \frac{1}{1 - \alpha} \ln \left(\sum_{j=1}^k p_j^\alpha \right) = \frac{1}{1 - \alpha} \ln \left(\frac{\sum_{j=1}^k \left(p_j^{(l \rightarrow m)} \right)^\alpha}{\sum_{i=1}^k p_i^\alpha} \right). \end{aligned} \tag{8}$$

By plugging the relative frequencies from Equations (1) and (5) into Equation (8), we obtain

$$\begin{aligned} \Omega_\alpha^{(l \rightarrow m)} &= \frac{1}{1 - \alpha} \ln \left[\frac{(n_l - 1)^\alpha + (n_m + 1)^\alpha + \sum_{j=1, j \neq l, m}^k n_j^\alpha}{\sum_{j=1}^k n_j^\alpha} \right] \\ &= \frac{1}{1 - \alpha} \ln \left[\frac{(n_l - 1)^\alpha + (n_m + 1)^\alpha + \sum_{j=1}^k n_j^\alpha - n_l^\alpha - n_m^\alpha}{\sum_{j=1}^k n_j^\alpha} \right] \\ &= \frac{1}{1 - \alpha} \ln \left[\frac{(n_l - 1)^\alpha - n_l^\alpha + (n_m + 1)^\alpha - n_m^\alpha}{\sum_{j=1}^k n_j^\alpha} + 1 \right]. \end{aligned} \tag{9}$$

As seen in Equation (9), the variable $\Omega_\alpha^{(l \rightarrow m)}$ does not depend (contrary to the $\Gamma_\alpha^{(i)}$) on n but depends only on the number of elements of each phenomenon j . In Equation (9), let us design the nominator $\sum_{j=1}^k n_j^\alpha$, which is constant and related to the original distribution (histogram) of elements and to the parameter α , as \mathcal{C}_α . It gives us the final form

$$\Omega_{\alpha}^{(l \rightarrow m)} = \frac{1}{1 - \alpha} \ln \left[\frac{(n_l - 1)^{\alpha} - n_l^{\alpha} + (n_m + 1)^{\alpha} - n_m^{\alpha}}{C_{\alpha}} + 1 \right]. \quad (10)$$

Equation (10) demonstrates that, for a particular distribution, $\Omega_{\alpha}^{(l \rightarrow m)}$ is a function only of the parameter α and frequencies of occurrences of the phenomena n_l and n_m in the original distribution, between which the exchange of the element occurs. Equation (10) further shows that if the exchange of the element occurs between phenomena l and m of the same (similar) frequencies of occurrence (i.e., $n_l \approx n_m$), the value of $\Omega_{\alpha}^{(l \rightarrow m)}$ equals 0. If we remove a rare point and supply it by a high-frequency point (i.e., $n_l \ll n_m$), the value of $\Omega_{\alpha}^{(l \rightarrow m)}$ is negative, and *vice versa*. Low values of parameter α separate low-frequency events as $\Omega_{\alpha}^{(l \rightarrow m)} = 0$, whereas high α emphasize high-frequency events as $\Omega_{\alpha}^{(l \rightarrow m)} \gg 0$ or $\Omega_{\alpha}^{(l \rightarrow m)} \ll 0$ and merge rare events into $\Omega_{\alpha}^{(l \rightarrow m)} = 0$. With respect to the previous discussion and practical utilization of this notion, we emphasize that, for real systems with large n , the $\Omega_{\alpha}^{(l \rightarrow m)}$ are rather small numbers.

In the 3D plots of Figure 1, we demonstrate $\Omega_{\alpha}^{(l \rightarrow m)}$ -transformations of four thoroughly studied distributions—the Cauchy, Gauss (symmetrical), Lévy, and Rayleigh distribution (asymmetric; all specified in Section 4.1)—for $\alpha = \{0.5; 1.0; 2.0; 4.0\}$, where each point presents the exchange of the element between bins l and m (Algorithm 1). In this case, the (a)symmetry of the distribution is always maintained.

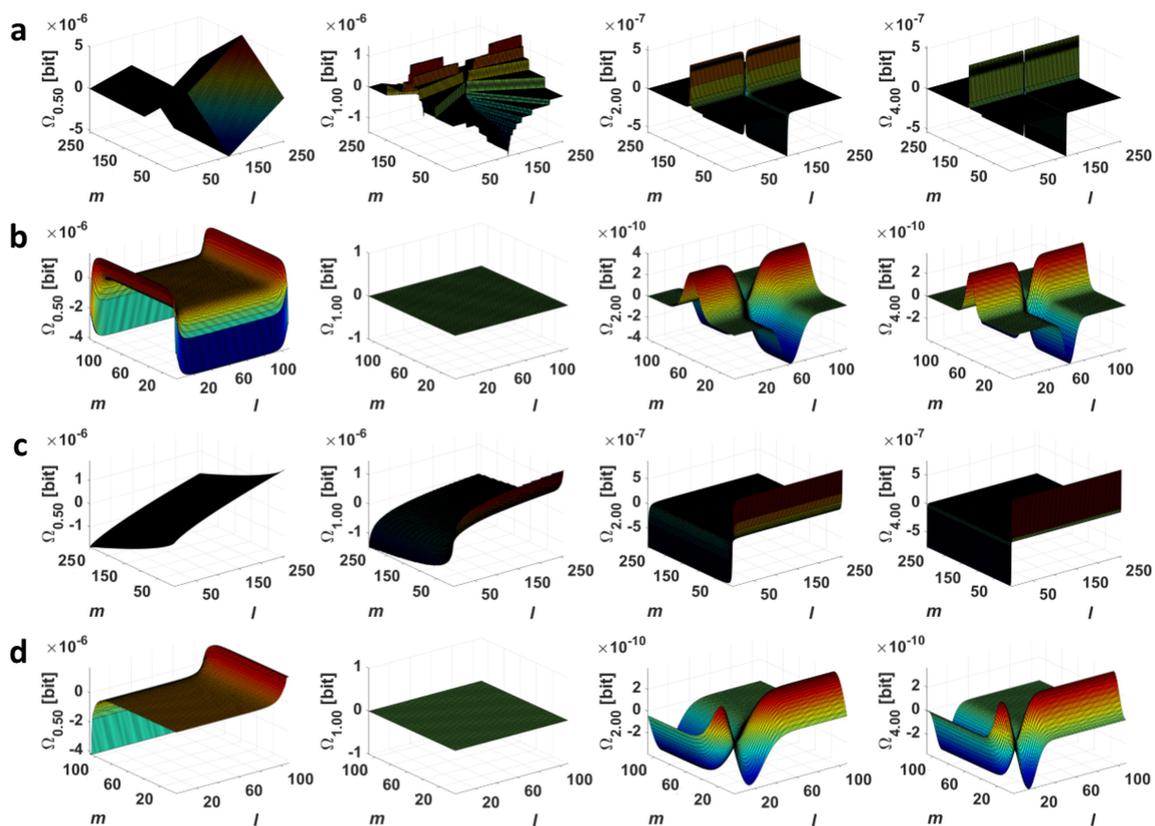


Figure 1. The Ω_{α} -transformations of the discrete (a) Cauchy; (b) Gauss; (c) Lévy; and (d) Rayleigh distribution for $\alpha = \{0.5; 1.0; 2.0; 4.0\}$ (Section 4.1).

