Dissecting landscape art history with information theory

Byunghwee Lee[a] , Min Kyung Seo[b,1] , Daniel Kim[c] , In-seob Shin[b] , Maximilian Schich[d,e], Hawoong Jeong[a,f,2] , and Seung Kee Han[b,2]

[a]Department of Physics, Korea Advanced Institute of Science and Technology, Daejeon 34141, Korea; [b]Department of Physics, Chungbuk National University, Cheongju, Chungchungbuk-do 28644, Korea; [c]Department of DigITal, Merck Sharp and Dohme Korea, Seoul 04637, Korea; [d]Cultural Data Analytics, Tallinn University, Tallinn 10120, Estonia; [e]Edith O'Donnell Institute for Art History, The University of Texas at Dallas, Richardson, TX 75044; and [f]The Asia Pacific Center for Theoretical Physics, Pohang, Gyeongsangbuk-do 37673, Korea

Painting has played a major role in human expression, evolving subject to a complex interplay of representational conventions, social interactions, and a process of historization. From individual qualitative work of art historians emerges a metanarrative that remains difficult to evaluate in its validity regarding emergent macroscopic and underlying microscopic dynamics. The full scope of granular data, the summary statistics, and consequently, also their bias simply lie beyond the cognitive limit of individual qualitative human scholarship. Yet, a more quantitative understanding is still lacking, driven by a lack of data and a persistent dominance of qualitative scholarship in art history. Here, we show that quantitative analyses of creative processes in landscape painting can shed light, provide a systematic verification, and allow for questioning the emerging metanarrative. Using a quasicanonical benchmark dataset of 14,912 landscape paintings, covering a period from the Western renaissance to contemporary art, we systematically analyze the evolution of compositional proportion via a simple yet coherent information-theoretic dissection method that captures iterations of the dominant horizontal and vertical partition directions. Tracing frequency distributions of seemingly preferred compositions across several conceptual dimensions, we find that dominant dissection ratios can serve as a meaningful signature to capture the unique compositional characteristics and systematic evolution of individual artist bodies of work, creation date time spans, and conventional style periods, while concepts of artist nationality remain problematic. Network analyses of individual artists and style periods clarify their rhizomatic confusion while uncovering three distinguished yet nonintuitive supergroups that are meaningfully clustered in time.

art history | paintings | proportion | information theory | composition

Understanding how artistic expressions and design principles have changed over time is a central question in art history, aesthetics, and cultural evolution (1–7) as individual artists reflected aesthetic values through their artworks, while aggregate notions of zeitgeist remain theoretically contested. In visual art, an artist often determines the main compositional characteristics of an artwork through an interplay of nonexplicit latent variables that are often imperfectly summarized in categorical concepts, including visual elements, such as the formalist notions of line, shape, tone, color, pattern, texture, form, etc. The artistic outcome as a whole, which is consequently the subject of art historical, critical, and aesthetic description, takes into account a potentially great variety of such latent factors summing up to a latent system of organizational principles. Among the diverse principles of organization, compositional techniques focusing on spatial arrangements within artworks have long been studied (8–10). Formalist art history on compositional techniques has focused on the rule of thirds, the golden ratio, the rule of odds, symmetry, modularity, etc. Vast amounts of qualitative research related to such principles of organization have been conducted in the fields of art history and aesthetics particularly before such

formalist analyses fell out of fashion. Meanwhile, only a few large quantitative and macroscopic studies on the spatial composition of paintings have been presented so far. Addressing an open challenge, this study develops a quantitative framework to answer two long-standing questions (11, 12): "Are there culturally and temporally transcendent design principles in painting?," and "How do such organizing principles evolve over time?"

