Article

Complexity Analysis of Escher's Art

António M. Lopes [1],*,† and J. A. Tenreiro Machado [2],†

[1] UISPA–LAETA/INEGI, Faculty of Engineering, University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[2] Institute of Engineering, Polytechnic of Porto, Department of Electrical Engineering, R. Dr. António Bernardino de Almeida, 431, 4249-015 Porto, Portugal; jtm@isep.ipp.pt
* Correspondence: aml@fe.up.pt; Tel.: +351-22-041-3486
† These authors contributed equally to this work.

check for updates

Abstract: Art is the output of a complex system based on the human spirit and driven by several inputs that embed social, cultural, economic and technological aspects of a given epoch. A solid quantitative analysis of art poses considerable difficulties and reaching assertive conclusions is a formidable challenge. In this paper, we adopt complexity indices, dimensionality-reduction and visualization techniques for studying the evolution of Escher's art. Grayscale versions of 457 artworks are analyzed by means of complexity indices and represented using the multidimensional scaling technique. The results are correlated with the distinct periods of Escher's artistic production. The time evolution of the complexity and the emergent patterns demonstrate the effectiveness of the approach for a quantitative characterization of art.

Keywords: complexity; information theory; multidimensional scaling; art

1. Introduction

During human history, artists conceived harmonies of objects and forms in their works [1]. Artworks are manifestations of the artists' creativity, reflecting their thoughts and culture [2]. Often we have a glimpse of a kind of mathematical exercise embedded in the artworks [3–9]. It is well known the symbiosis of art and science produced by Leonardo da Vinci [10], or the magic complexity of the music composed by Bach [11]. The fascination of many artists/scientists for art/mathematics is ubiquitous in human history [12–14]. Nowadays, we can take advantage of computational algorithms to stimulate synergies between art and science.

Art is a token of a complex system produced by mankind and influenced by a plethora of social, cultural, economic and technological inputs that interact in time and space. The study of artworks may help interpreting the world and the human mind. However, an assertive analysis of art poses conceptual and practical difficulties and reaching quantitative conclusions represents a huge challenge.

The tools to analyze complex systems have been successfully adopted in different areas, including economics, life and social sciences [15], with the objective of finding fundamental principles and universalities that govern the systems' dynamics [16,17]. Complexity is helpful to quantitatively describe nonlinear systems and to detect changes in their dynamics.

Several complexity indices have been adopted to analyze art, namely entropy [18,19], Kolmogorov complexity [20,21], fractal dimension [22,23], and others [24,25]. Such indices are not independent, but they capture different aspects of the system state, complementing each other and leading to a deeper assessment of the subject under study [26]. The quantitative analysis of art dates back to 1933, when Birkhoff proposed an aesthetic measure as the ratio between order (i.e., number of regularities) and complexity (i.e., number of elements) of an image. Nevertheless, only recently

quantitative techniques were applied, impelled by the availability of digital data and the development of computational tools. Taylor et al. [22] verified that Jackson Pollock's (1912–1956) dripped patterns are fractals and that the fractal dimension of the paintings increased over the course of his artistic career. Dodgson [18] adopted the concepts of entropy and correlation to describe Bridget Riley's (1961–2012) stripe paintings. Cucker [3] suggested that geometry is an important source of rules for artistic creation. Wallraven et al. [27] used multidimensional scaling (MDS) and clustering techniques to categorize paintings. Kim et al. [28] analyzed a large database of images, finding that the color-usage distribution is remarkably different among historical periods of western paintings. Lee et al. [29] examined almost 180 thousand paintings, focusing on the evolution of the color contrast. Among other findings, they observed a sudden increase in the diversity of color contrast after 1850. Machado and Lopes [20] studied paintings from the viewpoint of information theory and fractional calculus. Sigaki et al. [19] addressed the local order patterns of almost 140 thousand artwork images using complexity indices and observed a clear and robust time evolution.

Escher (1898–1972) is one of the most celebrated modern graphic artists [30,31]. He is known for his works in woodcuts, lithographs and mezzotints representing fantastic, unusual and impossible objects, with various perspectives, generating optical illusions. Escher is considered a mathematical artist, especially geometric, influenced by his relationships with mathematicians and by its own interests and abilities in mathematics. His artworks explore aspects such as infinity, perspective, symmetry, reflection, hyperbolic geometry, truncated and stellated polyhedra, and tessellations [32,33]. Escher made about 450 lithographs, woodcuts and wood engravings, and over 2000 drawings and sketches. Despite some open discussion about his 54-year long artistic career, Escher's work is classified in several periods tightly related to the places where he lived in. We observe five distinct periods [34]:

- $\mathcal{P}_1$—The early work period (1916–1922)—Escher lived in Arnhem and Haarlem, in the Netherlands. Most pieces of this period consist of woodcuts and linocuts produced when he was a student. The artwork is varied in theme, going from portraits to drawings that hint at cubism. Escher developed the linocut printmaking technique and learned to represent figures in black and white. One aspect of this process was to learn to 'think backwards', since images carved into woodblocks for printing must be carved backwards, as though seen in a mirror (e.g., 'Skull', c. 1920).

- $\mathcal{P}_2$—The Italian period (1922–1937)—Escher lived in Italy from 1924 to 1935. During those years he traveled across the country and created a large portfolio of lithographs and wood engravings based on drawings of Italian buildings, landscapes and seascapes (e.g., 'Castrovalva', c. 1930). In this period Escher was focused in representing the reality and his portfolio includes also plants, animals and portraits. An early attempt to make some different drawings, with the interpenetration of distinct worlds, came just in the final phase of this period (e.g., 'Candle Mirror', c. 1934).

- $\mathcal{P}_3$—The metamorphosis period (1937–1945)—Escher left Italy in 1935, lived in Switzerland and Belgium between 1935 and 1941, and went back to the Netherlands in 1941. In 1936 Escher revisited Spain and traveled to Alhambra and Cordoba. This trip inspired him to the subject of tessellations with great impact in his art. During this period he focused on representing the world as how it could be instead as how it really was. The artworks include cycles and the transformation of 3-dim into 2-dim forms. The symmetry and the perfect fit of shapes are also characteristic marks of this period (e.g., 'Day and Night', c. 1938).

- $\mathcal{P}_4$—The subordinated to perspective period (1946–1956)—in this period Escher worked with engravings, using unusual and multiple viewpoints, vanishing points and perspectives. Some works suggest the infinity of space through multiple vanishing points and bundles of straight lines. Escher stressed the sense of depth through the use of colors, progressively blurred throughout the pictures and creating the idea of an aerial perspective (e.g., 'Depth', c. 1955). He also demonstrated interest in geometric solids due to his studies in mineralogy and crystallography.

- $\mathcal{P}_5$—The approximation to infinity period (1956–1970)—in this period Escher made several engravings that have as central theme the infinity, where he explored ideas from hyperbolic geometry (e.g., 'Circle Limit III', c. 1959). This period is also characterized by the production of impossible figures, with perspectives, reflections, conflicts of dimension, illusion, and the shape of space (e.g., 'Art Gallery 1956', c. 1956, 'Waterfall', c. 1961).

In this paper we adopt complexity indices, dimensionality reduction and visualization techniques for studying the evolution of Escher's art. In the first phase, 457 artworks, produced between 1916 and 1969, were converted into digital format, discretized and represented in grayscale. The images were then processed by computing six distinct complexity indices, whose evolution was correlated with the periods of the artist's career. In a second phase, a MDS algorithm is adopted for visualizing complexity. The MDS was fed with dissimilarity information calculated with two different distance measures. The MDS maps were interpreted under the light of the emerging clusters and correlated with the periods of Escher's art. It should be noted that other indices can be used for quantifying complexity and different techniques can be adopted for dimensionality reduction, clustering and visualization. We can mention, for example, the use of time–frequency signal processing and hierarchical clustering for studying tidal data [35], the Lempel–Ziv complexity, sample entropy, signal harmonics power ratio, and fractal dimension for analyzing temperature time series [26], and information theory, fractional calculus and hierarchical clustering for studying art [20].