Algorithm 1: Calculation of a point divergence gain matrix (Ω_α) for typical histograms.

```

Input:  $n$ -bin histogram  $\mathbf{h}$ ;  $\alpha$ , where  $\alpha \geq 0 \wedge \alpha \neq 1$ 
Output:  $\Omega_\alpha$ 

1  $\Omega_\alpha = \text{zeros}(n, n);$     % create a zero square matrix  $\Omega_\alpha$  of the size of  $n \times n$ 
2  $C_\alpha = \text{sum}(\mathbf{h}.\wedge\alpha);$     % calculate the constant  $C_\alpha$  for the given distribution and  $\alpha$ 

3 for  $l = 1$  to  $n$  do
4     if  $\mathbf{h}(l) \neq 0$  then
5         for  $l = 1$  to  $n$  do
6              $\Omega_\alpha(l, m) =$ 
7                  $\log_2(((\mathbf{h}(l) - 1)^\wedge\alpha - \mathbf{h}(l)^\wedge\alpha + (\mathbf{h}(m) + 1)^\wedge\alpha - \mathbf{h}(m)^\wedge\alpha) / C_\alpha + 1) / (1 - \alpha));$ 
8             end
9         else
10             $\Omega_\alpha(l,:) = \text{NaN};$ 
11        end
12    % if the bin  $l$  of the histogram  $\mathbf{h}$  is occupied, calculate  $\Omega_\alpha$  at each position  $(l, m)$  according to
13    Equation (10), else set the not-a-number into the row  $l$  of the  $\Omega_\alpha$  matrix
14 end

```

Now we will consider the specific case $\alpha = 2$ (collision entropy) for which Equation (10) can be simplified to

$$\Omega_2^{(l \rightarrow m)} = -\ln \left[\frac{2}{C_2} (n_m - n_l + 1) + 1 \right] = -\ln \left[\frac{2}{C_2} (\Delta n^{(l \rightarrow m)} + 1) + 1 \right]. \tag{11}$$

For a specific difference $\Delta n^{(x \rightarrow y)} = D$, Equation (11) can be approximated by the 1st-order Taylor sequence

$$\begin{aligned} \Omega_2^{(l \rightarrow m)} &\approx -\ln \left[\frac{2}{C_2} (D + 1) + 1 \right] - \frac{2}{2(D + 1) + C_2} (\Delta n^{(l \rightarrow m)} - D) \\ &= -\frac{2}{2D + 2 + C_2} \Delta n^{(l \rightarrow m)} + \frac{2D}{2D + 2 + C_2} - \ln \left[\frac{2D}{C_2} + C_2 + 1 \right]. \end{aligned} \tag{12}$$

Equations (11) and (12) show that, for each unique $\Delta n^{(x \rightarrow y)}$, the $\Omega_2^{(l \rightarrow m)}$ depends only on the difference between the bins l and m , which the exchange of the element occurs between, and this dependence is almost linear. In other words, this explains why, for all distributions in Figure 2, the dependencies $\Omega_2^{(l \rightarrow m)} = f(n_m, n_m - n_l)$ are planes.

For $\alpha \rightarrow 1$, the Rényi entropy becomes the ordinary Shannon entropy [18] and we obtain (cf. Equation (4))

$$\mathcal{H}_1(P) = -\sum_{j=1}^k p_j \ln p_j = -\sum_{j=1}^k \frac{n_j}{n} \ln \frac{n_j}{n} = -\sum_{j=1, j \neq l, m}^k \frac{n_j}{n} \ln \frac{n_j}{n} - \frac{n_m}{n} \ln \frac{n_m}{n} - \frac{n_l}{n} \ln \frac{n_l}{n} \tag{13}$$

and

$$\mathcal{H}_1(P^{(l \rightarrow m)}) = -\frac{n_m + 1}{n} \ln \frac{n_m + 1}{n} - \frac{n_l - 1}{n} \ln \frac{n_l - 1}{n} - \sum_{j=1, j \neq l, m}^k \frac{n_j}{n} \ln \frac{n_j}{n}. \tag{14}$$

The difference of these entropies (cf. Equation (9)) is gradually giving

$$\begin{aligned}
 \Omega_1^{(l \rightarrow m)} &= -\frac{n_m + 1}{n} \ln \frac{n_m + 1}{n} - \frac{n_l - 1}{n} \ln \frac{n_l - 1}{n} + \frac{n_m}{n} \ln \frac{n_m}{n} + \frac{n_l}{n} \ln \frac{n_l}{n} \\
 &= -\frac{n_m + 1}{n} \ln(n_m + 1) + \frac{n_m + 1}{n} \ln n - \frac{n_l - 1}{n} \ln(n_l - 1) + \frac{n_l - 1}{n} \ln n + \frac{n_m}{n} \ln n_m \\
 &\quad - \frac{n_m}{n} \ln n + \frac{n_l}{n} \ln n_l - \frac{n_l}{n} \ln n \\
 &= \underbrace{\left(\frac{n_m + 1}{n} + \frac{n_l - 1}{n} - \frac{n_m}{n} - \frac{n_l}{n} \right)}_{=0} \ln n - \frac{n_m}{n} \ln(n_m + 1) - \frac{1}{n} \ln(n_m + 1) - \frac{n_l}{n} \ln(n_l - 1) \\
 &\quad + \frac{1}{n} \ln(n_l - 1) + \frac{n_m}{n} \ln n_m + \frac{n_l}{n} \ln n_l \\
 &= \frac{1}{n} \left(n_m \ln \frac{n_m}{n_m + 1} + n_l \ln \frac{n_l}{n_l - 1} + \ln \frac{n_l - 1}{n_m + 1} \right). \tag{15}
 \end{aligned}$$

One can see that relation (15) is defined for $n_l \in \mathbb{N} \setminus \{0, 1\}$ and $n_m \in \mathbb{N}^+$ and is approximately equal to 0 for $n_l, n_m \gg 0$ (the Cauchy and Rayleigh distribution for $\alpha = 1$ in Figure 3).

For $n_l \in \mathbb{N}^+$ and $n_m \in \mathbb{N}_0$, from Equation (10), further implies:

1. If $\alpha = 0$, then $\Omega_0^{(l \rightarrow m)} = 0$.
2. If $\alpha \rightarrow \infty$, then $\Omega_\infty^{(l \rightarrow m)} \rightarrow 0$.

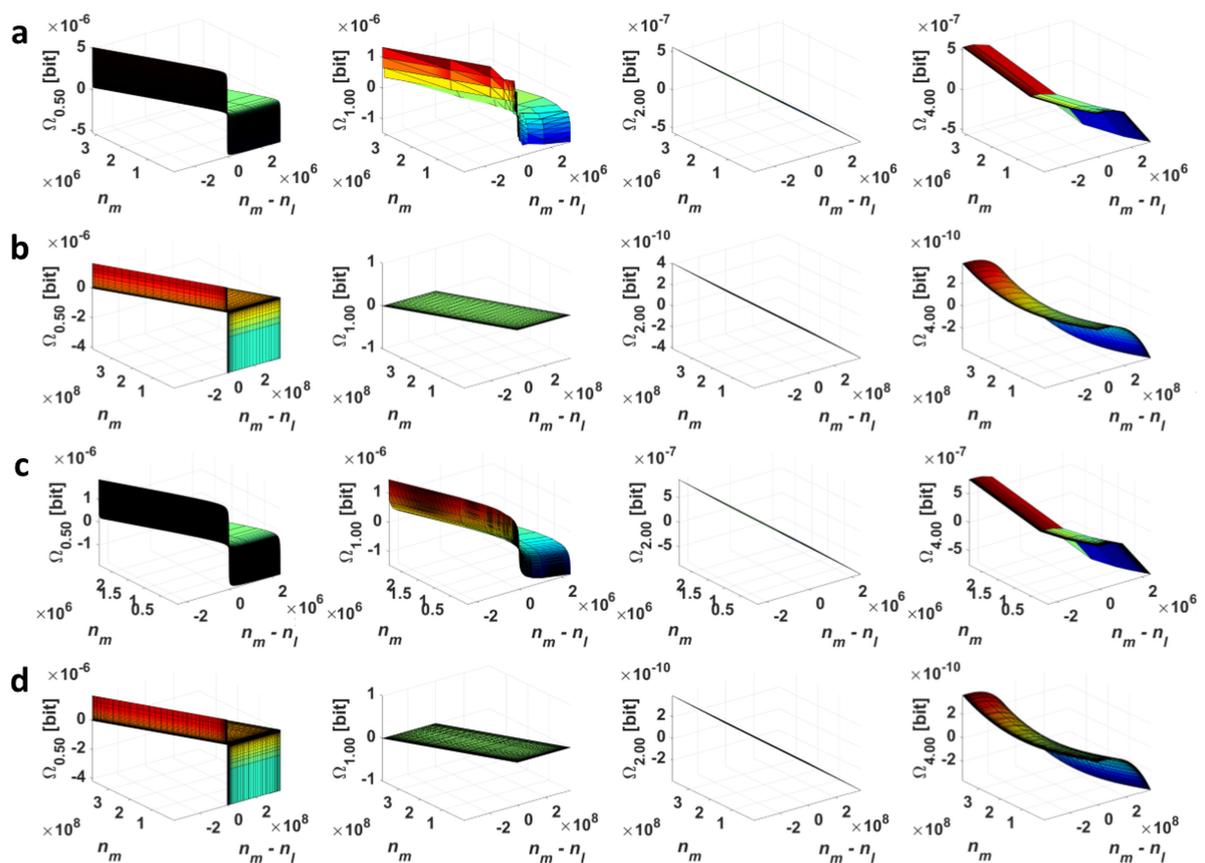


Figure 2. The dependencies $\Omega_\alpha = f(n_m, n_m - n_l)$ for the discrete (a) Cauchy; (b) Gauss; (c) Lévy; and (d) Rayleigh distribution at $\alpha = \{0.5; 1.0; 2.0; 4.0\}$ (Section 4.1).