Thanks to the recent proliferation of unprecedented numbers of large-scale digital scans of paintings (13–15), researchers have been able to develop and apply quantitative and statistical methods to study the visual arts, complementing qualitative research both validating and disproving previous insights (16). Computational assessments of visual art so far contributed to characterize diverse statistical properties of paintings such as the fractal dimension, the Fourier power spectrum, and frequency distributions in color space (17–21). Moreover, recent statistical analyses have been applied to quantify the evolution of artistic style and representation (3–7, 22–25), to authenticate and estimate creation dates (26–28); to reproduce characterizing

Significance

A foundational question in art and aesthetics is if there are culturally and temporally transcendent characteristics in the organizing principles within art and if yes, how these principles evolve over time. We propose a simple yet coherent information-theoretic framework that captures compositional proportion as used for dissecting landscape paintings by artists. The analysis of 14,912 landscape paintings representing the canonical historiography of Western art uncovers that the preferred compositional proportion within the historiography of landscape paintings systematically evolves over time. The network analysis of similarity distributions reveals clear clustering structures of individual artist's artworks, absolute time periods, and conventional style periods, suggesting meaningful supergroups of art historical concepts, which remained so far nonintuitive to a broader audience.

styles of specific artists (29), and to classify the systematic novelty of artists (30). Moreover, beyond the characterization of artworks, art historical metadata including exhibition trajectories and auction price history shed new light on the dynamics behind the careers of artists (31–33).

Some researchers devised quantitative measures for visual characteristics of artworks using concepts of information theory, including an area aptly titled critical and creative aesthetics as information processing (34).

A notable departure of study in computational aesthetics dates back to 1933 when the American mathematician Birkhoff (35) conceptualized a quantitative aesthetic measure to understand the order and complexity of artworks (36, 37). He regarded beauty as a mathematical phenomenon and introduced an aesthetic measure M, defined as the ratio between "order" O and "complexity" C, where O and C were measured based on the number of structural regularities and elements of an artwork, respectively. It is this sense of aesthetics aiming to formalize "orderliness" in artworks that has since driven further developments in the general theory of computational aesthetics, rather than the difficult or even evasive concept of "beauty," which is rooted in older literature (Baumgarten) and is more familiar to a broader audience.

Starting from Birkhoff's original idea, Bense (37) and Moles developed informational aesthetics and formulated the order and complexity of an artwork in terms of redundancy and Shannon's notion of information (38). Bense (37) considered a creative action in painting a transition process from an initial state (palette) to a final color distribution on a physical support (canvas). Inspired by Bense's idea, Rigau et al. (39) considered Bense's creative process an information channel between a color palette and an image's spatial regions. They presented a greedy algorithm that progressively partitions an image into quasihomogeneous regions by extracting the mutual information between color and regions in each step of partition until the painting is completely revealed. The procedure takes a full image as a unique initial partition and progressively subdivides it into a vertical or horizontal direction by a line that gives the maximum mutual information. Therefore, during the partitioning process, information acquisition increases, and the uncertainty of color over regions reduces. The information acquisition rate over each partition step was used to define the painting's degree of order (39). More recently, Shin et al. (40) developed a faster algorithm for the partitioning procedure (*SI Appendix*, section II has a detailed explanation of the algorithm) so that analyzing a massive set of paintings is now feasible in a much shorter timescale.

In this study, we employ Rigau's dissection algorithm, rooted in Bense's information theoretic concept regarding the process of painting, to present a systematic framework assessing the compositional proportion of landscape paintings as an essential organizing principle. Analyzing a large set of digital scans of landscape paintings, we reveal the emerging macroscopic trend of frequency distributions of compositional proportions as implied in a dataset of 14,912 paintings by 1,476 painters, covering a period from the Western renaissance to contemporary art (mostly from 1500 to 2000 CE). We are aware that our study works with digital surrogates of the original palette (red, green, and blue [RGB] values in the RGB color space) and the original paintings (digital scans). We are also aware that our study cuts out considerations of reception aesthetics (i.e., how different audiences perceive individual artworks in a variety of situations). Nevertheless, our study makes an important contribution: our study characterizes a dataset that essentially summarizes a collective community consensus regarding the historiography of landscape painting, as it crystallizes from the feedback of qualitative research and visual resource librarianship. We do not claim that our data reflect art historical fact. Instead, they reflect a

consensus of the rhizomatic metanarrative of landscape painting that, however, remains so far invisible except to the connoisseur who is familiar with the corpus as a whole and who by coincidence has been trained using the identical corpus, which of course, is highly unlikely. As such, our study reveals the metanarrative inherent in the chosen dataset to a broad multidisciplinary audience while offering a benchmark or null model for further research, qualitative and quantitative, including research dealing with originals and reception aesthetics.