Considering these ideas, Section 2 introduces the mathematical background, emphasizing the concept of complexity and the MDS technique. Section 3 analyzes the evolution of Escher's art in the perspective of six complexity indices. Section 4 uses MDS for dimensionality reduction and data visualization, and interprets the generated maps in the perspective of the periods of the artist's career. Finally, Section 5 presents the conclusions.

2. Mathematical Background

2.1. Classic Information Indices

The information theory proposed by Shannon [36] was recently adopted in the study of complex systems [37,38].

Let us consider a discrete 1-dim random variable X with sample space $\{x_1, \ldots, x_i, \ldots, x_N\}$ and probability distribution $P(X)$. An event, x_i, with probability of occurrence $P(x_i)$ has the information content:

$$h\left[P\left(x_i\right)\right] = -\log P\left(x_i\right).$$

(1)

The Shannon entropy is the arithmetic average, or expected value, of $h\left[P\left(x_i\right)\right]$:

$$H(X) = E\left[-\log P\left(x_i\right)\right] = \sum_{i=1}^{N} -P\left(x_i\right)\log P\left(x_i\right),$$

(2)

where the operator $E\left(\cdot\right)$ represents the expected value.

The joint entropy of a two-dimensional discrete random variable (X, Y) with sample spaces $\{x_1, \ldots, x_i, \ldots, x_N\}$ and $\{y_1, \ldots, y_j, \ldots, y_M\}$, and joint probability distribution $P(X, Y)$ is [39]:

$$S(X, Y) = \sum_{i=1}^{N} \sum_{j=1}^{M} -P\left(x_i, y_j\right)\log P\left(x_i, y_j\right).$$

(3)

The mutual information of (X, Y), with marginal probability distributions $P(X)$ and $P(Y)$, respectively, is given by [40,41]:

$$I(X;Y) = \sum_{i=1}^{N} \sum_{j=1}^{M} P\left(x_i, y_j\right) \log \frac{P\left(x_i, y_j\right)}{P(x_i)P(y_j)} \tag{4}$$

and assesses the information shared by (X, Y). When X and Y are independent, there is no shared information between them, and the mutual information is $I(X;Y) = 0$.

In the follow-up, the complexity indices $S(X, Y)$ and $I(X;Y)$ will be interpreted as state variables and the locus (S, I) will be called the entropy-mutual information plane.

The Jensen–Shannon divergence measures the similarity between two probability distributions $P(X)$ and $P(Y)$ and is defined as [41]:

$$JSD\left[P(X) \parallel P(Y)\right] = \frac{1}{2}\left[\sum_{i=1}^{N} P(x_i) \log P(x_i) + \sum_{i=1}^{N} P(y_i) \log P(y_i)\right] - \sum_{i=1}^{N} P(z_i) \log P(z_i), \tag{5}$$

where X and Y are random variables with sample spaces $\{x_1, \ldots, x_i, \ldots, x_N\}$ and $\{y_1, \ldots, y_i, \ldots, y_N\}$, and $Z = \frac{1}{2}(X + Y)$.

2.2. Permutation Entropy and Statistical Complexity

The permutation entropy (PE) was originally proposed as a robust index to assess the complexity of time series [42].

For a time series, $\{x_k : k = 1, \ldots, W\}$, $x_k \in \mathbb{R}$, we define two parameters: the embedding dimension, $d \geq 2$, $d \in \mathbb{N}$, and the embedding delay, $\tau \in \mathbb{N}$, that represent the length of the time series partitioning sequences, and the separation time between their elements, respectively. Let us denote by $\Psi = \{\Pi_1, \ldots, \Pi_{d!}\}$ the set of all possible permutations of the ordinals $\{1, \ldots, d\}$, and by $[\mathcal{I}]$ the Iverson bracket [43], such that $\mathcal{I} = \begin{cases} 1, & \text{if } \mathcal{I} \text{ is true} \\ 0, & \text{if } \mathcal{I} \text{ is false} \end{cases}$. The procedure for calculating PE is as follows:

1. For each $k = 1, \ldots, K$, with $K = W - (d-1)\tau$,

    1.1. Compose the sequence $\{x_k, x_{k+\tau}, \ldots, x_{k+(d-1)\tau}\}$;

    1.2. Construct the $2 \times d$ dimensional array $\begin{bmatrix} x_k & x_{k+\tau} & \cdots & x_{k+(d-1)\tau} \\ 1 & 2 & \cdots & d \end{bmatrix}$;

    1.3. Sort the array by increasing order of the elements in the first row;

    1.4. Denote by π_k the sequence of numbers in the second row of the sorted array;

2. Compute the probability distribution $P = [p_1, p_2, \ldots, p_{d!}]$, where $p_i = \frac{1}{K} \sum_{k=1}^{K} [\pi_k = \Pi_i]$, $i = 1, \ldots, d!$;

3. Calculate $PE(P) = \frac{1}{\log d!} \sum_{i=1}^{d!} -p_i \log p_i$.

The value of PE lies in the interval $0 \leq PE \leq 1$. The lower value $PE = 0$ indicates that the time series is regular or predictable, while the upper value $PE = 1$ corresponds to a random time series. The embedding dimension must be chosen such that $W \gg d!$ in order to obtain reliable values of PE. For practical purposes the values $d \in \{3, \ldots, 7\}$ and $\tau = 1$ are suggested [42].

For two-dimensional data, as is the case of images, the generalization of the procedure is straightforward. Let us consider a greyscale image represented by the $N_x \times N_y$ matrix A. We define four parameters: two for the embedding dimensions, $d_x, d_y \geq 2$, with $d_x, d_y \in \mathbb{N}$, and two for the embedding delays, $\tau_x, \tau_y \in \mathbb{N}$. The symbolic sequences for calculating the probabilities, $P = [p_1, p_2, \ldots, p_{(d_x d_y)!}]$, are now obtained from the spatial information in overlapping $d_x \times d_y$ dimensional submatrices. We have a total of $(d_x d_y)!$ permutations of the ordinals $\{1, \ldots, d_x d_y\}$ and we choose d_x and d_y such

that $N_x N_y \ll (d_x d_y)!$ for having reliable values for PE (for details see [44,45]). For images, the PE is a measure of 'randomness' in the layout of the pixels. If the pixels appear in a random (in the same) order, then $PE \to 1$ ($PE \to 0$).

Another complexity measure is the statistical complexity (C) proposed in [46,47]:

$$C(P) = \frac{1}{\kappa} \cdot JSD\,[P \parallel U] \cdot PE(P), \tag{6}$$

where $JSD\,[P \parallel U]$ is the Jensen–Shannon divergence between $P = [p_1, p_2, \ldots, p_{(d_x d_y)!}]$ and the uniform distribution $U = [u_1, u_2, \ldots, u_{(d_x d_y)!}] = \frac{1}{(d_x d_y)!} \cdot [1, 1, \ldots, 1]$, whereas

$$\kappa = \max_P \{JSD\,[P \parallel U]\} = -\frac{1}{2}\left[\frac{(d_x d_y)!+1}{(d_x d_y)!} \log\left[(d_x d_y)! + 1\right] + \log\left[(d_x d_y)!\right] - 2\log\left[2(d_x d_y)!\right]\right]. \tag{7}$$

is a normalization constant.

Often the locus (PE, C) is adopted for characterizing data, being called as the complexity-entropy plane [46,48].

2.3. Kolmogorov Complexity-Based Indices

The Kolmogorov complexity, $K(A)$, of an object A provides a measure of information independently of any probabilistic assumptions about the data sequences in A. The complexity $K(A)$ is defined as the size of the shortest program that, given an empty object at its input, computes A in an universal computer and then stops [49–51]. The exact value of $K(A)$ is not computable [49–54]. Therefore, approximations based on the Lempel–Ziv [55], linguistic [56] and compression [57] algorithms are used to obtain the upper bounds $K(A)$.

