# Toward Auvers Period: Evolution of van Gogh's Style

Jaume Rigau[1]    Miquel Feixas[1]    Mateu Sbert[1]    Christian Wallraven[2]

[1]Graphics and Imaging Laboratory, University of Girona, Spain
[2]MPI for Biological Cybernetics, Tuebingen, Germany; Brain and Cognitive Engineering, Korea University, Korea

**Abstract**

*In this paper, we analyze the evolution of van Gogh's style toward the Auvers final period using informational measures. We will try to answer the following questions: Was van Gogh exploring new ways toward changing his style? Can informational measures support the claim of critics on the evolution of his palette and composition? How "far" was van Gogh's last period from the previous ones, can we find out an evolutionary trend? We will extend here the measures defined in our previous work with novel measures taking into account spatial information and will present a visual tool to examine the palette. Our results confirm the usefulness of an approach rooted in information theory for the aesthetic study of the work of a painter.*

Categories and Subject Descriptors (according to ACM CCS): Image Processing and Computer Vision [I.4.9]: Applications—Computer Applications [J.5]: Arts and Humanities—

## 1. Introduction

Computer-aided analysis of art has gathered a large amount of interest recently with surprisingly simple algorithms and techniques being able to detect, for example, non-original Pollock [TMJ99] and Breughel [HGR10] paintings, to model the aesthetic perception of photographs [DJLW06], or to categorize art periods [WFC*09]. Apart from the economical interest in the outcome of these algorithms, more importantly such computational analyses can help to quantify underlying principles in aesthetic perception —principles that have so far largely defied a mathematical modeling. Whether taking insights from computer vision [DJLW06], or from efficient, neurally plausible coding mechanisms [HGR10], novel methods from statistical machine learning and especially from information theory are starting to relate artistic developments to specific changes in the palette, brush, or texture statistics. In this paper, we present results on how novel information-theory based measures can be used to trace the aesthetic development of Van Gogh throughout his oeuvre.

In order to correlate aesthetic judgments of pictorial art with mathematical models, different measures and techniques based mainly on information theory have been investigated to determine the information content of a painting [RFS08a]. These measures were based on the entropy of the palette, the Kolmogorov complexity of an image measured using an image compressor, and the compositional complexity of the painting. Some of these measures were shown to correlate surprisingly well with the six different periods of van Gogh's paintings [RFS08b], as classified by critics. However, only a reasonable subset of the paintings was taken into account, and in addition interesting questions, already pointed out by critics previously, surfaced in our results. For instance, was van Gogh heading in Auvers for a new style? Was Paris a laboratory for his subsequent paintings?

In this paper we want to investigate further —using this time the full set of color digital images of van Gogh's paintings available in *The Vincent van Gogh Gallery* of David Brooks [Bro10]— whether key features of van Gogh periods can be determined by an extended set of informational measures. We will focus mainly on van Gogh's Auvers period and will try to investigate whether our measures can support the claim of art critics on his evolution of palette and composition. We will also study how far van Gogh's last period was from his other periods, and try to trace his artistic development. To this end, we will employ our previously defined measures together with a set of novel measures that take into account spatial information. In addition, we will also introduce a novel visual tool to easier analyze the palette.

This paper is organized as follows: Section 2 reviews some previous work on aesthetic measures. In Section 3, the information-theoretic measures used in the paper are presented. Section 4 interprets the measures to shed light on van Gogh's style evolution. Section 5 presents conclusions and future work.

## 2. Informational Aesthetics

Ever since a measure of aesthetics was defined by George D. Birkhoff [Bir33] as the ratio between *order* and *complexity*, different authors have introduced diverse measures that quantify the degree of order and complexity of a work of art [Ben69, Mol68, MC98, Kos98, SN04] (see also Greenfield's [Gre05] and Hoenig's [Hoe05] surveys). Using information theory, Bense [Ben69] transformed Birkhoff's measure into an informational measure based on entropy. He assigned a complexity to the repertoire or palette, and an order to the distribution of its elements on the work of art. According to Bense, in any artistic process of creation, there exists a determined repertoire of elements (such as a palette of colours, sounds, phonemes, etc.) that is transmitted to the final product; thus, the creative process is also a selective process.

Rigau et al. [RFS08a] presented a set of information-theoretic measures to study some informational aspects of a painting related to its palette and composition. Some of these measures, based on the entropy of the palette, the compressibility of the image, and an information channel to capture the composition of a painting, were used to discriminate different painting styles [RFS08a] and to analyze the evolution of van Gogh's artwork [RFS08b], revealing a significant correlation between the values of the measures and van Gogh's artistic periods. These measures are reviewed in the next section. In two recent papers, we have also shown how these measures can not only help to categorize art into different periods [WFC*09], but also how they are able to model loci of interest when observers view an artwork, that is, where gaze is attracted in an artwork [WCR*09].

## 3. Information in a Painting

To further study the evolution of van Gogh's artwork, we use five measures based on palette entropy, compressibility, compositional complexity, randomness (entropy rate), and structural complexity (excess entropy). The first three were already used in [RFS08a]. While the entropy of the palette only takes into account the color diversity, the other measures also consider its spatial distribution. In fact, these measures are not fully independent but offer complementary views of complexity in an image, as we will see in the analysis of the results in Sec. 4.