We choose landscape paintings as the scope of our study for two reasons. First, landscape paintings more often consist of clear horizontal or vertical components compared with other paintings such as portrait, still-life, or abstract paintings. For instance, landscape paintings often have a horizontal boundary between a foreground and a background (or possibly, a middle ground) or vertical frames composed of trees, cliffs, or buildings. Thus, analyzing landscape paintings using the currently available informational partitioning algorithm has advantages of directional simplicity of elements in the composition, which also facilitates the interpretation of results. Second, the colors used in the subregions of landscape paintings are often more distinctly separated than in other genres of painting, resulting in the extraction of larger mutual information. To prove this point, we provide comparative results throughout this paper, looking at landscapes and abstract paintings, using an additional auxiliary large-scale dataset (*SI Appendix*, section I).

Analyzing 14,912 landscape paintings, we specifically address the following questions. Arising from the fact that the historiography of landscape painting is often subdivided by nationality, are there nationally specific or transcendent characteristics in the composition of landscape paintings, such as Dutch flatland vs. Swiss mountains? What are the characteristic frequencies of compositional proportion employed by individual landscape painters in general? How do the characteristic, perhaps dominantly used compositional proportions change over time? Can we use the frequency distributions of compositional proportion as a signature to characterize diverse conceptual groups in art history such as plain time spans of creation dates, conventional style periods, or individual artists' oeuvres?

In the following sections, we show that the distribution of compositional proportion successfully captures distinguishing compositional properties of artists and conventional style periods in art history as recorded in our dataset (i.e., arguably in line with public consensus). We also demonstrate that there exists a systematically changing trend in the compositional proportion distribution of paintings that allows us to unveil temporally clustered structures above the level of the life of artists and transcending the duration of conventional style periods in landscape painting.

Materials and Methods

Datasets. Digital scans of landscape paintings were collected from two major online sources: WikiArt (WA) (13) and the Web Gallery of Art (WGA) (14). WA landscape paintings were collected in October 2018 with a total of 12,431 paintings by 1,071 artists assigned to 61 nationalities. WGA data were collected in May 2016 with a total of 3,610 paintings by 816 artists assigned to 20 nationalities. We note that both websites continuously update the dataset, so a larger volume of landscape paintings is available on the date of publication of this research.