Lossless compression algorithms approximate $K(A)$ by the size of the compressed object, $K(A) \approx \text{size}\{\Phi(A)\}$, where $\Phi(\cdot)$ denotes compression [50,51]. However, for obtaining a good approximation, the compressor has to be 'normal'. This means that, given the object A and the concatenation of A with itself, AA, the compressor has to generate compressed objects such that $\text{size}[\Phi(A)] \approx \text{size}[\Phi(AA)]$ [50,51]. For having a complexity index independent of $\text{size}[A]$ we use the complexity ratio (CR):

$$CR = \frac{\text{size}[\Phi(A)]}{\text{size}[A]}. \tag{8}$$

The information distance between two objects $\{A_1, A_2\}$ can be computed using the conditional Kolmogorov complexity, $K(A_1|A_2)$. This corresponds to the size of the shortest program to compute A_1, given that A_2 is provided [58,59]. Therefore, when we have almost similar objects A_1 and A_2, the task is less complex and the size of the program is smaller. The inequality $K(A_1|A_2) \leq K(A_1)$ always holds and the normalized information distance, NID, is formulated as an universal metric [58]:

$$NID(A_1, A_2) = \frac{\max\{K(A_1|A_2), K(A_2|A_1)\}}{\max\{K(A_1), K(A_2)\}}. \tag{9}$$

The NID is a distance and thus it satisfies the conditions:

1. $NID(A_1, A_2) \geq 0$; moreover, we have (i) $NID(A_1, A_2) = 0$, if and only if $A_1 = A_2$; and (ii) $NID(A_1, A_2) = 1$, if and only if A_2 is an empty object (non-negativity);
2. $NID(A_1, A_2) = NID(A_2, A_1)$ (symmetry);
3. $NID(A_1, A_2) \leq NID(A_1, A_3) + NID(A_3, A_2)$ (triangle inequality).

The NID is based on $K(\cdot)$ and it is not computable [58]. To surpass this limitation, the normalized compression distance:

$$NCD(A_1, A_2) = \frac{\text{size}[\Phi(A_1 A_2)] - \min\{\text{size}[\Phi(A_1)], \text{size}[\Phi(A_2)]\}}{\max\{\text{size}[\Phi(A_1)], \text{size}[\Phi(A_2)]\}} \tag{10}$$

was proposed for approximating the *NID* [60,61]. The *NCD* is nonnegative, and we have $0 < NCD(A_1, A_2) < 1 + \epsilon$, where $\epsilon > 0$ denotes the error introduced by the compressor [61,62].

Stemming from (10), we propose the complexity index $NCD_R(A, R)$, where R represents a reference object of the same size of A. We tested for R pixels with random noise and all-identical shade of gray. The results showed that $NCD_R(A, R)$ is quite insensitive to the type of R, since it consists of a normalized distance. The locus (CR, NCD_R) will be designated by compression distance-ratio plane.

2.4. Multidimensional Scaling

The MDS is a technique for dimensionality reduction, clustering and computational visualization of multidimensional data [63–69]. Given L objects x_i, $i = 1, \ldots, L$, in a r-dim space and a measure of dissimilarity between the *i*th and *j*th objects, $\delta_{ij}(x_i, x_j)$, the procedure starts by calculating a $L \times L$ symmetric matrix, $\Delta = [\delta_{ij}]$ of object-to-object dissimilarities. The matrix Δ gives the input information to the MDS numerical algorithm. The MDS represents objects by means of points located in a q-dim space ($q < r$) at distances θ_{ij}, and iterates multiple configurations in order to maximize a fitness function and to achieve a map of points that approximates the original ones. By other words, the MDS calculates the matrix of distances $\Theta = [\theta_{ij}]$ that try to mimic $\Delta = [\delta_{ij}]$. A fitness function widely used is the raw stress:

$$\mathcal{R} = \left[\theta_{ij} - f(\delta_{ij}) \right]^2 , \tag{11}$$

where $f(\cdot)$ denotes some linear or non-linear transformation.

The MDS interpretation follows the patterns of points obtained in the MDS locus. Two similar (dissimilar) objects are shown as two points that are close to (far from) each other. Thus, the MDS interpretation does not follow neither the coordinates of the points, nor the shape of the clusters. In fact, we can translate, rotate and magnify the map, since the object-to-object distances are identical. Moreover, the MDS axes have neither units nor special physical meaning.

The MDS quality can be quantified by means of the Shepard and stress plots. The Shepard diagram compares θ_{ij} and δ_{ij}, for a particular value of q. A narrow scattering of the points represents a good fit between θ_{ij} and δ_{ij}. The stress diagram represents the locus of $\mathcal{R}$ versus q. Usually users adopt $q = 2$ or $q = 3$, because such values allow a direct visualization and establish a compromise between achieving low values of $\mathcal{R}$ and q.

3. Complexity of Escher's Art

This Section addresses the evolution of Escher's art through time in the perspective of six complexity indices. Section 3.1 characterizes the data set and the conversion scheme of the artwork into digital format. Section 3.2 analyzes the time evolution of the complexity indices. Section 3.3 discusses the complexity of Escher's art under the light of the loci (S, I), (PE, C) and (CR, NCD_R).

3.1. Data Description

The study involved 457 artworks created by Escher between the years 1916 and 1969. The digital images were obtained from the *visual arts encyclopedia*, at the website www.wikiart.org, on 5 March 2019. WikiArt is one of the largest visual arts databases available online for free. Each image file, stored in JPEG format, is read into a $N_x \times N_y \times N_z$ dimensional matrix, A. For $N_z = 3$, A represents color images and its elements are 8-bit integers in the range between 0 and 255, corresponding to red, green and blue (RGB) intensities. The color images are converted to grayscale and represented by two-dimensional matrices ($N_z = 1$), where the elements of A denote values between black and white (Gr), generated by the ITU-R BT.601 RGB to gray conversion scheme, $Gr = 0.2989R + 0.5870G + 0.1140B$. This pre-processing yields some loss of information in the originals, not only in terms of color and texture, but also about the three-dimensional textural surface of the painting. However, since most Escher paintings are black and white or with shades of gray, the procedure, necessary

for reducing the volume of information and convert all works into an uniform format, does not compromise significantly the analysis. For example, Figure 1 depicts the 'Another World II' (c. 1947), and its corresponding grayscale image.
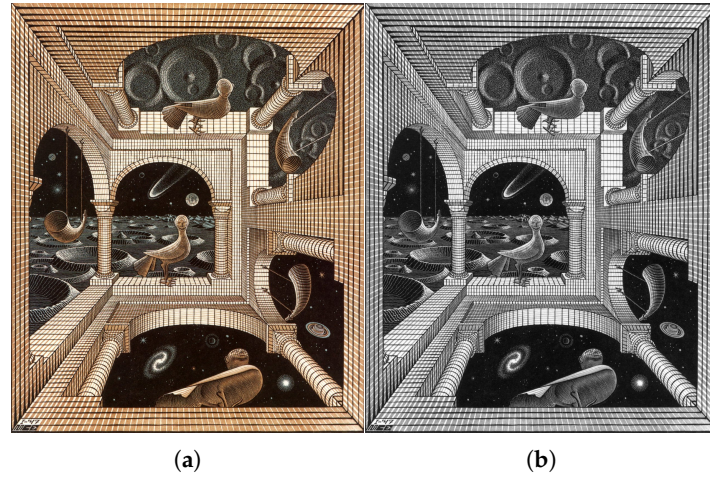

(**a**)　　　　　　　　　　　　　(**b**)

Figure 1. The artwork 'Another World II' (c. 1947): (**a**) color image; (**b**) grayscale image.

3.2. Time Evolution of the Complexity Indices

Here the artworks are characterized by means of the complexity indices in the set $\{S, I, PE, C, CR, NCD_R\}$. Their time evolution is then correlated with the Escher's artistic periods $\mathcal{P}_i$, $i = 1, \dots, 5$.