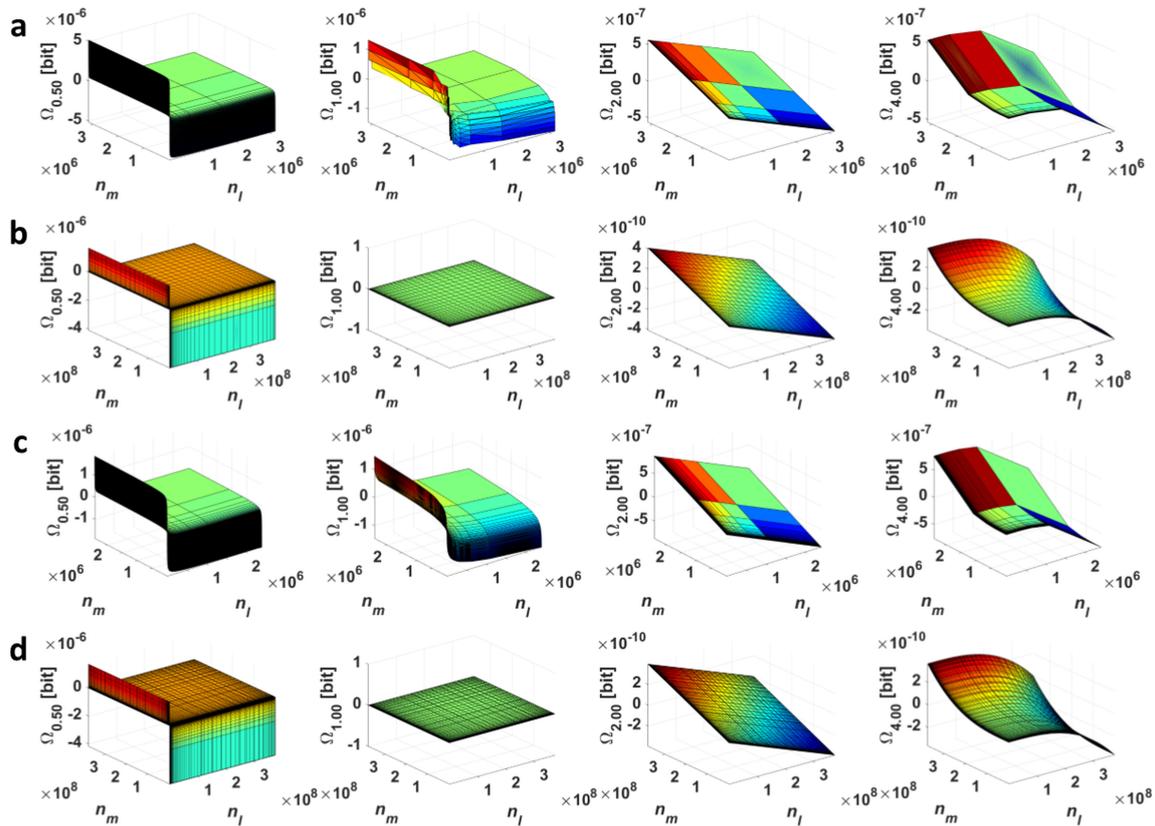


Figure 3. The dependencies $\Omega_\alpha = f(n_l, n_m)$ for the discrete (a) Cauchy; (b) Gauss; (c) Lévy; and (d) Rayleigh distribution at $\alpha = \{0.5; 1.0; 2.0; 4.0\}$ (Section 4.1).

2.2. Point Divergence Gain Entropy and Point Divergence Gain Entropy Density

In this section, we introduce two new variables that help us to investigate changes between two (typically consecutive) points of time series. A typical example can be provided by video processing, where each element of a time or spatial series is represented by a frame. Let us have two data frames $\mathcal{I}_a = \{a_1, \dots, a_n\}$ and $\mathcal{I}_b = \{b_1, \dots, b_n\}$ (For simplicity, we use only one index which corresponds to a one-dimensional frame. In case of images, we have typically two-dimensional frames and the elements are described by two indexes, e.g., x and y positions.). At each position $i \in \{1, \dots, n\}$, it is possible to replace the value a_i by the value of the following frame, i.e., b_i . The resulting $\Omega_\alpha^{(a_i \rightarrow b_i)}$ then quantifies how much information is gained/lost, when, at the i -th position, we replace the value a_i for the value b_i . A Point Divergence Gain Entropy (PDGE, I_α) is defined as a sum of absolute values of all PDGs for all pixels, i.e.,

$$I_\alpha(\mathcal{I}_a; \mathcal{I}_b) = \sum_{i=1}^n |\Omega_\alpha^{(a_i \rightarrow b_i)}| = \sum_{l=1}^k \sum_{m=1}^k n_{lm} |\Omega_\alpha^{(l \rightarrow m)}|, \tag{16}$$

where n_{lm} denotes the number of present substitutions $l \rightarrow m$, when we transform $\mathcal{I}_a \rightarrow \mathcal{I}_b$. The absolute value ensures that the contribution of the transformation of a rare point to a frequent point (negative Ω_α) and a frequent point to a rare point (positive Ω_α) do not cancel each other and both contribute to the resulting PDGE. Typically, appearance or disappearance of a rare point (and replacement by a frequent value—typically background colour) carries important information about the experiment. The PDGE can be understood as an absolute information change.

Moreover, it is possible to introduce other macroscopic quantity—a Point Divergence Gain Entropy Density (PDGED, P_α), where we do not sum over all pixels, but only over all realized transitions $l \rightarrow m$. Thus, the PDGED can be defined as

$$P_\alpha(\mathcal{I}_a; \mathcal{I}_b) = \sum_{l=1}^k \sum_{m=1}^k \chi_{lm} |\Omega_\alpha^{(l \rightarrow m)}|, \quad (17)$$

where

$$\chi_{lm} = \begin{cases} 1, & n_{lm} \geq 1, \\ 0, & n_{lm} = 0. \end{cases} \quad (18)$$

Let us emphasize that two transitions $a_1 \rightarrow b_1$ and $a_2 \rightarrow b_2$, where the frequencies of the occurrences of the phenomena a_1 and a_2 are equal and of the phenomena b_1 and b_2 are equal as well, give two unique values of the $\Omega_\alpha^{(a_i \rightarrow b_i)}$. In the computation of the PDGED, this is arranged by a hash function (Algorithm 2). We can understand the quantity PDGED as an absolute information change of all realized transitions of phenomena $m \rightarrow l$.

Algorithm 2: Calculation of a point information gain matrix (Ω_α) and values P_α and I_α for two consecutive images of a time-spatial series.