From a given color image $\mathcal{I}$ of $N$ pixels, we use its sRGB and HSV representations to study the behavior of the proposed measures:

- sRGB color representation is based on a repertoire of $256^3$ colors and its alphabet is given by $\mathcal{X}_{\mathrm{rgb}}$. From this space, we also consider the luminance function $Y_{709}$, which is a measure of the density of luminous intensity of a pixel computed as a linear combination of its RGB channels (we use the Rec. 709: $Y = 0.212671R + 0.715160G + 0.072169B$). In this case, the alphabet is represented by $\mathcal{X}_\ell = [0, 255]$.
- HSV (hue, saturation, value) is a cylindrical-coordinate representation of sRGB which is more perceptually plausible than the sRGB cartesian representation. In this case, the alphabets are represented by $\mathcal{X}_{\mathrm{H}}$, $\mathcal{X}_{\mathrm{S}}$, and $\mathcal{X}_{\mathrm{V}}$, according to a given discretization of each parameter.

From the normalization of the corresponding histograms of the alphabets of the color representations, the probability distributions of the corresponding random variables ($X_{\mathrm{rgb}}$, $X_\ell$, $X_{\mathrm{H}}$, $X_{\mathrm{S}}$, and $X_{\mathrm{V}}$) are determined, which represent the *palette* features of a painting. The palette is considered as the finite and discrete range of colors used by the artist.

## 3.1. Redundancy of the Palette

The *entropy $H(C)$* of a random variable $C$ taking values $c$ in $\mathcal{C}$ with distribution $p(c) = probability[C = c]$ is defined by

$$H(C) = -\sum_{c \in \mathcal{C}} p(c) \log p(c), \qquad (1)$$

where logarithms are taken in base 2 and entropy is expressed in bits. In this paper, the set $\mathcal{C}$ will stand for color alphabets (e.g., $\mathcal{X}_{\mathrm{rgb}}$), where $C$ represents its corresponding random variable (e.g., $X_{\mathrm{rgb}}$). The maximum entropy $H_{\max}$ of a random variable is $\log |\mathcal{C}|$. The palette entropy $H(C)$ fulfills $0 \leq H(C) \leq H_{\max}$ and can be interpreted as the average color uncertainty of a pixel. Following Bense's proposal of using redundancy to measure *order* in an aesthetic object [Ben69], the *relative redundancy* of the palette is given by

$$M_B = \frac{H_{\max} - H(C)}{H_{\max}}. \qquad (2)$$

$M_B$ takes values in $[0, 1]$ and expresses the reduction of pixel uncertainty due to the choice of a palette with a given color probability distribution instead of a uniform distribution [RFS08a]. In our tests, $M_B$ has been computed from alphabet $\mathcal{X}_{\mathrm{rgb}}$.

## 3.2. Compressibility

The *Kolmogorov complexity $K(\mathcal{I})$* of an image $\mathcal{I}$ is the length of the shortest program to compute $\mathcal{I}$ on an appropriate universal computer [LV97]. From a Kolmogorov complexity perspective, the *order* in an image can be measured by the difference between the image size (obtained using a constant length code for each color) and its Kolmogorov complexity. The ratio between the order and the initial image size is given by

$$M_K = \frac{N \times H_{\max} - K(\mathcal{I})}{N \times H_{\max}}. \qquad (3)$$

$M_K$ takes values in $[0,1]$ and the higher the $M_K$ the higher the compression ratio [RFS08a]. Because of the noncomputability of $K$, we use real-world compressors (e.g., PNG or JPEG) to estimate it, that is, we approximate the value of $K$ by the size of the corresponding compressed file [LCL*04]. A compressor exploits both the palette redundancy and the color position in the canvas. In our experiments, JPEG compression is used because of its ability to discover such spatial "patterns". To avoid losing significant information, we use a JPEG compressor with the maximum quality option.

### 3.3. Compositional Complexity

The creative process described by Bense can be understood as the realization of an information channel between the palette and the set of regions of the image [RFS08a]. This channel is defined between the random variables $C$ (input) and $R$ (output), which represent the set of bins ($\mathcal{C}$) of the color histogram and the set of regions ($\mathcal{R}$) of the image, respectively. The *mutual information* $I(C,R)$ between $C$ and $R$ represents the *correlation* between colors and regions. For an image $\mathcal{I}$ decomposed into $n$ regions, the *ratio of mutual information* is defined by

$$M_s(n) = \frac{I(C,R)}{H(C)}, \qquad (4)$$

where $H(C)$ is the maximum value achievable for $I(C,R)$ (when each region coincides with a pixel) [RFS08a]. Given a partitioning algorithm, the inverse function

$$M_s^{-1}\left(\frac{I(C,R)}{H(C)}\right) = n \qquad (5)$$

yields the number of regions obtained from a given mutual information ratio. The number of regions needed to extract a given quantity of information is taken as a measure of *compositional complexity*. To compute the number of regions, a BSP partitioning algorithm is used to produce quasi-homogeneous regions extracting at each step the maximum mutual information of the painting [RFS08a]. The global composition of the image can be obtained after relatively few partitions, while the details or forms in the painting begin to appear at a finer scale. In our tests, $I(C,R)$ and $H(C)$ have been computed from alphabet $\mathcal{X}_\ell$.