For calculating S and I the probabilities $P(x_i, y_j)$ are obtained from the matrices $A = [a_{ij}]$, $i = 1, 2, \dots, N_x$, $j = 1, 2, \dots, N_y$, through the proportion $P(x_i, y_j) = \frac{a_{ij}}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} a_{ij}}$, where x_i and y_i are the i-th row and j-th column. Therefore $P(x_i, y_i)$ represents a normalized version of an image A in the perspective of probability. Figure 2 presents an illustrative example, showing $P(x_i, y_j)$ for the grayscale version of the artwork 'Another World II'.
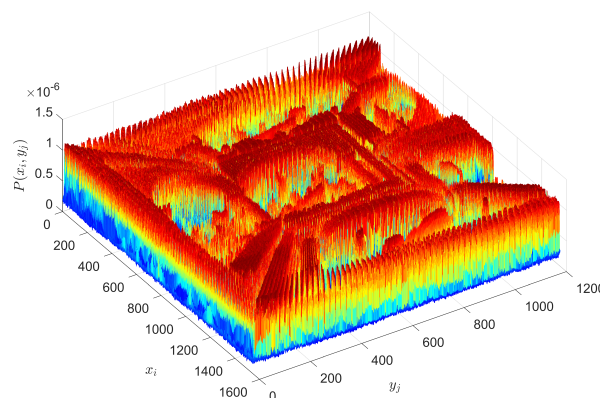

Figure 2. The two-dimensional histogram of the grayscale version of the 'Another World II' (c. 1947) by Escher.

For determining PE and C, we use the parameters $d_x = d_y = 2$ and $\tau_x = \tau_y = 1$, which were adjusted by numeric simulations. For computing CR and NCD_R we adopt the Windows implementation of the gzip compressor, version 1.3.12 (built upon the Lempel-Ziv coding algorithm LZ77), while the reference for the NCD_R is a white image.

Figure 3 depicts the *S* and *I* versus time, using box plots, for the period 1916–1969. In each box, the central mark indicates the median, the bottom and top edges correspond to the 25 and 75 percentiles, respectively, the whiskers span between the extreme data points (not considering the outliers), and the outliers are depicted using the marker '+'. The complexity indices were calculated considering artworks within five-year windows centered at the time stamp. Numerical experiments showed that this value was a good compromise for obtaining a readable and detailed graphical representation. Lower values increase the detail, but blur the charts, while higher values tend to filter too much the data and details are lost. We note that *S* has higher variations than *I*, discriminating better between periods. Indeed, we can find a relationship between the time evolution of *S* and *I* and the different periods of the Escher's artistic career, even knowing that this division is neither rigidly defined, nor absolutely consensual [34].
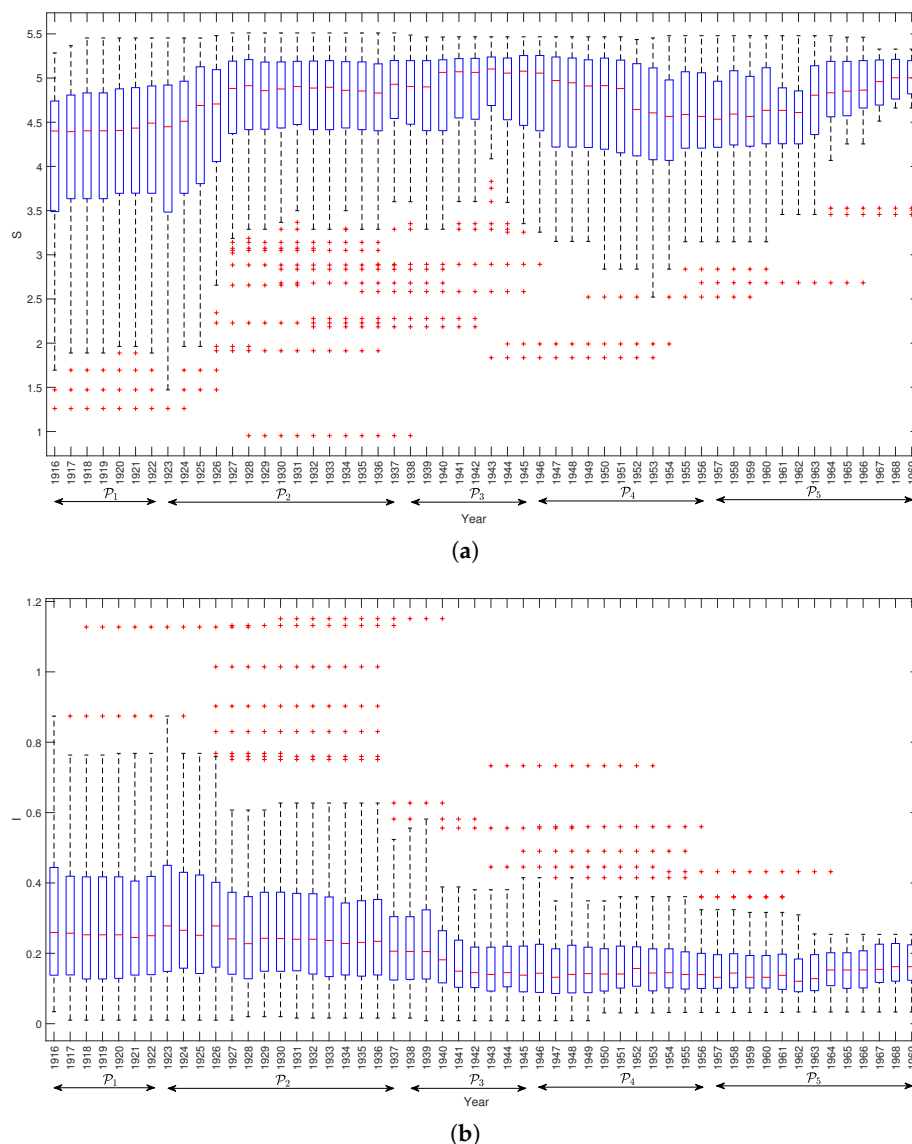

Figure 3. Two complexity indices versus time for the period 1916–1969: (**a**) *S*; (**b**) *I*.

We can mention five emerging periods:

- $\mathcal{P}_1$ (1916–1922) the joint entropy *S* and the mutual information *I* stay approximately constant;
- $\mathcal{P}_2$ (1922–1937) the *S* takes a leap up and then stabilizes, while *I* reveals a slight increase, with some oscillation, and then stabilizes;

- $\mathcal{P}_3$ (1937–1945) the value of S increases and I decreases;
- $\mathcal{P}_4$ (1946–1956) the S decreases considerable and reaches a new local minimum, while the value of I remains approximately constant;
- $\mathcal{P}_5$ (1956–1970) the value of S reveals an increasing trend, while I changes only a little bit, revealing a slight increase just at the end of the period.

For the other complexity indices we reach similar conclusions and, therefore, the corresponding box plots are omitted.

In synthesis, the complexity indices unveil direct correspondence with Escher's artistic periods. However, we note in the box plots some dispersion and outliers that may rise interesting questions. For example, what is it about the outliers that make the complexity indices to exhibit fluctuation? Should the outliers be interpreted as anomalies? What is it about the quality of the original images? Since they are from a public database, is the quality reliable? Does it have influence in the results?

3.3. Loci of the Complexity Indices

We analyzed the loci (S, I), (PE, C) and (CR, NCD_R) to assess the complexity of Escher's art. The resulting curves are depicted in Figure 4, where the labels denote the two last digits of the years (form 1916 to 1969) and the points correspond to the medians of the complexity indices of artworks within year-centered five-year length windows. This windowing procedure hides some details, but gives a clear picture of the global locus.