While the overall number of paintings from WGA is relatively smaller than from WA, the WGA dataset has a larger volume of paintings produced before 1800 CE. Therefore, we utilize both datasets in a complementary way. We carefully constructed a unified dataset by filtering out the duplicate paintings from both datasets. During the data preprocessing procedure, we also manually removed unsuitable borders and photo backgrounds from the entire digital reproductions in painting dataset. The final dataset includes 14,912 paintings by 1,476 painters from 61 nationalities (Fig. 1*A*) (*SI Appendix*, section I has a detailed description and the data curation process). To build a consistent nationality standard, the nationality of WGA artists was assigned based on Wikipedia's nationality information. We provide
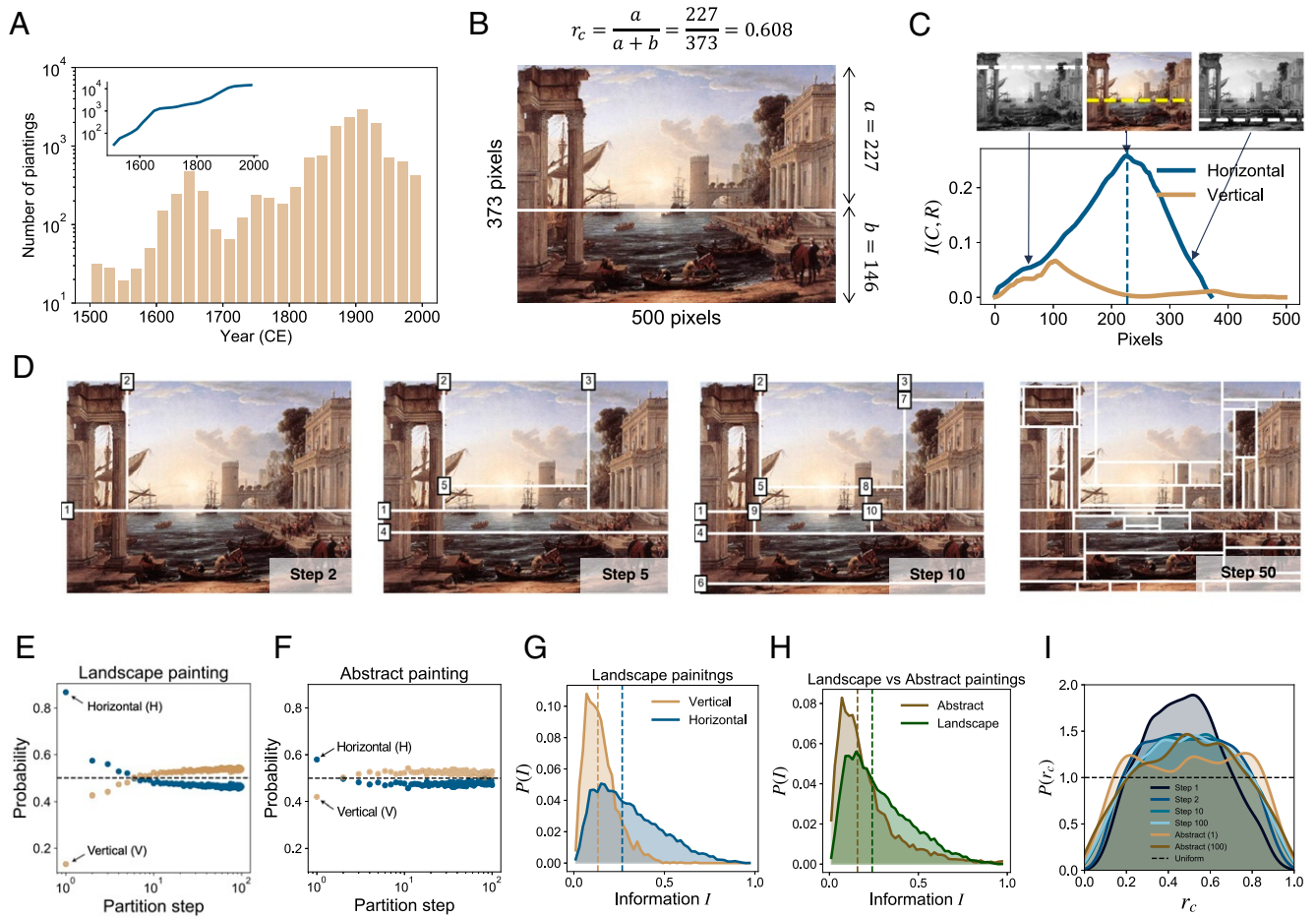
Fig. 1. Partitioning images horizontally and vertically using the dissection algorithm of Rigau et al. (39) based on compositional information $I(C, R)$ focusing on a dataset of 14,912 landscape paintings. (*A*) The number of landscape paintings in the dataset in 20-y bins. *Inset* shows the cumulative number of paintings over time. (*B*) The first bisection of a sample painting (*Seaport with the Embarkation of the Queen of Sheba* by Claude Lorrain dated 1648) selected for maximum compositional information $I(C, R)$. (*C*) The first partition direction and proportion are determined based on maximum $I(C,R)$. In case of Lorrain's painting, the horizontal partition at the 227th line from the top gives the maximum $I(C, R)$. (*D*) The partitioning process of a painting varies depending on the composition of the painting. The partitioning result for the painting is shown for steps 2, 5, 10, and 50. The process is deterministic for a given image. Painting images credit: The National Gallery. (*E* and *F*) Probability of partition direction over partition steps for landscapes (*E*) and abstract paintings (*F*). While our method is restricted to horizontal and vertical dissection, it captures greater arbitrariness in abstract paintings within these plots. (*G*) The probability density distribution of compositional information I at first partitions for different partition direction in landscape paintings. Horizontally partitioned paintings show a wider tail, resulting in larger median information. (*H*) The probability density distribution of compositional information from first partitions for landscapes and abstract paintings. The median compositional information of abstract painting is relatively smaller compared with landscape painting. Together, *G* and *H* indicate that the position of the horizon in landscapes is a meaningful characteristic that needs further investigation. (*I*) The probability density distribution of partition proportion r_c at different partitions steps from 14,912 landscape paintings. Kernel density estimation is used for calculating probability density. The null model based on uniformly random partitions is depicted by a black dashed line, while the distribution for abstract paintings is given in ochre. The plot further indicates that the initial partition in landscape paintings is more dominant, contributing more strongly to the overall composition.