### 3.4. Randomness

The notation used here is inspired by the work of Feldman and Crutchfield [CF03]. Given a chain $\ldots X_{-2}X_{-1}X_0X_1X_2\ldots$ of random variables $X_i$ taking values in $\mathcal{X}$, a block of $L$ consecutive random variables is denoted by $X^L = X_1 \ldots X_L$. The probability that the particular $L$-block $x^L$ occurs is denoted by $p(x^L)$. The Shannon entropy of length-$L$ sequences or *L-block entropy* is defined by

$$H(X^L) = -\sum_{x^L \in \mathcal{X}^L} p(x^L) \log p(x^L), \qquad (6)$$

where the sum runs over all possible $L$-blocks. The *entropy rate* is defined by

$$h^x = \lim_{L\to\infty} \frac{H(X^L)}{L} = \lim_{L\to\infty} h^x(L), \qquad (7)$$

where $h^x(L) = H(X_L|X_{L-1}, X_{L-2}, \ldots, X_1)$ is the entropy of a symbol conditioned on a block of $L-1$ adjacent symbols. The entropy rate of a sequence measures the average amount of information (i.e., irreducible randomness) per symbol $x$ and the optimal achievement for any possible compression algorithm [CT91, Fel02]. Entropy rate can be also seen as the uncertainty associated with a given symbol if all the preceding symbols are known.

The entropy rate of an image quantifies the average uncertainty surrounding a pixel, that is, the difficulty of predicting the color of its neighbor pixels. While a painting that is highly random is difficult to compress, a painting with low randomness has many correlations with pixel colors. It is interesting to note that $\log|\mathcal{C}| - h^x$ can be also considered as a measure of redundancy in a painting.

In the context of an image, $\mathcal{X}$ represents the color alphabet and $x^L$ is given by a set of $L$ *neighbor* pixel intensity values. In practice, we cannot compute $L$-block entropies for high $L$, due to the exponential size —$N^L$, where $N$ is the cardinality of $\mathcal{X}$— of the joint histogram. In our tests (see Sec. 4), the entropy rate has been estimated taking $L$-block samples radially around each pixel. This pixel represents the origin and becomes the first element of the block. To carry out the computations, we set $L = 3$ and $N = 256$. Using digital photography software, we have conducted experiments that showed a positive correlation between entropy rate and contrast.

### 3.5. Structural complexity

A complementary measure to the entropy rate is the *excess entropy*, which is a measure of the *structure* of a system. The *excess entropy* is defined by

$$E = \sum_{L=1}^{\infty} (h^x(L) - h^x) \qquad (8)$$

$$= \lim_{L\to\infty} (H(X^L) - h^x L) \qquad (9)$$

and captures how $h^x(L)$ converges to its asymptotic value $h^x$. Thus, when we take into account only a few number of symbols in the entropy computation, the system appears more random than it actually is. This excess randomness tells us how much additional information must be gained about the configurations in order to reveal the actual uncertainty $h^x$. The way in which $h^x(L)$ converges to its asymptotic form tells us about the structure or correlations of a system [CF03, FMC08].

Considered by many authors as a measure of the structural complexity of a system, the excess entropy is introduced here to measure the spatial structure of a painting. If

it is large the painting contains many regularities or correlations [FMC08]. Thus, excess entropy serves to detect ordered, low entropy density patterns in a painting. In the case of a completely random image, the excess entropy should vanish, showing that correlations are not present in the image. In our tests, the excess entropy has been estimated using Equ. (9) and taking $L = 5$ and $N = 32$. While $L = 3$ is enough to compute the entropy rate, excess entropy needs larger sequences, which implies reducing the number of bins due to computational restrictions.

## 4. Artistic Analysis

In this section, we analyze how the style of van Gogh evolves toward his last period. According to art critics, in Auvers, van Gogh changes his style in the following way: he sees the Northern landscape with a sharpened and heightened vision; softens the hue in landscapes (reflecting the response to the more subdued Northern light with whites, blues, violets, and soft greens); uses harsher primary colors; exhibits a certain unevenness and impetuosity of brushstroke; and simplifies the composition (see Ronald Pickvance [Pic86]). Can our measures support these claims on the evolution of the palette and composition?

To analyze the evolution of van Gogh's style, the measures presented in Sec. 3 have been applied to a set of images of van Gogh's paintings obtained from *The Vincent van Gogh Gallery* of David Brooks [Bro10]. In this website, van Gogh's oeuvre (861 paintings) is classified into six periods. From this set of images, we have excluded 61 black and white images which were not available in color yielding a total of 800 color images for our experiments.