Analyzing the locus (S, I) (Figure 4a), we verify that during $\mathcal{P}_1$ the complexity has a very limited evolution, for $\mathcal{P}_2$ it has a large excursion, for $\mathcal{P}_3$ it changes direction, for $\mathcal{P}_4$ it has another large excursion and, finally for $\mathcal{P}_5$ it evolves almost in opposite direction with respect to the former period. It is also interesting to see that between two consecutive periods P_i and P_{i+1}, $(i = 2, 3, 4)$ we always has a tangle revealing the artist search for the new direction of work. For the locus (PE, C) (Figure 4b), we verified that the complexity indices were strongly correlated and, therefore, the five periods of Escher's career appear somewhat overlapped. Therefore, the exercise of discriminating $\mathcal{P}_i$, $i = 1, \ldots, 5$, had some limitations. Moreover, we did not find an obvious presence of the inter-period tangles observed in Figure 4a. For (CR, NCD_R) (Figure 4c), we observed again a good match between the artistic periods of Escher and the locus. Furthermore, we noted also the presence of the inter-period tangles.
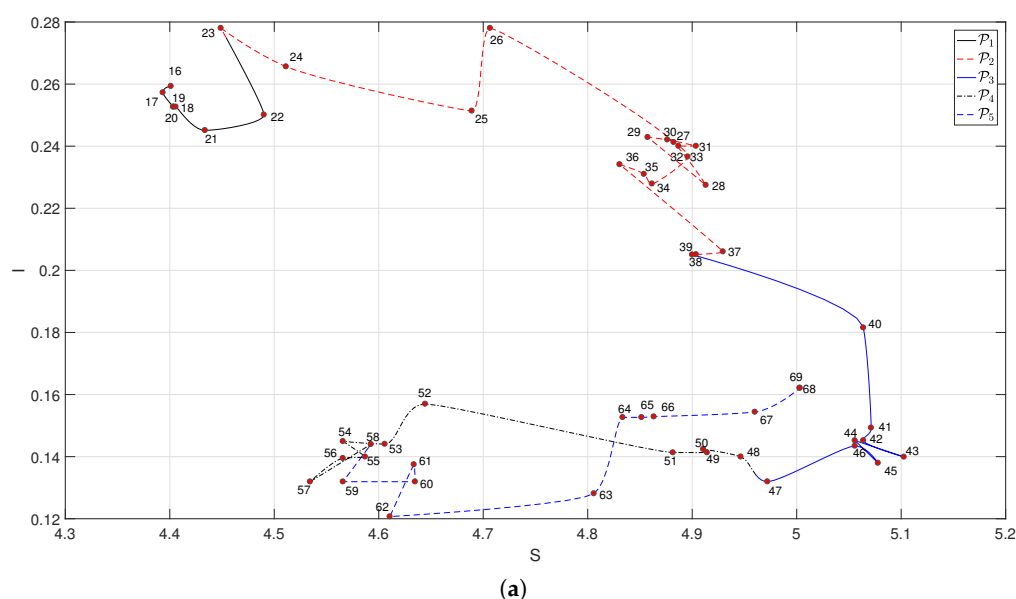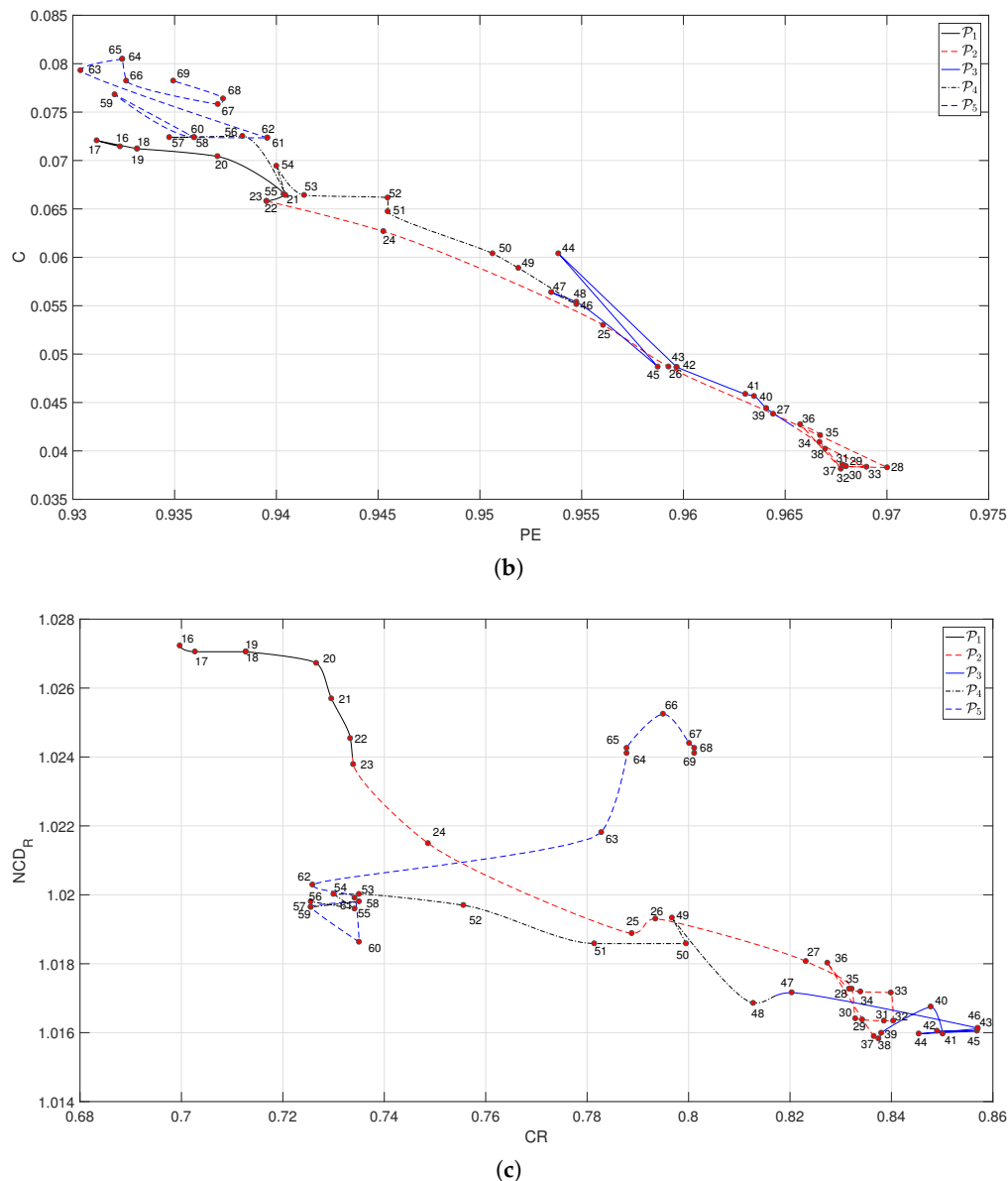

(a)

Figure 4. *Cont.*

(b)

(c)

Figure 4. The three complexity loci for the period 1916–1969: (**a**) (S, I); (**b**) (PE, C); (**c**) (CR, NCD_R).

4. MDS Visualization of Complexity

In this Section we use the MDS technique for visualizing complexity. In Section 4.1 the input for the MDS is the dissimilarity information between artworks computed with the NCD (10). In Section 4.2 the MDS is fed with the dissimilarity information between the six complexity indices $\{S, I, PE, C, CR, NCD_R\}$ calculated with the Euclidean distance.