Input: 2 consecutive images \mathbf{I}_1 and \mathbf{I}_2 of the size $m \times n$; α , where $\alpha \geq 0 \wedge \alpha \neq 1$

Output: Ω_α

```

1 h = hist( $\mathbf{I}_1$ ); % create an intensity histogram h of the image  $\mathbf{I}_1$ 
2  $C_\alpha$  = sum( $\mathbf{h} \wedge \alpha$ ); % calculate the constant  $C_\alpha$  for the given distribution and  $\alpha$ 
3  $\Omega_\alpha$  =  $\mathbf{I}_1 \cdot 0$ ; % create a zero matrix  $\Omega_\alpha$  of the size of the  $\mathbf{I}_1$ 
4 hashMap = containers.Map; % declare an empty hash-map (the key-value array)

5 for  $i = 1$  to  $(m \times n)$  do
6    $\Omega_\alpha(i) = \log_2(((\mathbf{h}(\mathbf{I}_1(i+1)) - 1)^\alpha - \mathbf{h}(\mathbf{I}_1(i+1))^\alpha + (\mathbf{h}(\mathbf{I}_2(i+1)))^\alpha -$ 
7      $\mathbf{h}(\mathbf{I}_2(i+1))^\alpha) / C_\alpha + 1) / (1 - \alpha)$ ;
8     % for each element  $i$  of the image  $\mathbf{I}_1$ , calculate a value  $\Omega_\alpha$  after replacement of the intensity
9     at the position  $i$  in the histogram of image  $\mathbf{I}_1$  by the intensity at the same position in the
10    image  $\mathbf{I}_2$  (Equation (10))
11     $v = \mathbf{I}(i)$ ; % read a value of the element (intensity) at the position  $i$ 
12     $checkSum = \text{calcCheckSum}(\mathbf{h}, v)$ ;
13    % calculate  $checkSum$  using a hash-function effective enough (e.g., MD4, MD5, SHA1)
14    if not  $hashMap.isKey(checkSum)$  then
15       $hashMap(checkSum) = \Omega_\alpha(i)$ ;
16      % if the hash-map does not contain the key, insert a new element with the key
17       $checkSum$ , where the inserted value is the  $\Omega_\alpha$  at the position  $i$ 
18    end
19  end

20  $I_\alpha = \text{sum}(\text{sum}(\text{abs}(\Omega_\alpha)))$ ;
21 % calculate  $I_\alpha$  as a sum of all elements in the matrix  $\Omega_\alpha$  (Equation (16))
22  $P_\alpha = \text{sum}(\text{abs}(\text{values}(\mathbf{hashMap})))$ ;
23 % calculate  $P_\alpha$  as a sum of all elements in the matrix  $\mathbf{hashMap}$  (Equation (17))

```

If the aim is to assess the influence of elements of a high occurrence on the time-spatial changes in the image series, it is recommended to use PDGE where each element is weighted by its number of occurrences. If the aim is to suppress the influence of these extreme values, it is better to compute PDGED.

Let us consider a time-series \mathcal{V} , where each time step contains one frame, so $\mathcal{V} = \{\mathcal{I}_1, \mathcal{I}_2, \dots\}$. The series \mathcal{V} can be, e.g., a sequence of images (a video) obtained from some experiment, etc. For each time step, it is possible to calculate $I_\alpha(t) = I_\alpha(\mathcal{I}_t; \mathcal{I}_{t+s})$, resp. $P_\alpha(t) = P_\alpha(\mathcal{I}_t; \mathcal{I}_{t+s})$, where s is the time lag. Typically, we assume $s = 1$, i.e., consecutive frames with a constant time step.

3. Application of Point Divergence Gain and Its Entropies in Image Processing

The generalized Point Divergence Gain $\Omega_\alpha^{(l \rightarrow m)}$ in Equation (10) was originally used for characterization of dynamic changes in image series, namely in z-stacks of raw RGB data of unmodified live cells obtained via scanning along the z-axis using video-enhanced digital bright-field transmission microscopy [6,7]. In these two references, this new mathematical approach utilizes 8- and 12-bit intensity histograms of two consecutive images for pixel-by-pixel intensity weighted (parameterized) subtraction of these images to suppress the camera-based noise and to enhance the image contrast (In case of calibrated digital camera-based images, where the value of each point of the image reflects a number of incident photons, or, in case of computer-based images, it can be sufficient to use a simple subtraction for evaluation of time-spatial changes in the image series.).

For this paper, we chose other (grayscale) digital image series (Table 1) in order to demonstrate other applications of the PDG mathematical approach in image processing and analysis. Moreover, we newly introduce applications of the additive macroscopic variables Point Divergence Gain Entropy I_α and Point Divergence Gain Entropy Density P_α .

Table 1. Specifications of image series.

Series	Source	Bit-Depth	Number of Img.	Resolution	Origin
Toy Vehicle	[19]	8-bit	10	512 × 512	camera
Walter Cronkite	[19]	8-bit	16	256 × 256	camera
Simulated BZ	[20–22]	8-bit	10,521	1001 × 1001	computer-based ^a
Ring-fluorescence		12-bit	1058	548 × 720	experimental ^b
Ring-diffraction		8-bit ^c	1242	252 × 280	experimental ^b

^a A set of a noisy hotch-potch machine simulation of the Belousov–Zhabotinsky reaction [20–22] at 200 achievable states with the internal excitation of 10, and phase transition, internal excitation, and external neighbourhood kind of noise of 0, 0.25, and 0.15, respectively. ^b The microscopic series of a 6-μm standard microring (FocalCheck™, cat. No. F36909, Life Technologies™ (Eugene, OR, USA)) were acquired using the CellObserver microscope (Zeiss, Oberkochen, Germany) at the EMBL (Heidelberg, Germany). For both light processes, the green region of the visible spectrum was selected using an emission and transmission optical filter, respectively. In case of the diffraction, the point spread function was separated and the background intensities was disposed using Algorithm 1 in [7]. ^c The 12-bit depth was reduced using a Least Information Lost algorithm [23], which, by shifting the intensity bins, filled all empty bins in the histogram obtained from the whole data series up and rescaled these intensities between their minimal and maximal value.

3.1. Image Origin and Specification

Owing to the relation of the $\Omega_\alpha^{(l \rightarrow m)}$ to the Rényi entropy, the I_α and P_α as macroscopic variables can determine a fractal origin of images by plotting $I_\alpha = f_I(\alpha)$ and $P_\alpha = f_P(\alpha)$ spectra. If we deal with an image multifractality, the dependency $I_\alpha = f_I(\alpha)$ or the dependency $P_\alpha = f_P(\alpha)$ shows a peak. In case of a unifractality, these dependences are monotonous. It is demonstrated in Figures 4 and 5. There can be no doubts that the origin of the simulated Belousov–Zhabotinsky reaction (Figure 4) is multifractal. This statement is further strengthened by the courses of the dependencies $I_\alpha = f_I(\alpha)$ and $P_\alpha = f_P(\alpha)$, where we can see peaks with maxima at $\alpha \in (1, 2)$. On the contrary, a pair of images in Figure 5 (moving toys of cars) is a mixture of the objects of different fractal origin. In this case, whereas the course of $f_I(\alpha)$ is monotonous and thus shows a unifractal characteristics, the dependence

$f_P(\alpha)$ has a maximum at $\alpha = 0.6$ and thus demonstrates some multifractal features in the image. This is due to the fact that, since each information contribution is counted only once, the P_α is more sensitive to the phenomena, which occur less frequently in the image. The monotonic course of the P_α would be achieved only when a sequence of time-evolved Euclidian objects was transformed into the values $\Omega_\alpha^{(l \rightarrow m)}$.