We will first consider the palette measures. From the entropy-based measure $M_B$ (Table 1), we can see how the palette evolves. There is a first palette simplification from the Early to Nuenen period, but starting with the Paris-period the palette entropy constantly decreases, obtaining its minimum in Auvers. It is important to note that the measure of entropy is logarithmic, that is, the constant although small increases in the measure translates into a much larger absolute increase in the variety of colors used.

In Fig. 1 we show the digital-image-palette (DIP, see Appendix A) based on the HSV representation. In this figure, we show a painting of each period (a), the DIP of this painting (b), the DIP of the period (c), and the normalized DIP (NDIP) of the period. The DIP representation has been obtained from a discretization of the hue in 360 bins ($\mathcal{X}_H$) and, for each bin, the average of both saturation and brightness is depicted together with the hue. The average of the achromatic values is represented by the gray-color of the circumference. For each painting, the frequency of bins has been weighted by the real size of canvas. Observe that the canvas size has been doubled from Paris on (Table 1). In the last row of Fig. 1, the global palette of all periods is shown. As

the figure shows, the palette gains in chromaticity (except for the somber palette of the Nuenen period) and evolves toward softer colors, becoming more and more constrained in hue space. At the same time, the palette also evolves toward more yellowish and brighter hues overall. All of this means that van Gogh was continuously evolving and optimizing his palette. Also, let us note the remarkable similarity of the global average to the Paris one, especially striking in the NDIP (Fig. 1.d) —in a way the Paris period represents van Gogh's oeuvre remarkably well.

To quantify the palette difference between periods we use a DIP-distance defined in Equ. (10). In Table 2, we show the distances between the DIPs of all periods and global artwork. The Nuenen period has the maximum average distance to the other periods, while the Paris period yields the minimum distance, even to the global palette, reinforcing the central role of Paris period in van Gogh's artwork. Interestingly, the distances from the Auvers period are more balanced toward all the other periods, being of course closer to Saint-Rémy which could be due to the fact that in Auvers van Gogh reflected on all of his previous periods. Indeed, before going to Auvers, he spent some days in Paris and had the opportunity of reviewing a large part of his previous paintings, as he explains in a letter to his sister Wil [Sol07].

With respect to composition, we can group the six van Gogh periods into three distinct groups (see Table 1): Early/Nuenen, Paris/Arles, and Saint-Rémy/Auvers. The composition from one group to the other one exhibits large changes, doubling (for low mutual information ratios as 0.05 and 0.1) the number of regions to extract the same amount of information. We see a peak of compositional complexity in Saint-Rémy period, followed by a slight decrease in Auvers. Again, this quantitative findings is in accordance with critics opinion about this period with respect to the simplification of composition [Pic86].

If we accept that entropy rate measure positively correlates with contrast (see Sec. 3), then we can obtain from Table 3 that contrast decreases from the Early period to Nuenen but later constantly increases (entropy rate evolves inversely similar to the palette redundancy $M_B$). The entropy rates achieve their maximum values in the last period, which is again in accordance to art critics' prevalent analysis: the simplification of composition was accompanied by an increase in contrast [Pic86]. Table 1 also shows how the complexity $M_K$, which expresses the compression ratio, behaves in an inverse way to the evolution of entropy rate $h^x$. This behavior agrees with the fact that the entropy rate expresses the optimal achievement for a compression algorithm.