4.1. MDS Visualization of Complexity Based on the NCD

We compute the 457×457 dimensional matrix $\Delta_{NCD} = [NCD(A_i, A_j)]$, where $NCD(A_i, A_j)$ is given by (10) and denotes the dissimilarity between the artworks A_i and A_j, $i, j = 1, \ldots, 457$. The matrix Δ_{NCD} is used as the input to the MDS numerical scheme. Since the MDS outputs a large number of points, we post-process the results by calculating the medians of the (x, y, z) coordinates that correspond to artworks within year-centered five-year length windows. Figure 5 depicts the resulting 54-point MDS three-dimensional map obtained with Δ_{NCD} for the period 1916–1969. We verify again the presence of the five periods $\mathcal{P}_1$–$\mathcal{P}_5$ and the inter-period tangle structures.
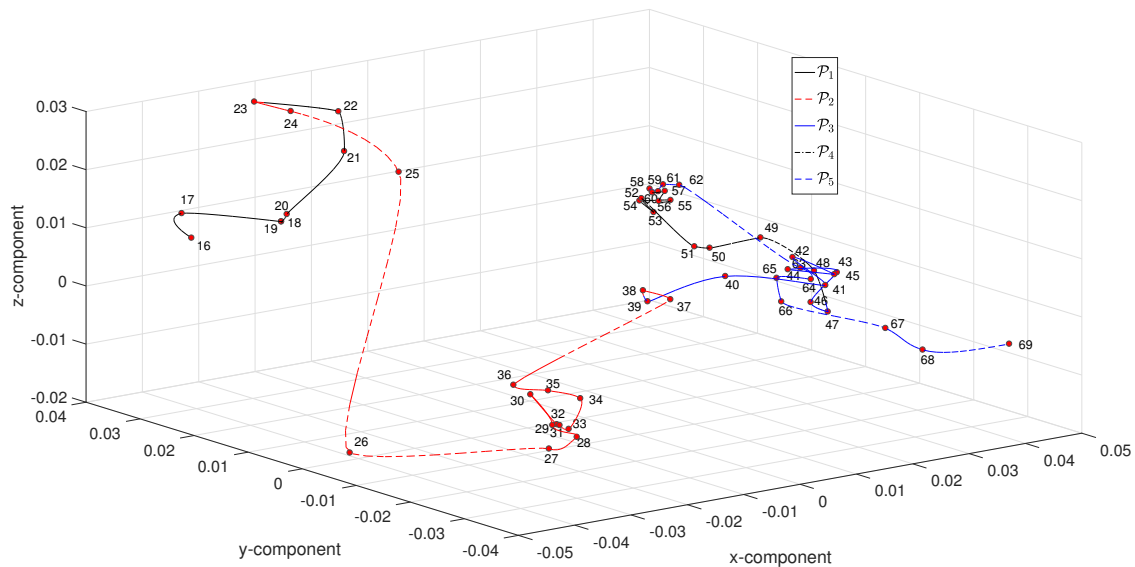
Figure 5. The multidimensional scaling (MDS) three-dimensional map obtained with Δ_{NCD} for the period 1916–1969.

Figure 6a,b show the corresponding MDS assessment charts. Since the Shepard diagram reveals a small scatter around the 45 degree line, we conclude that there exists a good fit between the original and the reproduced distances. The stress plot shows that the three-dimensional map ($n = 3$) is a good representation of the locus of points, since $n = 3$ corresponds to the elbow of the curve. Therefore, three-dimensional maps represent a good compromise between accuracy and readability.
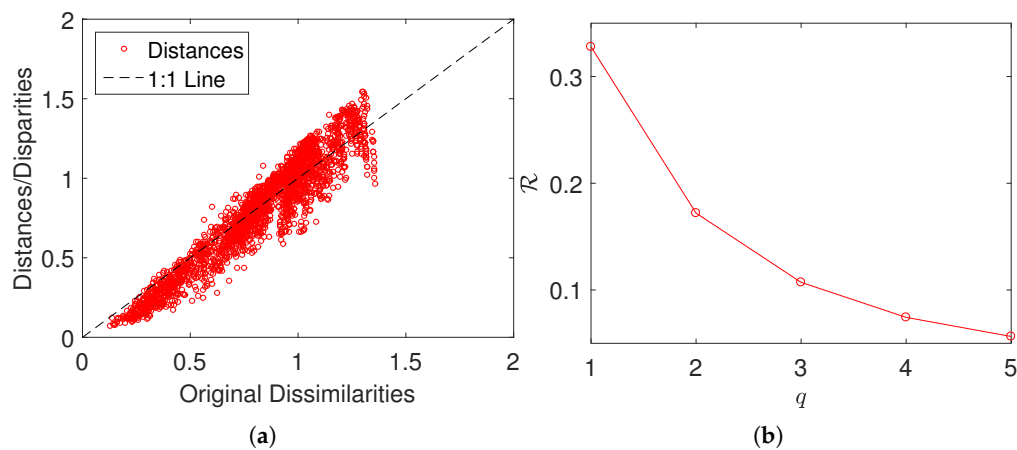
Figure 6. The MDS assessment charts obtained with Δ_{NCD}: (**a**) Sheppard; (**b**) stress.

4.2. MDS Visualization of Complexity Based on the Euclidean Distance

The six complexity indices $\{S, I, PE, C, CR, NCD_R\}$ reveal some correlation, as shown in Figure 7. Nonetheless, we conjecture that each index captures distinct details of the complex system and that a more complete assessment is accomplished when using all indices simultaneously. However, a six-dimensional visual representation is not possible and we decided to test the MDS technique for dimensionality-reduction and visualization.
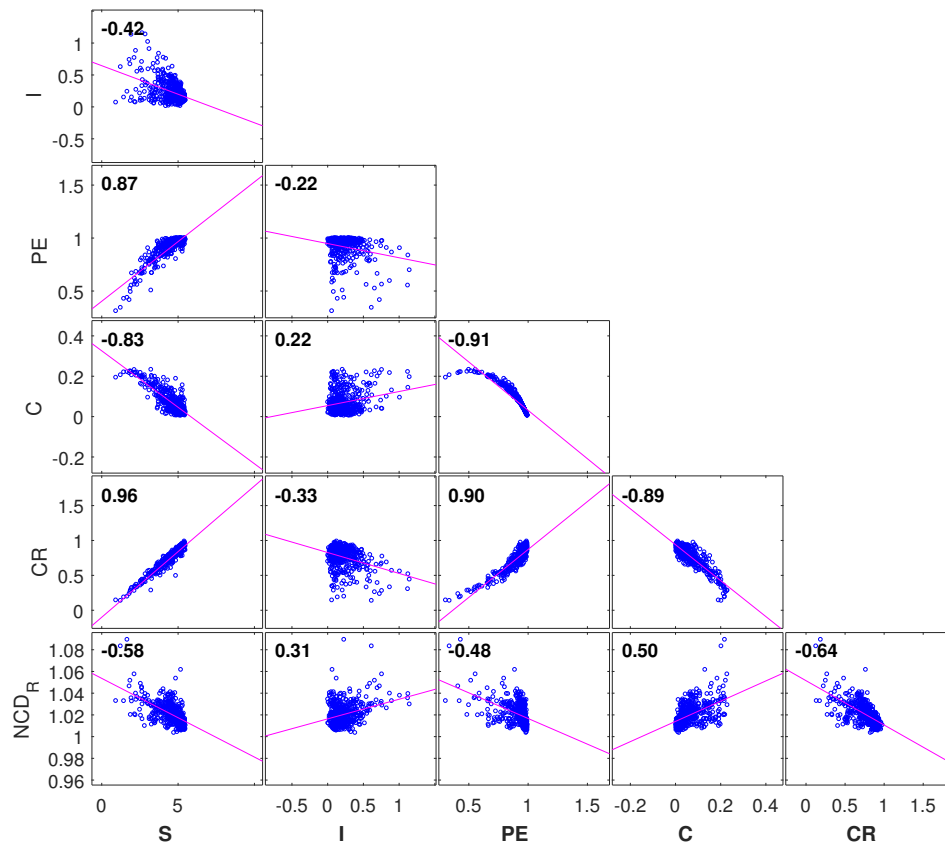
Figure 7. The correlation between the complexity indices $\{S, I, PE, C, CR, NCD_R\}$.