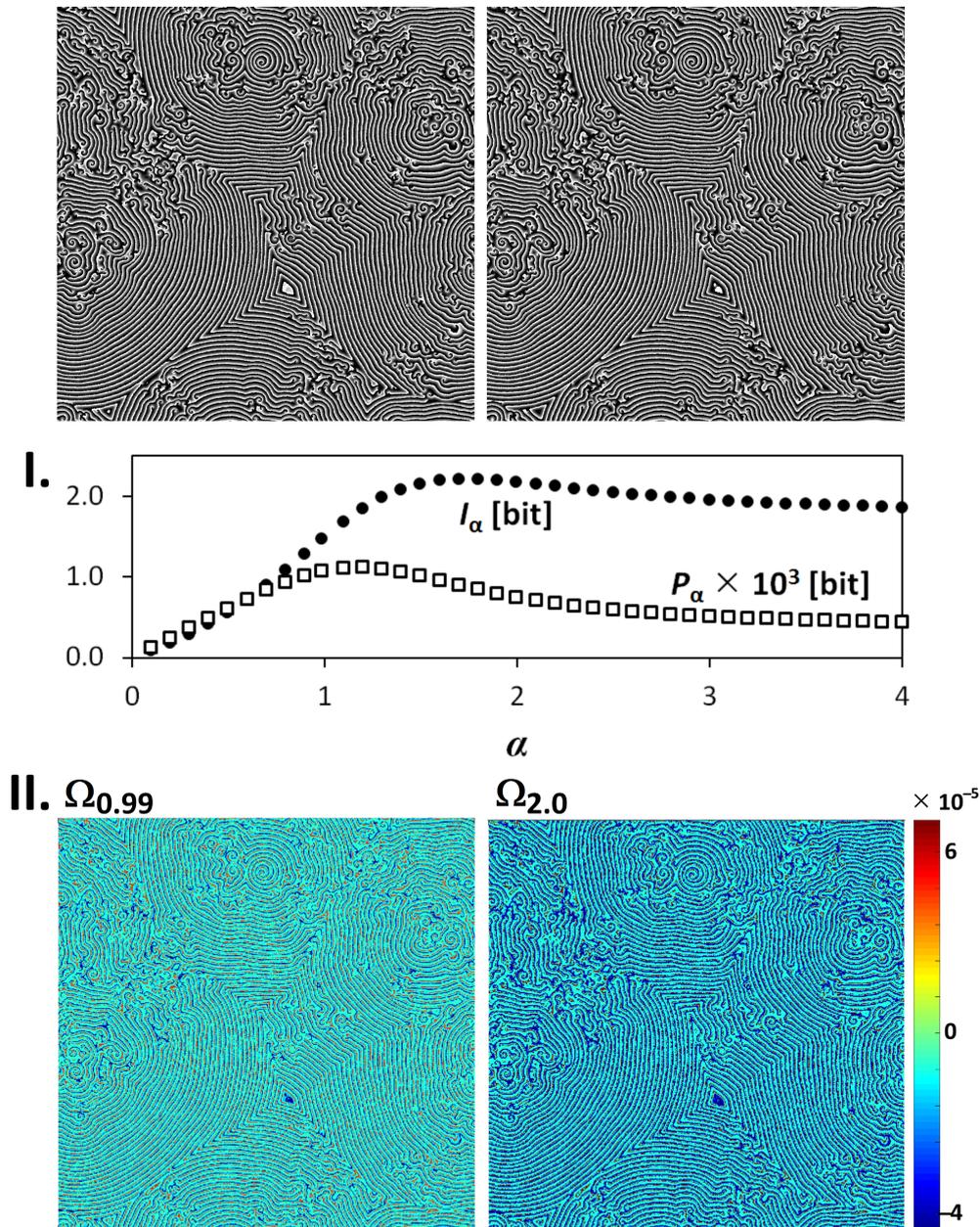


Figure 4. The I_α , P_α , and Ω_α for a pair of multifractal grayscale images. **I.** The I_α and P_α spectra, **II.** 8-bit visualization of Ω_α -values for $\alpha = \{0.99; 2.0\}$.

As mentioned in Section 2.2, the variables I_α and P_α measure absolute information change between a pair of images and characterize a similarity between these images. Therefore, these variables can find a practical utilization in auto-focusing in both light and electron digital microscopy. The in-focus object can be defined as an image with the global extreme of I_α or P_α . In other characteristics, this image fulfils the Nijboer–Zernike definition [24]: it is the smallest and darkest image in light or electron diffraction or the smallest and brightest image in light fluorescence (Section 3.3).

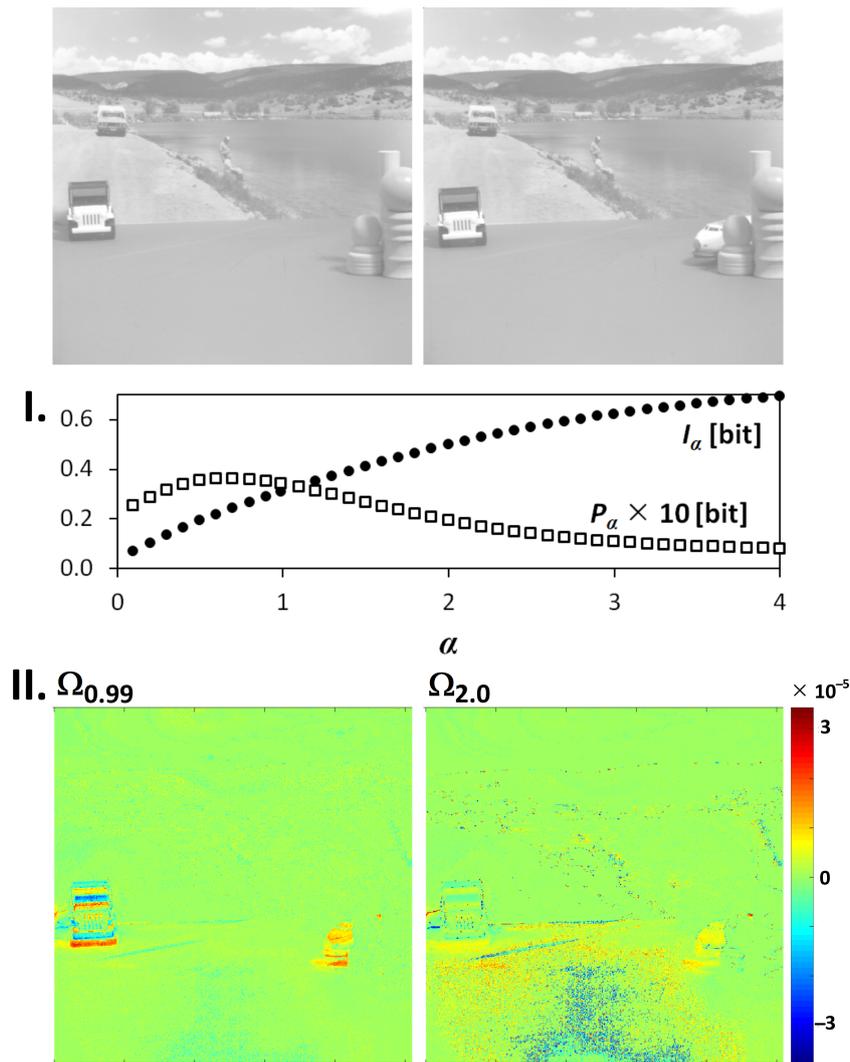


Figure 5. The I_α , P_α , and Ω_α for a pair of real-life grayscale images. I. the I_α and P_α spectra; II. 8-bit visualization of Ω_α -values for $\alpha = \{0.99; 2.0\}$.

3.2. Image Filtering and Segmentation

Segmentation is a type of filtering of specific features in an image. The parameter α and the related value of $\Omega_\alpha^{(l \rightarrow m)}$ enable us to filter the parts of two consecutive images, which are either stable or differently variable in time. This can be employed in a 3D image reconstruction by thresholding and joining $\Omega_\alpha^{(l \rightarrow m)} = 0$ from two consecutive images or in image tracking via thresholding of the highest and lowest $\Omega_\alpha^{(l \rightarrow m)}$ in a first image and the following image, respectively.

This is illustrated using simple examples in Figures 4 and 5 where the highest (red-coded) and lowest (blue-coded) values of the $\Omega_\alpha^{(l \rightarrow m)}$ show the position of the object in the second and the first image of the image sequence, respectively. Compared with the $\Omega_{0.99}^{(l \rightarrow m)}$, the variance between the extremes of the $\Omega_{2.00}^{(l \rightarrow m)}$ is wider and the number of points $\Omega_{2.00}^{(l \rightarrow m)} = 0$ is lower.

In digital light transmission microscopy, this mathematical method enabled us to find time stable intracellular objects inside live mammalian cells from consecutive pixels that fulfilled the equality $\Omega_\alpha^{(l \rightarrow m)} = 0$ for $\alpha = 4.00$ [6] or $\alpha = 5.00$ [7]. In these cases, the high value of α ensured merging rare points in the image, suppressing the camera noise that was reflected in the images and, thus, modelling the shape of organelles. The rest of image escaped the observation. In the next paper [25], this method was extended to widefield fluorescent data.

As in the case of the Point Information Gain [2], the process of image segmentation of objects of a certain shape can be further improved by usage of the surroundings of this shape from which the intensity histogram is created for each pixel in the image.

3.3. Clustering of Image Sets

Finally, we used the Point Divergence Gain to detect the most relevant information contained in a sequence of images, capturing, e.g., an experiment. For this end, we used I_α or P_α as quantities of information change in the consecutive images and applied the clustering methods on them. The values of I_α or P_α are small numbers (Section 2.1). Due to the computation rounding of small numbers of the I_α and the P_α and for a better characterization of the image multifractality, in clustering, we use α -dependent spectra of these variables than a sole number at one α .