As we interpret excess entropy as a measure of the degree of correlation and patterns, we can read from Table 3 how the Auvers period presents more brightness patterns, while Arles period shows more structure in chromaticity (hue and saturation). In Fig. 2 we show two paintings of Auvers period to illustrate the behavior of the entropy rate and excess
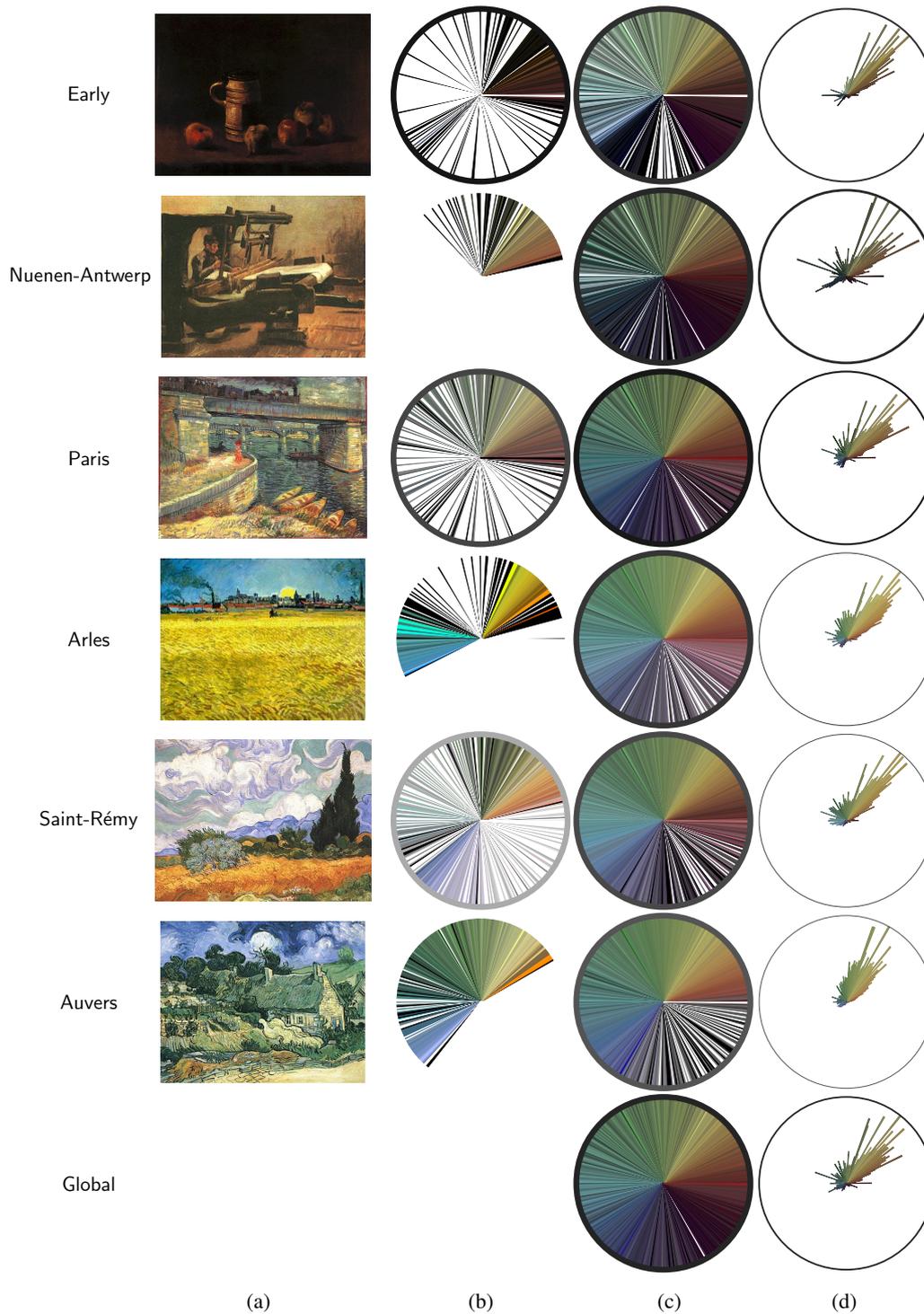
**Figure 1:** *Digital-image-palette of van Gogh's periods. (a) Painting example. (b) DIP of painting (a). (c) DIP of the period. (d) NDIP of the period. Global DIP and NDIP are shown in the last row. Painting images credit: © 1996-2010 David Brooks [Bro10].*

| Period | # | Size | $M_B$ | $M_K$ | $M_s^{-1}(0.05)$ | $M_s^{-1}(0.1)$ | $M_s^{-1}(0.15)$ | $M_s^{-1}(0.20)$ | $M_s^{-1}(0.25)$ |
|---|---|---|---|---|---|---|---|---|---|
| Early | 26 | 17.5 | 0.422 | 0.769 | 6.154 | 39.769 | 147.538 | 413.538 | 1019.000 |
| Nuenen | 172 | 23.7 | 0.486 | 0.794 | 5.727 | 33.878 | 153.953 | 479.378 | 1144.616 |
| Paris | 209 | 22.0 | 0.384 | 0.712 | 12.177 | 81.301 | 309.541 | 813.287 | 1688.938 |
| Arles | 181 | 40.2 | 0.351 | 0.688 | 13.376 | 93.834 | 344.094 | 869.221 | 1748.387 |
| St-Rémy | 137 | 42.8 | 0.342 | 0.665 | 27.766 | 185.985 | 587.219 | 1286.766 | 2331.175 |
| Auvers | 75 | 39.5 | 0.334 | 0.659 | 27.613 | 164.307 | 509.187 | 1130.560 | 2081.400 |
| Global | 800 | 31.1 | 0.388 | 0.713 | 14.983 | 98.300 | 344.911 | 851.988 | 1710.363 |

**Table 1:** *For each period and the global artwork, number of paintings, canvas size average ($dm^2$), $M_B$, $M_K$, and $M_s^{-1}$(0.1, 0.15, 0.20, and 0.25 are shown (N = 256 bins has been used).*

| Period | Early | Nuenen | Paris | Arles | St-Rémy | Auvers | Global | Avg |
|---|---|---|---|---|---|---|---|---|
| Early | 0.000 | 26.775 | 28.559 | 47.213 | 37.197 | 44.498 | 25.658 | 36.848 |
| Nuenen | 26.775 | 0.000 | 31.437 | 57.830 | 52.633 | 53.289 | 29.425 | 44.393 |
| Paris | 28.559 | 31.437 | 0.000 | 35.976 | 37.259 | 43.895 | 13.955 | 35.425 |
| Arles | 47.213 | 57.830 | 35.976 | 0.000 | 23.857 | 38.076 | 29.824 | 40.590 |
| St-Rémy | 37.197 | 52.633 | 37.259 | 23.857 | 0.000 | 31.401 | 28.120 | 36.469 |
| Auvers | 44.498 | 53.289 | 43.895 | 38.076 | 31.401 | 0.000 | 36.074 | 42.232 |
| Global | 25.658 | 29.425 | 13.955 | 29.824 | 28.120 | 36.074 | 0.000 | 32.611 |