In a first phase a 54×6 dimensional array, $T = [t_{ik}]$, is constructed, where t_{ik}, $i = 1, \ldots, 54$, $k = 1, \ldots, 6$, represents the i-th year and the k-th complexity index. The data in T is then normalized by the mean and standard deviation to avoid numerical saturation. Therefore, the columns of T, that is, v_k, are converted to:

$$\hat{v}_k = \frac{v_k - \mu(v_k)}{\sigma(v_k)}, \tag{12}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the arithmetic mean and the standard deviation, respectively. In a second phase, the lines of the normalized array, $\hat{u}_i$, are used for calculating the dissimilarity matrix $\Delta_T = [\delta(\hat{u}_i, \hat{u}_j)]$, $i, j = 1, \ldots, 54$, where $\delta(\hat{u}_i, \hat{u}_j) = \left[\sum_k (\hat{u}_{ik} - \hat{u}_{jk})^2 \right]^{\frac{1}{2}}$ denotes the Euclidean distance between the complexity indices $\hat{u}_i$ and $\hat{u}_j$. Other distances can be adopted, but several numerical experiments that the Euclidean distance yields good results. Finally, in a third phase, the matrix Δ_T is processed by means of the MDS for constructing the loci of objects that represent the the evolution of complexity.

Figure 8 depicts the MDS three-dimensional map obtained with Δ_T for the period 1916–1969. We verify roughly the same characteristics as before for the periods, but in this case we obtain a smoother trajectory, since the adoption of six indices works like a low-pass filter of noise present in the individual complexity indices.

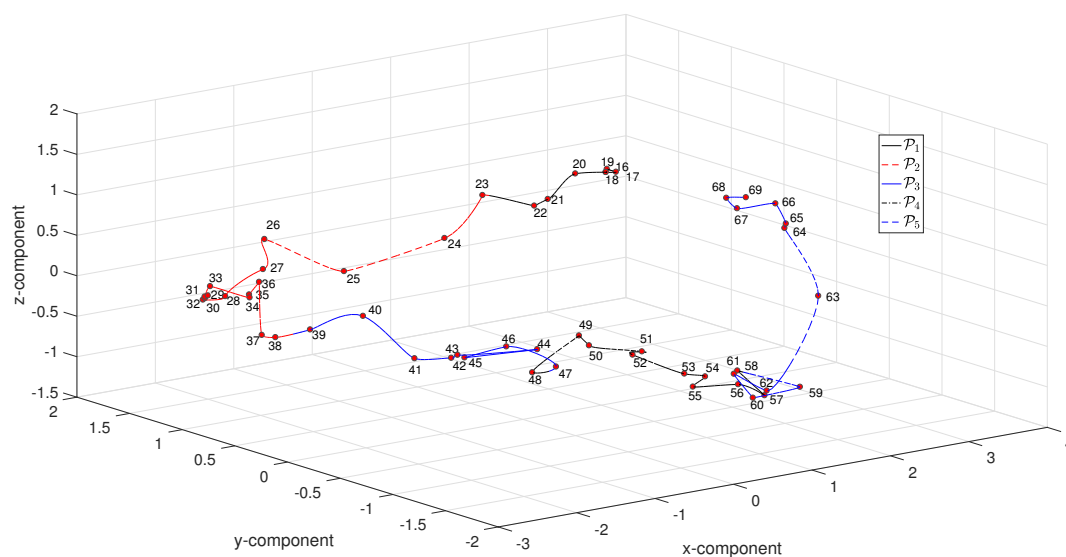The Sheppard and stress plots are omitted since they are similar to those shown in Figure 6a,b.

Figure 8. The MDS three-dimensional map obtained with Δ_T for the period 1916–1969.

5. Conclusions

We adopted six complexity indices, dimensionality-reduction and visualization techniques for studying the evolution of Escher's art. A total of 457 artworks, produced between years 1916 and 1969, were converted into digital format, discretized and represented in grayscale. The artworks were assessed by means of six distinct complexity indices and by the MDS technique. The results showed that the evolution of complexity is correlated with the periods of Escher's artistic career. On a different level, we conclude that the proposed indices represent reliable and assertive tools for assessing artwork, and motivate their adoption in other artistic manifestations of the human spirit.

References

1. Janson, H.W.; Janson, A.F. *History of Art: The Western Tradition*; Prentice Hall Professional: Upper Saddle River, NJ, USA, 2004.
2. Shiner, L. *The Invention of Art: A Cultural History*; University of Chicago Press: Chicago, IL, USA, 2001.
3. Cucker, F. *Manifold Mirrors: The Crossing Paths of the Arts and Mathematics*; Cambridge University Press: Cambridge, England, 2013.
4. Russoli, F. *Renaissance Painting*; Penguin (Non-Classics): London, UK, 1962.
5. Emmer, M. Visual harmonies: an exhibition on art and math. In *Imagine Math*; Springer: Milano, Italy, 2012; pp. 117–122.
6. Weiss, G. Geometry—Daughter of Art, Mother of Mathematics. In *The Visual Language of Technique*; Springer: Cham, Switzerland, 2015; pp. 41–57.
7. Kandinsky, W.; Rebay, H. *Point and Line Plane*; Dover Publications: New York, NY, USA, 1979.
8. Field, J.V. Linear perspective and the projective geometry of Girard Desargues. *Nuncius Ann. Storia Sci.* **1987**, *2*, 3–40. [CrossRef]
9. Jones, P.S. Brook Taylor and the mathematical theory of linear perspective. *Amer. Math. Mon.* **1951**, *58*, 595–606. [CrossRef]
10. Atalay, B. *Math and the Mona Lisa: The Art and Science of Leonardo da Vinci*; Smithsonian Institution: Washington, DC, USA, 2011.

11. Hofstadter, D.R. *Gödel, Escher, Bach: An Eternal Golden Braid, a Metaphorical Fugue on Minds and Machines in the Spirit of Lewis Carroll*; Penguin Books: New York, NY, USA, 1980.

12. Gamwell, L. *Exploring the Invisible: Art, Science, and the Spiritual*; Princeton University Press: New Jersey, NJ, USA, 2002.

13. Gamwell, L. *Mathematics and Art: A Cultural History*; Princenton University Press: New Jersey, NJ, USA, 2015.

14. Rodin, E.Y. *The Visual Mind: Art and Mathematics*; MIT Press: Cambridge, MA, USA, 1993.

15. Perc, M.; Jordan, J.J.; Rand, D.G.; Wang, Z.; Boccaletti, S.; Szolnoki, A. Statistical physics of human cooperation. *Phys. Rep.* **2017**, *687*, 1–51. [CrossRef]

16. Stanley, H.E. *Phase Transitions and Critical Phenomena*; Clarendon Press: Oxford, UK, 1971.

17. Machado, J.T.; Lopes, A.M. Analysis of natural and artificial phenomena using signal processing and fractional calculus. *Fract. Calc. Appl. Anal.* **2015**, *18*, 459–478. [CrossRef]

18. Dogson, N.A. Mathematical characterisation of Bridget Riley's stripe paintings. *J. Math. Arts* **2012**, *5*, 1–12.

19. Sigaki, H.Y.; Perc, M.; Ribeiro, H.V. History of art paintings through the lens of entropy and complexity. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E8585–E8594. [CrossRef]

20. Machado, J.T.; Lopes, A.M. Artistic painting: A fractional calculus perspective. *Appl. Math. Model.* **2019**, *65*, 614–626. [CrossRef]

21. Boon, J.P.; Casti, J.; Taylor, R.P. Artistic forms and complexity. *Nonlinear Dyn.-Psychol. Life Sci.* **2011**, *15*, 265.

22. Taylor, R.; Micolich, A.; Jones, D. Fractal expressionism. *Phys. World* **1999**, *12*, 1–3. [CrossRef]

23. De la Calleja, E.; Cervantes, F.; De la Calleja, J. Order-fractal transitions in abstract paintings. *Ann. Phys.* **2016**, *371*, 313–322. [CrossRef]

24. Montagner, C.; Linhares, J.M.; Vilarigues, M.; Nascimento, S.M. Statistics of colors in paintings and natural scenes. *JOSA A* **2016**, *33*, A170–A177. [CrossRef] [PubMed]

25. Koch, M.; Denzler, J.; Redies, C. $1/f^2$ Characteristics and isotropy in the Fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS ONE* **2010**, *5*, e12268. [CrossRef] [PubMed]

26. Lopes, A.; Tenreiro Machado, J. Complexity Analysis of Global Temperature Time Series. *Entropy* **2018**, *20*, 437. [CrossRef]

27. Wallraven, C.; Cunningham, D.W.; Fleming, R. Perceptual and Computational Categories in Art. In Proceedings of the Computational Aesthetics 2008: Eurographics Workshop on Computational Aesthetics, Lisbon, Portugal, 18–20 June 2008.