The dependence of the label of the cluster on the order of the image in the series is the smoothest for joint vectors $[I_\alpha, P_\alpha]$. The similarity of these vectors (and thus images as well) is described in a space of principal components, e.g., [26], and classified by standard clustering algorithms such as k-means++ algorithm [27]. In comparison to the entropies and entropy densities related to the $\Gamma_\alpha^{(i)}$, the clustering using the I_α and the P_α is more sensitive to changes in the patterns (intensities) and does not require other specification of images by local entropies computed from a specific type of surroundings around each pixel.

The described clustering method was examined on z-stacks obtained using light microscopy. The z-stacks were classified into 2–6 clusters (groups) when patterns of each image was described by 26 numbers, i.e., by vectors $[I_\alpha, P_\alpha]$ at 13 α (Figures 6a and 7a). These clusters were evaluated on the basis of the sizes of intensity changes between images. These five classification graphs of the gradually splitting clusters (Figures 6a and 7a, middle) further demonstrate the mutual similarity among the micrographs in each data series. The typical (middle) image of each cluster is shown in Figures 6b and 7b.

Firstly, we shall deal with a z-stack with 1057 images of a microring obtained using a widefield fluorescent microscope. The results of clustering illustrate a canonically repetitive properties of the so-called point spread function as the image of the observed object goes to and from its focus. In this case, the image group containing the real focus of the maximal I_α and P_α at low α (Section 3.1) is successfully determined by clustering into two clusters (Figure 6a). However, we will aim for a description of the results for five clusters. The central Cluster 5 (94 images) can be called an object's focal region with image levels where parts of the object have their own focus. The in-focus cluster is asymmetrically surrounded by Cluster 4 (131 and 53 images below and above Cluster 5, respectively), which was set on the basis of the occurrence of the lower peaks of I_α and P_α at low α . Cluster 3 (190 and 150 images below and above the focus, respectively) is typical of constant I_α and P_α for all α . Cluster 2 contains img. 176–214 and the last 126 images. These images are characteristic of constant I_α and decreasing/increasing P_α at $\alpha \geq 2$. Cluster 1 (the first 175 images) is prevalently dominated by increasing I_α and decreasing P_α at high α .

Before the calculation of the I_α and P_α , the undesirable background intensities were removed from the images obtained using optical transmission microscopy. The rest of each image was rescaled into 8 bits (Section 4.2). The results of clustering of these images (Figure 7a) are similar to fluorescent data (Figure 6a). The light transmission point spread function is symmetrical around its focus as well but the pixels at the same x, y -positions below and above the focus have opposite, dark vs. bright, intensities. Furthermore, the transitional regions between the clusters are longer than for the fluorescent data. The central, in-focus, part of the z-stack (img. 427–561 in Cluster 4) with the highest peaks of I_α and P_α is unambiguously separated using four clusters. The focus itself lies at the 505th image. This central part of the z-stack is surrounded by eight groups of images which were, due to their similarity, objectively classified into three clusters. Cluster 1 was formed by images 1–78, 376–426, and 562–661. These images show peaks of middle values of the I_α and P_α . Images 79–153, 292–375, and 662–703 were classified into Cluster 2 (dominated by the local minimum of the I_α at $\alpha < 1$). Cluster 3

is related to the images with the lowest values of the I_α together with the lowest values and local peaks of the P_α for $\alpha < 1$ and for $\alpha < 1$, respectively. This cluster contains images 154–291 and the last 537 images of the series.

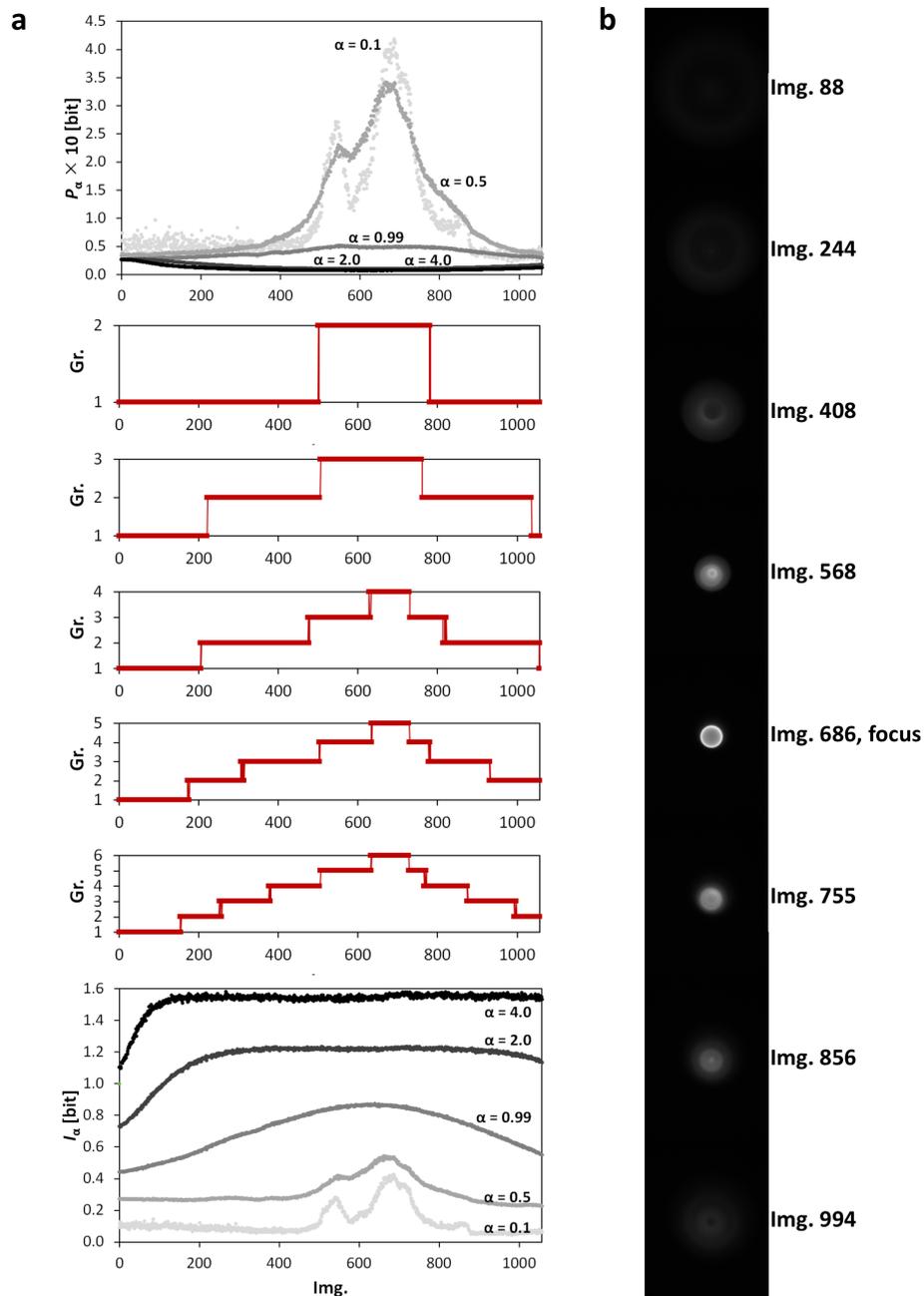


Figure 6. The results of clustering of a z-stack of grayscale microscopic images of a microring obtained using a fluorescence mode. (a) the dependencies of (upper) the P_α and (lower) the I_α vs. order of the image in the z-stack for $\alpha = \{0.5; 0.99; 2.0; 4.0\}$ and (middle) clustering (k-means, squared Euclidian distance, 2–6 groups) of the z-stack using connected spectra $[I_\alpha, P_\alpha]$ for $\alpha = \{0.1; 0.3; 0.5; 0.7; 0.99; 1.3; 1.5; 1.7; 2.0; 2.5; 3.0; 3.5; 4.0\}$; (b) the typical (middle) group’s images for clustering into five groups (in (a), middle). The original 12-bit images are visualized in 8 bits using the Least Information Loss conversion [23].

Let us mention that, in the clustering process, the I_α and P_α can recognize outliers such as incorrectly saved images or images with illumination artifacts.