the full list of paintings used in the analysis and corresponding metadata in Dataset S1.

We also collected 5,780 abstract paintings by 175 painters from the WA dataset in a separate auxiliary dataset. The abstract painting dataset includes artworks by abstract painters in the three WA categories: "abstract art," "postpainterly abstraction," and "abstract expressionism." Quantifying this corpus throughout this paper roughly translates to an imaginary art historian summarizing the art history of landscape painting while being exclusively trained on this widely available, broadly used, yet of course also limited selection of paintings. Differences of the intuitions of individual connoisseurs are to be expected. However, we believe the differences may spark the development of better benchmarks than WA and/or WGA.

Data Preparation. The informational partitioning algorithm of Rigau et al. (39) generally works with images of any color depth. However, for the main analyses of our study, the matrix elements of images corresponding to pixels

in RGB colors have been coarse grained into three-bit color (one bit for each RGB value) for the following two reasons. First, an image represented in high color depth with a large number of bits per color component in each single pixel would look more natural and realistic to human vision. However, because the partitioning algorithm we used in this study treats color "indexical" and not gradual, colors with slight differences are considered completely different colors. Therefore, we chose to quantize or coarse grain the painting images to a minimal representation of the color space. Using this method, we also avoid finite-size effects that appear if the pixel size of an image is not sufficiently larger than the number of distinct colors in the respective color system when calculating the compositional information and partitioning position. To minimize this finite-size effect, we choose the minimum number of colors that can represent an image in RGB color space while still capturing the macroscopic shape within the painting. To validate our results, we also checked other color depths such as eight-bit RGB, gray scale, or different thresholds to create three-bit images arriving at similar and robust results. *SI Appendix*, section III has detailed explanations on the

effect of color depth on the partitioning process and a comparison between results obtained from different color-depth systems.

Image Partitioning Using Compositional Information. We apply the image partitioning algorithm of Rigau et al. (39) based on mutual information of color palette and subregions in an image. Each image in the dataset was first converted into a matrix representation whose dimensions correspond to image width and height.

The full image is considered an initial partition, and the algorithm progressively subdivides the image according to the partitions that provide maximum information gain at each step. For a random variable C taken from the set of discrete colors in RGB color space, the color palette information of an image is defined as Shannon entropy $H(C)$:

$$H(C) = -\sum_{c \in C} P(c) \log_2 P(c). \tag{1}$$

The probability $P(c)$ is given by $P(c) = S_c/S$, where S_c denotes the number of pixels taking color c and S is the size of the image.