28. Kim, D.; Son, S.W.; Jeong, H. Large-scale quantitative analysis of painting arts. *Sci. Rep.* **2014**, *4*, 7370. [CrossRef] [PubMed]

29. Lee, B.; Kim, D.; Jeong, H.; Sun, S.; Park, J. Understanding the historic emergence of diversity in painting via color contrast. *arXiv* **2017**, arXiv:1701.07164.

30. Escher, M.C. *MC Escher: The Graphic Work*; Taschen: Cologne, Germany, 2000.

31. Schattschneider, D.; Emmer, M. *MC Escher's Legacy*; Springer: Berlin/Heidelberg, Germany, 2003.

32. Haak, S. Transformation geometry and the artwork of MC Escher. *Math. Teach.* **1976**, *69*, 647–652.

33. Nicki, R.M.; Forestell, P.; Short, P. Uncertainty and preference for 'ambiguous' figures, 'impossible' figures and the drawings of MC Escher. *Scand. J. Psychol.* **1979**, *20*, 277–281. [CrossRef]

34. Ernst, B. *The Magic Mirror of MC Escher*; Taschen America Llc: Los Angeles, CA, USA, 2007.

35. M Lopes, A.; Tenreiro Machado, J. Tidal Analysis Using Time–Frequency Signal Processing and Information Clustering. *Entropy* **2017**, *19*, 390. [CrossRef]

36. Shannon, C. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]

37. Gray, R.M. *Entropy and Information Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany 1990.

38. Vallianatos, F.; Papadakis, G.; Michas, G. Generalized statistical mechanics approaches to earthquakes and tectonics. *Proc. Math. Phys. Eng. Sci.* **2016**, *472*, 20160497. [CrossRef]

39. Korn, G.A.; Korn, T.M. *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*; Courier Corporation: Middlesex County, MA, USA, 1968.

40. Shannon, C.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.

41. Cover, T.; Thomas, J. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1991.

42. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [CrossRef]

43. Berger, S.; Schneider, G.; Kochs, E.; Jordan, D. Permutation Entropy: Too Complex a Measure for EEG Time Series? *Entropy* **2017**, *19*, 692. [CrossRef]

44. Ribeiro, H.V.; Zunino, L.; Lenzi, E.K.; Santoro, P.A.; Mendes, R.S. Complexity-entropy causality plane as a complexity measure for two-dimensional patterns. *PLoS ONE* **2012**, *7*, e40689. [CrossRef] [PubMed]

45. Zunino, L.; Ribeiro, H.V. Discriminating image textures with the multiscale two-dimensional complexity-entropy causality plane. *Chaos Solitons Fractals* **2016**, *91*, 679–688. [CrossRef]

46. Lopez-Ruiz, R.; Mancini, H.L.; Calbet, X. A statistical measure of complexity. *Phys. Lett. A* **1995**, *209*, 321–326. [CrossRef]

47. Martin, M.; Plastino, A.; Rosso, O. Generalized statistical complexity measures: Geometrical and analytical properties. *Phys. A Stat. Mech. Appl.* **2006**, *369*, 439–462. [CrossRef]

48. Rosso, O.; Larrondo, H.; Martin, M.; Plastino, A.; Fuentes, M. Distinguishing noise from chaos. *Phys. Rev. Lett.* **2007**, *99*, 154102. [CrossRef]

49. Kolmogorov, A.N. Three approaches to the quantitative definition ofinformation'. *Probl. Inf. Transm.* **1965**, *1*, 1–7.

50. Antão, R.; Mota, A.; Machado, J.T. Kolmogorov complexity as a data similarity metric: Application in mitochondrial DNA. *Nonlinear Dyn.* **2018**, *93*, 1059–1071. [CrossRef]

51. Pinho, A.J.; Ferreira, P.J. Image similarity using the normalized compression distance based on finite context models. In Proceedings of the 18th IEEE International Conference on Image Processing, Brussels, Belgium, 11–14 September 2011; pp. 1993–1996.

52. Solomonoff, R.J. A formal theory of inductive inference. Part I. *Inf. Control.* **1964**, *7*, 1–22. [CrossRef]

53. Chaitin, G.J. On the length of programs for computing finite binary sequences. *J. ACM* **1966**, *13*, 547–569. [CrossRef]

54. Wallace, C.S.; Boulton, D.M. An information measure for classification. *Comput. J.* **1968**, *11*, 185–194. [CrossRef]

55. Lempel, A.; Ziv, J. On the complexity of finite sequences. *IEEE Trans. Inf. Theory* **1976**, *22*, 75–81. [CrossRef]

56. Gordon, G. Multi-dimensional linguistic complexity. *J. Biomol. Struct. Dyn.* **2003**, *20*, 747–750. [CrossRef] [PubMed]

57. Dix, T.I.; Powell, D.R.; Allison, L.; Bernal, J.; Jaeger, S.; Stern, L. Comparative analysis of long DNA sequences by per element information content using different contexts. *BMC Bioinform.* **2007**, *8*, S10. [CrossRef] [PubMed]

58. Bennett, C.H.; Gács, P.; Li, M.; Vitányi, P.M.; Zurek, W.H. Information distance. *IEEE Trans. Inf. Theory* **1998**, *44*, 1407–1423. [CrossRef]

59. Fortnow, L.; Lee, T.; Vereshchagin, N. Kolmogorov complexity with error. In *Annual Symposium on Theoretical Aspects of Computer Science*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 137–148.

60. Li, M.; Chen, X.; Li, X.; Ma, B.; Vitányi, P. The similarity metric. In Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms, Baltimore, MD, USA, 12–14 January 2003; pp. 863–872.

61. Cilibrasi, R.; Vitányi, P.M. Clustering by compression. *IEEE Trans. Inf. Theory* **2005**, *51*, 1523–1545. [CrossRef]

62. Cebrián, M.; Alfonseca, M.; Ortega, A. Common pitfalls using the normalized compression distance: What to watch out for in a compressor. *Commun. Inf. Syst.* **2005**, *5*, 367–384.

63. Baker, F.; Porollo, A. CoeViz: A Web-Based Integrative Platform for Interactive Visualization of Large Similarity and Distance Matrices. *Data* **2018**, *3*, 4. [CrossRef]

64. Fiori, S. Visualization of Riemannian-manifold-valued elements by multidimensional scaling. *Neurocomputing* **2011**, *74*, 983–992. [CrossRef]

65. Berrar, D.; Ohmayer, G. Multidimensional scaling with discrimination coefficients for supervised visualization of high-dimensional data. *Neural Comput. Appl.* **2011**, *20*, 1211–1218. [CrossRef]

66. Saeed, N.; Nam, H.; Haq, M.I.U.; Muhammad Saqib, D.B. A Survey on Multidimensional Scaling. *ACM Comput. Surv.* **2018**, *51*, 47. [CrossRef]

67. Machado, J.T.; Lopes, A.M. Multidimensional scaling analysis of soccer dynamics. *Appl. Math. Model.* **2017**, *45*, 642–652. [CrossRef]

68. Tenreiro Machado, J.; Lopes, A.; Galhano, A. Multidimensional scaling visualization using parametric similarity indices. *Entropy* **2015**, *17*, 1775–1794. [CrossRef]

69. Machado, J.; Mendes Lopes, A. Fractional Jensen–Shannon analysis of the scientific output of researchers in fractional calculus. *Entropy* **2017**, *19*, 127. [CrossRef]