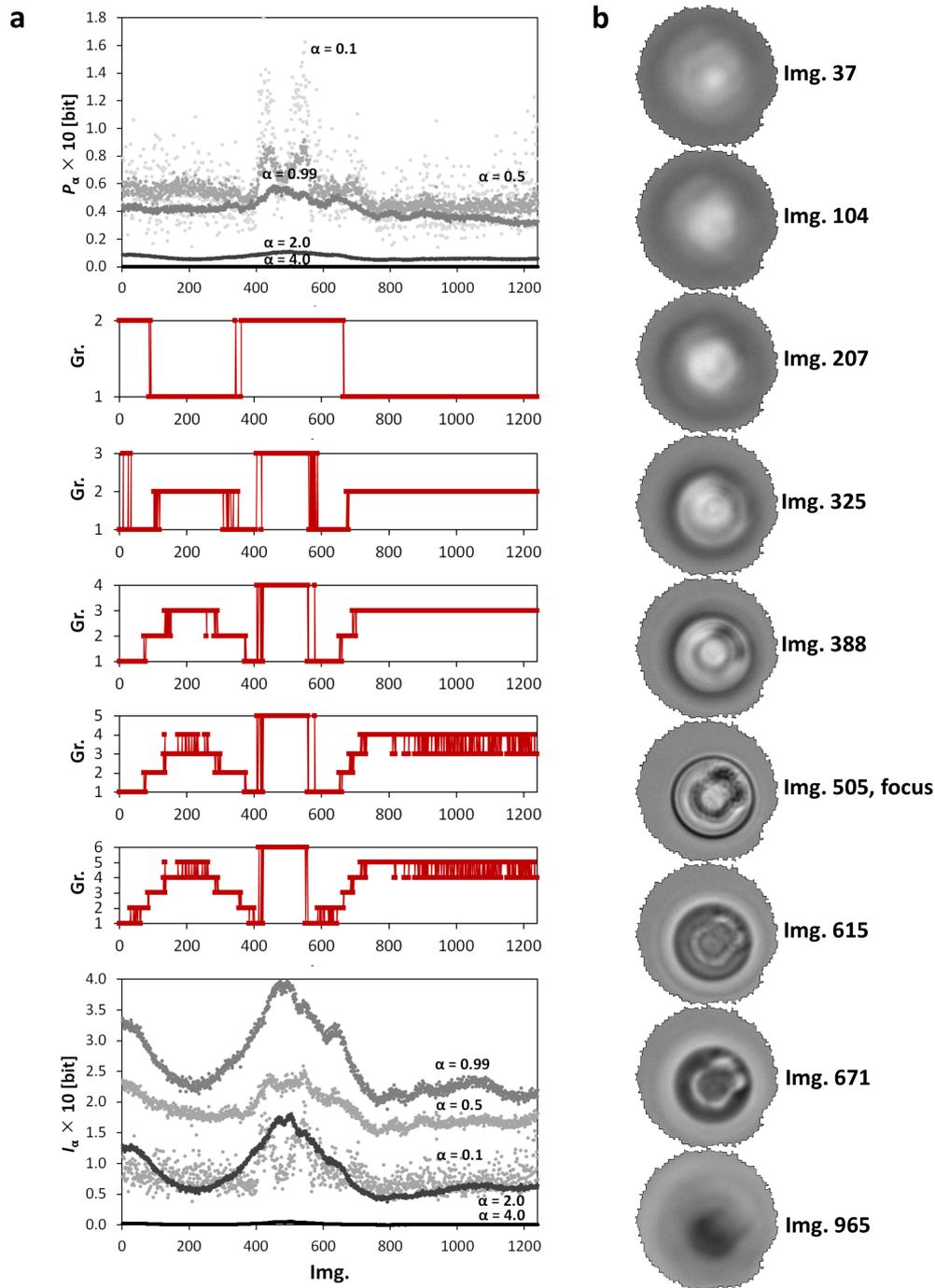


Figure 7. The results of clustering of a z-stack of grayscale microscopic images of a microring obtained using a diffraction mode. (a) the dependencies of (upper) the P_α and (lower) the I_α vs. order of the image in the z-stack for $\alpha = \{0.5;0.99;2.0;4.0\}$ and (middle) clustering (k-means, squared Euclidian distance, 2–6 groups) of the z-stack using connected spectra $[I_\alpha, P_\alpha]$ for $\alpha = \{0.1;0.3;0.5;0.7;0.99;1.3;1.5;1.7;2.0;2.5;3.0;3.5;4.0\}$; (b) the typical (middle) group’s images for clustering into 5 groups (in (a), middle). The original 12-bit images are visualized in 8 bits using the Least Information Loss conversion [23].

4. Materials and Methods

4.1. Processing of Typical Histograms

For the Cauchy, Lévy, Gauss, and Rayleigh distributions, dependences of the $\Omega_\alpha^{(l \rightarrow m)}$ on the number of elements in bins l and m were calculated for $\alpha = \{0.1, 0.3, 0.5, 0.7, 0.99, 1.3, 1.5, 1.7, 2.0, 2.5, 3.0, 3.5, 4.0\}$ using a `pdg_histograms.m` Matlab[®] 2014 script (Mathworks, Natick, MA, USA). The following probability density functions $f(x)$ were studied:

1. Lévy distribution:

$$f(x) = \text{round} \left[10^c \frac{\exp\left(-\frac{1}{2x}\right)}{\sqrt{2\pi x^3}} \right], \quad x \in \mathbb{N}, \quad \begin{cases} x \in [1, 256], & c \in \{5, 7\}, \\ x \in [1, 85], & c = 3, \end{cases} \quad (19)$$

2. Cauchy distribution:

$$f(x) = \text{round} \left[10^c \frac{1}{\pi(1+x^2)} \right], \quad x \in \mathbb{Z}, \quad \begin{cases} x \in [-127, 127], & c = 7, \\ x \in [-44, 44], & c = 3.5, \end{cases} \quad (20)$$

3. Gauss distribution:

$$f(x) = \text{round} \left[10^c \frac{\exp\left(-\frac{x^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}} \right], \quad x \in \mathbb{Z}, \quad \begin{cases} x \in [-4, 4], & c = 4, \quad \sigma = 1, \\ x \in [-29, 29], & c = 3, \quad \sigma = 10, \\ x \in [-36, 36], & c = 4, \quad \sigma = 10, \\ x \in [-64, 64], & c = 10, \quad \sigma = 10, \end{cases} \quad (21)$$

4. Rayleigh distribution:

$$f(x) = \text{round} \left[10^c \frac{x}{b^2} \exp\left(-\frac{x^2}{2b^2}\right) \right], \quad x \in \mathbb{N}, \quad x \in [1, 108], \quad c = 10, \quad b = 16. \quad (22)$$

In Figure 1, the Cauchy and Lévy distributions at $c = 7$ and the Gauss distribution at parameters $c = 10$ and $\sigma = 10$ are depicted.

4.2. Image Processing and Analysis

Image analysis based on calculation of the $\Omega_\alpha^{(l \rightarrow m)}$, I_α , and P_α is demonstrated on five standard grayscale multi-image series (Table 1). All images were processed using Whole Image mode in an Image Info Extractor Professional software (Institute of Complex Systems, FFPW, USB, Nové Hradý, Czech Republic). A pair of images 5000–5001 of a simulated Belousov–Zhabotinsky (BZ) reaction and a pair of images motion01.512–motion02.512 were recalculated for 40 values of $\alpha = \{0.1, 0.2, \dots, 0.9, 0.99, 1.1, 1.2, \dots, 4.0\}$. The rest of series were processed for 13 values of $\alpha = \{0.1, 0.3, 0.5, 0.7, 0.99, 1.3, 1.5, 1.7, 2.0, 2.5, 3.0, 3.5, 4.0\}$. The transformation at 13 α was followed by clustering of the matrices $[P_\alpha, I_\alpha]$ vs. `Img`. by k -means method (squared Euclidian distance metrics). Due to a high data variance in the BZ simulation, the clustering was preceded by the z-score standardization of the matrices over α . The resulted indices of clusters were reclassified to be consecutive (i.e., the first image of the series and the first image of the following group are classified into gr. 1 and 2, respectively, etc.).