Considering a painting process as a mapping from the color set C to the set R composed of a finite number of regions in an image, the conditional entropy $H(C|R)$ is defined as

$$H(C|R) = -\sum_{c \in C, r \in R} P(c, r) \log_2 P(c|r), \tag{2}$$

where the joint distribution $P(c, r) = Pr[C = c, R = r]$ and the conditional probability $P(c|r) = Pr[C = c|R = r]$. The compositional information gained from introducing a set of partition R is provided by the compositional information defined as the mutual information:

$$I(C, R) = H(C) - H(C|R). \tag{3}$$

If the image is decomposed into n regions, $R = r_1, r_2, \ldots, r_n$, the compositional information is given by the generalized Jensen–Shannon divergence (JSD):

$$I(C, R) = \text{JSD}(C_1, C_2, \ldots, C_n)$$
$$\equiv H(\sum_{i=1}^{n} \pi_i C_i) - \sum_{i=1}^{n} \pi_i H(C_i, r_i)' \tag{4}$$

where π_i is the size of region r_i normalized by the full size S and the regional Shannon entropy $H(C_i, r_i)$ for the color set C_i in the region r_i is given by

$$H(C_i, r_i) = -\sum_{c \in C_i} P(c|r_i) \log_2 P(c|r_i). \tag{5}$$

Therefore, given a partition that dissects an image into two regions, $I(C, R)$ gained from the partitioning is maximized if the dissected subregions are composed of a completely distinct color to other subregions. On the other hand, if the dissected subregions both have same color frequency distribution with the original image, the partitioning offers no meaningful information [$I(C, R) = 0$].

The procedure to find partitioning positions of an image is as follows. During the first partition process, one should calculate the compositional information gain over all possible partitions in both horizontal and vertical directions on the entire image resulting in $w - 1 \times h - 1$ trials, where w and h are the numbers of pixels of weight and height, respectively, of the image. Then, the algorithm selects a partition that gives the maximum information. From the second partitioning process, the algorithm is repeatedly applied to remaining subblocks, and the subblocks are partitioned with lines that offer maximum information for each step. In principle, the partitioning process can be continued until the image is fully decomposed into regions of homogeneous colors. Because the conventional partitioning process takes considerable time to find partitioning positions for a large dataset, we used the Shin et al. (40) line updating bipartitioning (LUB) algorithm to calculate the mutual information for each partition step for faster calculation (detailed descriptions of the LUB algorithm are described in *SI Appendix,* section II).

In this study, we focus on the first two partitions of painting images to analyze compositional proportion of a painting. This allows for meaningful interpretation as the first two sections classify the images in relation to the most dominant compositional features of the painting. After an image or a subblock of the image is partitioned into two subregions with ratios $a : b$, we define the compositional proportion r_c of the bipartitioned image as

$$r_c = \frac{a}{a + b}. \tag{6}$$

Here, a is the height (width) of the first subregion from the top (left), b is the height (width) of the lower (right) subregion for a horizontally (vertically) partitioned image, and the proportion is normalized to (0, 1). The closer the r_c is to one, the more the painting is divided at the bottom (right).

Results

Fig. 1 B and C shows an example of how the partitioning process takes place in a sample painting by Claude Lorrain (1620 to 1650) (the three-bit and other color-depth images in which the actual algorithm works are shown in *SI Appendix,* Fig. S7). The original image is considered an initial region, and the algorithm progressively subdivides the image according to the partitions that provide maximum information gain at each step as described in *Materials* and *Methods*. In the partitioning process, a painting image is split into quasihomogeneous regions via either a horizontal or a vertical line. Lorrain's painting is found to have larger compositional information in horizontal partitions than vertical partitions (Fig. 1C). Therefore, the painting is first partitioned in the horizontal direction at the position that gives the maximum information. Then, the compositional proportion r_c is determined by the normalized partitioning position given by Eq. **6**. In case of the sample painting, the partitioning takes place at the 227th pixel from the top, while the height of the original image is 373. Therefore, r_c is calculated as $227/373 \approx 0.608$.

Landscape Paintings Are Well Characterized by Dominant Partitions. In principle, an image can be partitioned repeatedly until the image is fully divided into subblocks of completely homogeneous colors. As illustrated in Fig. 1D, a subblock of an image can be partitioned either vertically or horizontally at any partition proportion r_c in each step. However, we find that early partitions in landscape painting show a distinct characteristic compared with later steps.