**Table 2:** *DIP-distance matrix between periods and the global artwork. The average column is only computed from period columns.*

entropy. Observe first that entropy rates of top painting are higher, specially for the brightness. This matches with the high contrasted spots in the foliage of the trees due to the diversity in the illumination and chroma of the leaves. On the other hand, the sheaves of wheat and the background present a more uniform color which translates in lower entropy rate values. The excess entropy of the top painting is also higher revealing more patterns than the bottom one. This is due to the fact that the apparent randomness of the color of the pixel of the leaves disappears when we take into account the correlations in the sequences of pixels. This is, we discover order out of apparent randomness. In the bottom image, either the sequences of pixels studied are too short, due to computational limitations ($L = 5$ and $N = 32$), or the uniformity is higher from the beginning. In either case, the uniformity discovered out of randomness is lower.

Addressing the question whether van Gogh was exploring new ways toward changing his style, we can answer for the Auvers period that the measures, indeed, reflect the fact that van Gogh traded off simplified composition against an extended palette and increased contrast. Furthermore, palette extension, contrast increase, and compositional complexity increase can be seen as van Gogh's aesthetic development from his Paris period to Saint-Rémy.

As we have seen, the case of the Paris period is interesting in that, for almost all considered measures, this period closely approximates the global average. Given that for art critics this period constitutes an exploratory phase for van
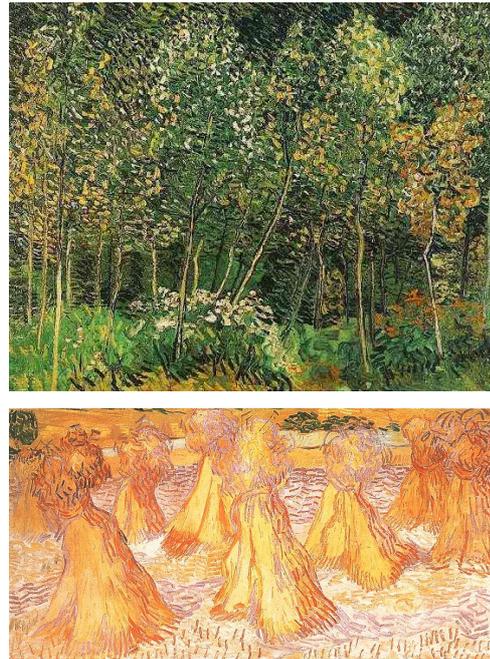


**Figure 2:** *Entropy rate and excess entropy values of two paintings of Auvers period: (top) $h_{HSV}^x = (5.718, 7.696, 9.279)$ and $E_{HSV} = (1.189, 3.602, 5.719)$; (bottom) $h_{HSV}^x = (4.156, 7.215, 5.000)$ and $E_{HSV} = (0.492, 1.733, 0.366)$.*

| Period | $h_H^x$ | | $h_S^x$ | | $h_V^x$ | | $E_H$ | | $E_S$ | | $E_V$ | |
|--------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Name | $\bar{x}$ | $s(x)$ | $\bar{x}$ | $s(x)$ | $\bar{x}$ | $s(x)$ | $\bar{x}$ | $s(x)$ | $\bar{x}$ | $s(x)$ | $\bar{x}$ | $s(x)$ |
| Early | 5.251 | 0.724 | 6.551 | 0.443 | 6.802 | 0.452 | 1.791 | 0.376 | 1.785 | 0.523 | 1.961 | 0.477 |
| Nuenen | 4.899 | 0.810 | 6.219 | 0.602 | 6.215 | 0.776 | 1.739 | 0.345 | 1.590 | 0.557 | 1.669 | 0.534 |
| Paris | 5.564 | 0.709 | 6.802 | 0.469 | 6.972 | 0.405 | 1.771 | 0.360 | 1.898 | 0.477 | 1.886 | 0.382 |
| Arles | 5.855 | 0.598 | 7.034 | 0.305 | 7.272 | 0.286 | 1.935 | 0.309 | 2.063 | 0.381 | 2.151 | 0.366 |
| St-Rémy | 5.859 | 0.652 | 6.972 | 0.322 | 7.427 | 0.248 | 1.829 | 0.354 | 1.906 | 0.385 | 2.180 | 0.312 |
| Auvers | 5.925 | 0.531 | 7.116 | 0.309 | 7.471 | 0.305 | 1.880 | 0.279 | 2.009 | 0.339 | 2.223 | 0.305 |
| Global | 5.561 | 0.804 | 6.780 | 0.502 | 6.996 | 0.547 | 1.822 | 0.361 | 1.877 | 0.492 | 1.984 | 0.520 |

**Table 3:** *For each period and the global artwork, average of entropy rate $h^x$ and excess entropy $E$ for hue (H), saturation (S), and brightness value (V) are shown. Entropy rate has been computed using $L = 3$ and $N = 256$, and excess entropy using $L = 5$ and $N = 32$. The standard deviation is shown for each measure.*

Gogh, our results show that in this period, indeed, past and future styles are being contained and tested.