5. Conclusions

In this paper, we derived novel variables from the Rényi entropy—a Point Divergence Gain $\Omega_\alpha^{(l \rightarrow m)}$, a Point Divergence Gain Entropy I_α , and a Point Divergence Gain Entropy Density P_α . We have

discussed their theoretical properties and made a brief comparison with the related quantity called Point Information Gain Γ_{α}^i [2]. Moreover, we have shown that the $\Omega_{\alpha}^{(l \rightarrow m)}$ and related quantities can find their applications in multidimensional data analysis, particularly in video processing. However, due to element-by-element computation, we can characterize time-spatial (4-D) changes much more sensitively than using, e.g., the previously derived Γ_{α}^i . The $\Omega_{\alpha}^{(l \rightarrow m)}$ can be considered as a microstate of the information changes in the space-time. However, the $\Omega_{\alpha}^{(l \rightarrow m)}$, I_{α} , and P_{α} show a property that is similar to the Γ_{α}^i and its relative macroscopic variables. Due to the derivation from the Rényi entropy, they are good descriptors of multifractality. Therefore, they can be utilized to characterize patterns in datasets and to classify the (sub)data into groups of similar properties. This has been successfully utilized in clustering of multi-image sets, image filtration, and image segmentation, namely in microscopic digital imaging.

Acknowledgments: This work was supported by the Ministry of Education, Youth and Sports of the Czech Republic—projects CENAKVA (No. CZ.1.05/2.1.00/01.0024), CENAKVA II (No. LO1205 under the NPU I program), the CENAKVA Centre Development (No. CZ.1.05/2.1.00/19.0380)—and from the European Regional Development Fund in frame of the project Kompetenzzentrum MechanoBiologie (ATCZ133) in the Interreg V-A Austria—Czech Republic programme. J.K. acknowledges the financial support from the Czech Science Foundation Grant No. 17-33812L and the Austrian Science Fund, Grant No. I 3073-N32.

Author Contributions: Renata Rychtáriková is the main author of the text and tested the algorithms; Jan Korbel is responsible for the theoretical part of the article; Petr Macháček is the developer of the Image Info Extractor Professional software; Dalibor Štys is the group leader who derived the point divergence gain. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rychtáriková, R. Clustering of multi-image sets using Rényi information entropy. In *Bioinformatics and Biomedical Engineering (IWBBIO 2016)*; Lecture Notes in Computer Science Series; Ortuño, F., Rojas, I., Eds.; Springer: Cham, Switzerland, 2016; pp. 527–536.
2. Rychtáriková, R.; Korbel, J.; Macháček, P.; Císař, P.; Urban, J.; Štys, D. Point information gain and multidimensional data analysis. *Entropy* **2016**, *18*, 372.
3. Jizba, P.; Kleinert, H.; Shefaat, M. Rényi's information transfer between financial time series. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 2971–2989.
4. Jizba, P.; Korbel, J. Multifractal diffusion entropy analysis: Optimal bin width of probability histograms. *Phys. A Stat. Mech. Appl.* **2014**, *413*, 438–458.
5. Jizba, P.; Korbel, J. Modeling Financial Time Series: Multifractal Cascades and Rényi Entropy. In *ISCS 2013: Interdisciplinary Symposium on Complex Systems; Emergence, Complexity and Computation Series*; Sanayei, A., Zelinka, I., Rössler, O., Eds.; Springer: Berlin/Heidelberg, Germany, 2014.
6. Rychtáriková, R.; Náhlík, T.; Smaha, R.; Urban, J.; Štys, D., Jr.; Císař, P.; Štys, D. Multifractality in imaging: Application of information entropy for observation of inner dynamics inside of an unlabeled living cell in bright-field microscopy. In *ISCS14: Interdisciplinary Symposium on Complex Systems*; Sanayei, A., Zelinka, I., Rössler, O.E., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; pp. 261–267.
7. Rychtáriková, R.; Náhlík, T.; Shi, K.; Malakhova, D.; Macháček, P.; Smaha, R.; Urban, J.; Štys, D. Super-resolved 3-D imaging of live cells' organelles from bright-field photon transmission micrographs. *Ultramicroscopy* **2017**, *179*, 1–14.
8. Chevalier, C.; Bect, J.; Ginsbourger, D.; Vazquez, E.; Picheny, V.; Richet, Y. Fast parallel kriging-based stepwise uncertainty reduction with application to the identification of an excursion set. *Technometrics* **2014**, *56*, 455–465.
9. Eidsvik, J.; Mukerji, T.; Bhattacharjya, D. *Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis*; Cambridge University Press: Cambridge, UK, 2015.
10. Helle, K.B.; Pebesma, E. Optimising sampling designs for the maximum coverage problem of plume detection. *Spat. Stat.* **2015**, *13*, 21–44.
11. Jizba, P.; Arimitsu, T. The world according to Rényi: Thermodynamics of multifractal systems. *Ann. Phys.* **2004**, *312*, 17–59.

12. Rényi, A. On measures of entropy and information. In Proceedings of the Fourth Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 20 June–30 July 1961; Volume 1, pp. 547–561.
13. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.
14. Csiszár, I. I-divergence geometry of probability distributions and minimization problems. *Ann. Probab.* **1975**, *3*, 146–158.
15. Harremoës, P. Interpretations of Rényi entropies and divergences. *Phys. A Stat. Mech. Appl.* **2006**, *365*, 5–62.
16. Van Erven, T.; Harremoës, P. Rényi divergence and Kullback-Leibler divergence. *J. Latex Class Files* **2007**, *6*, 1–24.
17. Van Erven, T.; Harremoës, P. Rényi divergence and majorization. In Proceedings of the IEEE International Symposium on Information Theory, Austin, TX, USA, 13–18 June 2010; pp. 1335–1339.
18. Shannon, C.E. A mathematical theory of communication. *ACM SIGMOBILE Mob. Comput. Commun.* **2001**, *5*, 3–55.
19. Volume 1: Textures. Available online: <http://sipi.usc.edu/database/database.php?volume=textures&image=61#top> (accessed on 31 January 2018).
20. Štys, D.; Jizba, P.; Zhyrova, A.; Rychtáriková, R.; Štys, K.M.; Náhlík, T. Multi-state stochastic hotchpotch model gives rise to the observed mesoscopic behaviour in the non-stirred Belousov-Zhabotinsky reaction. *arXiv* **2016**, arXiv:1602.03055. Available online: <https://arxiv.org/abs/1602.03055> (accessed on 31 January 2018).
21. Štys, D.; Náhlík, T.; Zhyrova, A.; Rychtáriková, R.; Papáček, Š.; Císař, P. Model of the Belousov-Zhabotinsky reaction. In Proceedings of the International Conference on High Performance Computing in Science and Engineering, Soláň, Czech Republic, 25–28 May 2015; Kozubek, T., Blaheta, R., Šístek, J., Eds.; Springer: Cham, Switzerland, 2016; pp. 171–185.
22. Štys, D.; Štys, K.M.; Zhyrova, A.; Rychtáriková, R. Optimal noise in the hodgepodge machine simulation of the Belousov-Zhabotinsky reaction. *arXiv* **2016**, arXiv:1606.04363. Available online: <https://arxiv.org/abs/1606.04363> (accessed on 31 January 2018).
23. Štys, D.; Náhlík, T.; Macháček, P.; Rychtáriková, R.; Saberioon, M. Least Information Loss (LIL) conversion of digital images and lessons learned for scientific image inspection. In Proceedings of the International Conference on Bioinformatics and Biomedical Engineering, Granada, Spain, 20–22 April 2016; Ortuño, F., Rojas, I., Eds.; Springer: Cham, Switzerland, 2016; pp. 527–536.
24. Braat, J.J.M.; Dirksen, P.; van Haver, S.; Janssen, A.J.E.M. Extended Nijboer-Zernike (ENZ) Analysis & Aberration Retrieval. Available online: <http://www.nijboerzernike.nl> (accessed on 12 December 2017).
25. Rychtáriková, R.; Steiner, G.; Kramer, G.; Fischer, M.B.; Štys, D. New insights into information provided by light microscopy: Application to fluorescently labelled tissue section. *arXiv* **2017**, arXiv:1709.03894. Available online: <https://arxiv.org/abs/1709.03894> (accessed on 31 January 2018).
26. Jolliffe, I.T. *Principal Component Analysis*, 2nd ed.; Springer Series in Statistics; Springer: New York, NY, USA, 2002.
27. Arthur, D.; Vassilvitskii, S. k-means++: The advantages of careful seeding. In Proceedings of the 18th ACM-SIAM, Philadelphia, PA, USA, 7–9 January 2007; pp. 1027–1035.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).