First, the partition direction of landscape paintings in the early partition steps is mostly found to be horizontal in direction (Fig. 1E), while there was no directional preference in abstract paintings in the first step (Fig. 1F); 86.8% of landscape paintings were horizontally partitioned at the first partition step with larger compositional information than vertically partitioned cases on average (Fig. 1G). The frequent observation of a dominant horizontal partition is mainly due to the fact that landscape paintings usually include a horizon, where constituting colors begin to vary significantly. After around 10 partition steps, the ratio of landscape paintings that are partitioned in the horizontal direction saturated at the point slightly below 0.5. However, abstract paintings did not show any directional preference in composition over all partition steps including particularly the first step. It is interesting that the ratio of horizontal vs. vertical partitions saturates slightly below 0.5 in both painting genres (Fig. 1 E and F). We speculate that this saturation effect is mainly caused by two reasons. First, more images exist in landscape aspect ratio as opposed to portrait aspect ratio, notwithstanding the genre of landscapes or abstract subjects. This means that simply by chance more should be partitioned in a vertical direction. Another reason, especially for landscape paintings, is that whereas large-scale objects in landscape paintings such as sky, earth, and ocean are horizontally placed, there are typically smaller vertical objects often with edges left and right such as trees, plants, and buildings. Consequently, the frequent existence of vertically oriented objects, particularly in the foreground, could cause more frequent partition in the vertical direction in the later partition steps. We provide more detailed verification regarding the two explanations in *SI Appendix,* section III.

Landscape painting further exhibits distinct characteristics in the amount of compositional information. The compositional information gained from the first partitions in landscape painting

is much larger than that of modern abstract paintings (Fig. 1*H*). This indicates that the initially partitioned subregions in landscape paintings tend to be composed of more distinct colors than other subregions. In contrast, the abstract paintings have relatively less quantifiable information (Fig. 1*H*). For example, consider the famous abstract–expressionist paintings of Jackson Pollock (1912 to 1956). Pollock's drip paintings were painted by pouring and splashing paint onto large canvases rolled out across the floor, resulting in a fractal pattern without larger-scale compositional dissections of the canvas. Since the composing colors are distributed relatively randomly in space, the subregions of Pollock's drip paintings tend to have similar color distributions in the early partition steps compared with the color distribution of the original image as a whole, resulting in little compositional information with regard to meaningful horizontal or vertical dissection.

The distributions of partition proportion r_c also differ over partition steps and genres. Fig. 1*I* shows the distribution r_c from different partition steps for the total set of landscape paintings. In the first partition step, the distribution of r_c has a relatively sharp peak indicating that artists have employed certain proportions more than other proportions when composing their paintings. As the number of partition steps grows, the distribution of r_c gets wider and becomes flatter. On the contrary, the partitioning position of abstract paintings follows an almost uniform distribution even from the first partition step, implying no typical structure at the macroscopic scale (Fig. 1*I*).

Typical Compositions of Landscape Paintings Based on First and Second Partitions Gradually Change over Time While Transcending Concepts of "Nationality." In this section, we investigate frequency distributions of the dominant composition and compositional proportion in landscape paintings over the entire period of our dataset. First, as we found that the early partitions are a meaningful characteristic in landscape paintings, we study their composing direction and the joint probability distributions of compositional proportion r_c up to the second partition. Partitioning up to the second partition classifies paintings into one of four groups based on the order of two partition directions: horizontal–horizontal (H–H), horizontal–vertical (H–V), vertical–horizontal (V–H), and vertical–vertical (V–V), where H and V stand for a horizontal partition and a vertical partition, respectively. While abstract paintings have a symmetric directional preference on average, landscape paintings show a strong asymmetry in the partitioning direction (Fig. 2*A*). The proportion of the H–H-type paintings is significantly higher followed by H–V-, V–H-, and V–V-type paintings: H–H > H–V > V–H > V–V.

However, after we explore the temporal change of the distribution of the partition types, a gradual change in composition selected by landscape painters is observed (Fig. 2*B*). Since the 1600s, H–V-type composition was the most dominant, but its proportion continued to decrease, while the number of H–H-type works gradually increased. The cross-over is found during the midnineteenth century. On the contrary, the fractions of V–H- and V–V-type compositions were kept stable over time.