## 5. Conclusions

In this paper we have traced the artistic development of van Gogh's style using information-theoretic measures, introducing two novel measures for contrast and structure in an image, as well as a novel, visual tool to analyze the palette. Our results show how the Paris period (a period in which van Gogh's painting underwent crucial changes that would define his characteristic style later) is, indeed, a period of experimentation: it seems to contain a representation of all previous and later periods. From Paris till Auvers, the evolution of the style shows an enrichment of the chromaticity and an increase in the contrast. The compositional complexity follows the same trend in Arles and Saint-Rémy. In Auvers, van Gogh moved toward simplifying the composition. We can thus say that van Gogh in the last period traded a simplification in palette with an increase in contrast and chromaticity. In the future, we plan to explore the Paris period in more detail and to introduce long range correlations in entropy rate and excess entropy measures to better capture the structure of a painting. We will also apply the proposed measures to identify different periods of other artists.

## Appendix A: Digital-Image-Palette

In order to represent the palette of an image, we define the Digital-Image-Palette (DIP) based on the next rules:

- The HSV color representation is selected to depict the colors of the palette with hue, saturation, and value. We consider the cylindric representation with $H \times S \times V$ in the range $[0°, 360°] \times [0, 1] \times [0, 1]$.
- The hue $h$ of an $hsv$ value refers to a pure color without tint or shade (addition of white or black pigment, respectively); the value $v$ represents the brightness relative to the brightness of a similarly illuminated white; and the saturation $s$ represents the colorfulness relative to its own brightness $v$.

- The space is discretized into $N$ bins $H_i$ (e.g., 360) where each one corresponds to a cylindrical sector. A bin $H_i$ represents all the colors that fall inside it.
- The achromatic colors (gray-scale) have an undefined hue and a null saturation. Thus, we consider $N$ chromatic bins and one achromatic: $M = N + 1$ bins.
- $F_i$ is the frequency of $H_i$ weighted by the real size of canvas in order to avoid the heterogeneous scales of the images with respect to the real size of the paintings.
- The hue $h_i$ assigned to a sector $H_i$ is given by the angle of the middle of the arc of the sector.
- A point in the HSV space is projected into the plane $S \times V$ of its corresponding $H_i$. This projected point is represented by $\vec{sv}$ containing the saturation and brightness information.

The DIP is obtained according to the next steps:

1. For each pixel $p \in \mathcal{I}$ do
   a. $hsv = HSV(RGB(p))$
   b. $H_i \leftarrow h, \ i \in \{1, \ldots, M\}$
   c. Increase $F_i$
   d. Add $\vec{sv}$ into $H_i$

2. For each $H_i$ do
   a. $\vec{SV}_i$ = vectorial sum for all $\vec{sv}$ in $H_i$
   b. $\vec{\bar{sv}}_i$ = normalization of $\vec{SV}_i$ from $F_i$
   c. $hsv_i = (h_i, \pi_1(\vec{\bar{sv}}_i), \pi_2(\vec{\bar{sv}}_i))$
   d. Paint sector $H_i$ with color $hsv_i$

The visual representation of a DIP is composed by the set of sectors $H_i$ in a circle of unitary radius for the chromatic colors, and by a circumference painted with the achromatic value (Fig. 3). The frequency $F_i$ is normalized ($f_i$) to represent a normalized DIP (NDIP). Its visualization uses variable radius (chromatic colors) and the circumference width (achromatic colors) to express that normalization.

The dissimilarity, or DIP-distance, between two DIPs $i$ and $j$ is defined by

$$d_{ij} = \frac{1}{M} \sum_{k=1}^{M} |f_{i_k} \times \vec{sv}_{i_k} - f_{j_k} \times \vec{sv}_{j_k}|. \qquad (10)$$
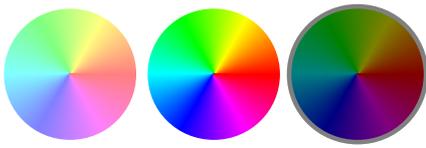
**Figure 3:** *DIP representation examples. For all hues, (left) $s = 0.5$ and $v = 1$, (center) $s = 1$ and $v = 1$, and (right) $s = 1$ and $v = 0.5$. The achromatic value v is represented on the border of the circumference.*

## References

[Ben69]   BENSE M.: *Introduction to the Information-theoretical Aesthetics. Foundation and Application in the Text Theory*. Rowohlt Taschenbuch Verlag GmbH., 1969.

[Bir33]   BIRKHOFF G. D.: *Aesthetic Measure*. Harvard University Press, 1933.

[Bro10]   BROOKS D.: The Vincent van Gogh Gallery. http://www.vggallery.com, 2010.

[CF03]   CRUTCHFIELD J. P., FELDMAN D. P.: Regularities unseen, randomness observed: Levels of entropy convergence. *Chaos 15* (2003), 25–54.

[CT91]   COVER T. M., THOMAS J. A.: *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[DJLW06]   DATTA R., JOSHI D., LI J., WANG J. Z.: Studying aesthetics in photographic images using a computational approach. In *Computer Vision - ECCV 2006*, vol. 3953-III of *Lecture Notes in Computer Science*. Springer-Verlag, 2006, pp. 288–301.

[Fel02]   FELDMAN D. P.: *A Brief Introduction to: Information Theory, Excess Entropy and Computational Mechanics*. Lecture notes, Department of Physics, University of California, Berkeley (CA), USA, 2002. http://hornacek.coa.edu/dave/.

[FMC08]   FELDMAN D. P., MCTAGUE C. S., CRUTCHFIELD J. P.: The organization of intrinsic computation: Complexity-entropy diagrams and the diversity of natural information processing. *Chaos 18* (October 2008), 043106.

[Gre05]   GREENFIELD G.: On the origins of the term "computational aesthetics". In *Computational Aesthetics 2005. Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (May 2005), Eurographics Association, pp. 9–12.

[HGR10]   HUGHES J. M., GRAHAM D. J., ROCKMORE D. N.: Quantification of artistic style through sparse coding analysis in the drawings of pieter bruegel the elder. *Proceedings of the National Academy of Sciences of the USA 107*, 4 (January 2010), 1279–1283.

[Hoe05]   HOENIG F.: Defining computational aesthetics. In *Computational Aesthetics 2005. Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (May

2005), Neumann L., Sbert M., Gooch B., Purgathofer W., (Eds.), Eurographics Association, pp. 13–18.

[Kos98]   KOSHELEV M.: Towards the use of aesthetics in decision making: Kolmogorov complexity formalizes Birkhoff's idea. *Bulletin of the European Association for Theoretical Computer Science 66* (October 1998), 166–170.

[LCL*04]   LI M., CHEN X., LI X., MA B., VITÁNYI P. M. B.: The similarity metric. *IEEE Transactions on Information Theory 50*, 12 (2004), 3250–3264. First refered in Symposium on Discrete Algorithms '03 (Li03A).

[LV97]   LI M., VITÁNYI P. M. B.: *An Introduction to Kolmogorov Complexity and Its Applications*. Graduate Texts in Computer Science. Springer-Verlag, 1997.

[MC98]   MACHADO P., CARDOSO A.: Computing aesthetics. In *Advances in Artificial Intelligence. 14th Brazilian Symposium on Artificial Intelligence, SBIA'98, Porto Alegre, Brazil, November 4-6, 1998. Proceedings* (1998), vol. 1515 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 219–228.

[Mol68]   MOLES A.: *Information Theory and Esthetic Perception*. University of Illinois Press, 1968.

[Pic86]   PICKVANCE R.: *Van Gogh in Saint-Rémy and Auvers*. The Metropolitan Museum of Art (Harry N. Abrams, Inc.), 1986. Published in conjunction with the exhibition *Van Gogh in Saint-Rémy and Auvers*, held at The Metropolitan Museum of Art, New York, 25 November 1986 – 22 March 1987.

[RFS08a]   RIGAU J., FEIXAS M., SBERT M.: Informational aesthetics measures. *IEEE Computer Graphics and Applications 28*, 2 (March/April 2008), 24–34.

[RFS08b]   RIGAU J., FEIXAS M., SBERT M.: Informational dialogue with Van Gogh's paintings. In *Computational Aesthetics 2008. Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (June 2008), Eurographics Association, pp. 115–122.

[SN04]   SVANGÅRD N., NORDIN P.: Automated aesthetic selection of evolutionary art by distance based classification of genomes and phenomes using the universal similarity metric. In *Applications of Evolutionary Computing, EvoWorkshops2004: EvoBIO, EvoCOMNET, EvoHOT, EvoIASP, EvoMUSART, EvoSTOC* (2004), vol. 3005 of *Lecture Notes in Computer Science*, Springer-Verlag, pp. 447–456.

[Sol07]   SOLANA G.: Epilogue in auvers. In *Van Gogh. Los Últimos Paisajes*. Fundación Colección Thyssen-Bornemisza, 2007, pp. 153–163.

[TMJ99]   TAYLOR R. P., MICOLICH A. P., JONAS D.: Fractal analysis of pollock's drip paintings. *Nature 399* (June 1999), 422.

[WCR*09]   WALLRAVEN C., CUNNINGHAM D., RIGAU J., FEIXAS M., SBERT M.: Aesthetic appraisal of art — from eye movements to computers. In *Computational Aesthetics 2009. Eurographics Workshop on Computational Aesthetics in Graphics, Visualization and Imaging* (May 2009), Eurographics Association, pp. 137–144.

[WFC*09]   WALLRAVEN C., FLEMING R., CUNNINGHAM D., RIGAU J., FEIXAS M., SBERT M.: Categorizing art: Comparing humans and computers. *Computers & Graphics 33*, 4 (August 2009), 484–495.