In order to check whether the global difference over time is attributed to the influence of dominant nationalities in local time periods, we examined the distribution of the composition types over time for each country. Interestingly, the time-varying nature of the composition in the landscape paintings was also observed among individual nationalities. Fig. 2*C* shows the distribution of composition types of the dominant nationalities over five time periods, where each time bin was set to include the same number of paintings. Here, we choose the dominant nationalities of each period as those that included more than 30 artworks in each period, but we note that the results were robust for other criteria from 20 to 60. Overall, our results suggest that typical compo-

sitions in landscape painting differ significantly based on the partition types and exhibit a gradual temporal evolution while transcending concepts of nationality. This also suggests that most of the difference in the use of composition across nationality is possibly rooted in the global temporal structure of the data, while apparent differences in nationality are simply rooted in changes of localized production and/or art historical selection (*SI Appendix*, Fig. S3 shows the temporal distribution of the datasets across nationalities).

To explore typical compositional proportion for each of the four pairs of partition direction, we plot the four types of joint distributions of partition ratio r_c accompanied by typical sample paintings found near the peaks of the distributions (Fig. 2 *D–G*). In the H–H-type paintings, the proportion of first horizontal r_c was largest near the 0.40 position from the top. In the next partition step, the second horizontal r_c is found near the 0.47 position from the top in the lower subregion as shown in the sample painting (Fig. 2*D*). As such, the H–H-type corresponds to the composition of paintings with three horizontal layers, usually depicting a foreground, a middle ground, and a background, as suitable to represent perspective. H–V- and V–H-type paintings have dominant vertical features in the composition therefore balancing both horizontal and vertical elements. In H–V-type paintings, the first partition was found most frequently around the 0.55 position from the top (Fig. 2*E*). Therefore, the first division cuts slightly down from the middle. However, unlike in H–H-type paintings, the second vertical split in H–V paintings could occur evenly over wide range of ratios. V–H-type paintings most frequently have a first vertical partition position near 0.5 from the left and a second horizontal partition at 0.6 from the top (Fig. 2*F*). Finally, V–V-type paintings constitute a distinct symmetric shape in the joint probability distribution indicating that the first and second partition can occur equally on any side (left or right) of the painting (Fig. 2*G*). Overall, the dissection algorithm captures the dominant dissections in painting composition reasonably well. This is surprising as the algorithm obviously has some limitations, as it cannot capture dominant diagonal elements or curved structures of a painting, whereas a human observer corresponding to a more flexible algorithm would certainly do. Similarly for the sample images in Fig. 2 *F* and G, our algorithm does not capture the exact center position of the tree trunk, which human experts would perhaps naturally consider as an intuitive position to dissect the image frame. Here, it is important to note that our algorithm strictly splits areas of positions at maximum color difference, not at maximum symmetry. This effect is similar to Voronoi vs. Delaunay tessellation and of course, a possible subject of future research (41).

The Dominant Horizon Evolves Systematically in Landscape Art History. So far, we have explored characteristics of the first two dominant proportions r_c in our full time-aggregated set of landscape paintings. Then, how has the use of compositional proportion in landscape paintings changed over time? We investigate the change of the distribution of r_c in three different types of art historical concepts: discrete creation date time spans of 20 y, conventional style periods, and the oeuvre of individual artists, all of these as found in the original source datasets (*SI Appendix*, section I has detailed statistics). While we are aware that all of these concepts of date, style attribution, and authorship are more or less agreed on conventions, often subject to ongoing scholarly discussion, we simply treat them as a priori groupings of paintings. We focus on the great majority of Western landscape paintings that are partitioned in the horizontal direction in the first step (86.8%) from 1500 to 2000 CE and their corresponding dissection ratio r_c in the following analyses to allow for clear and intuitive understanding. Essentially, with exceptions, we are looking at the position of the horizon over time (i.e., the physical
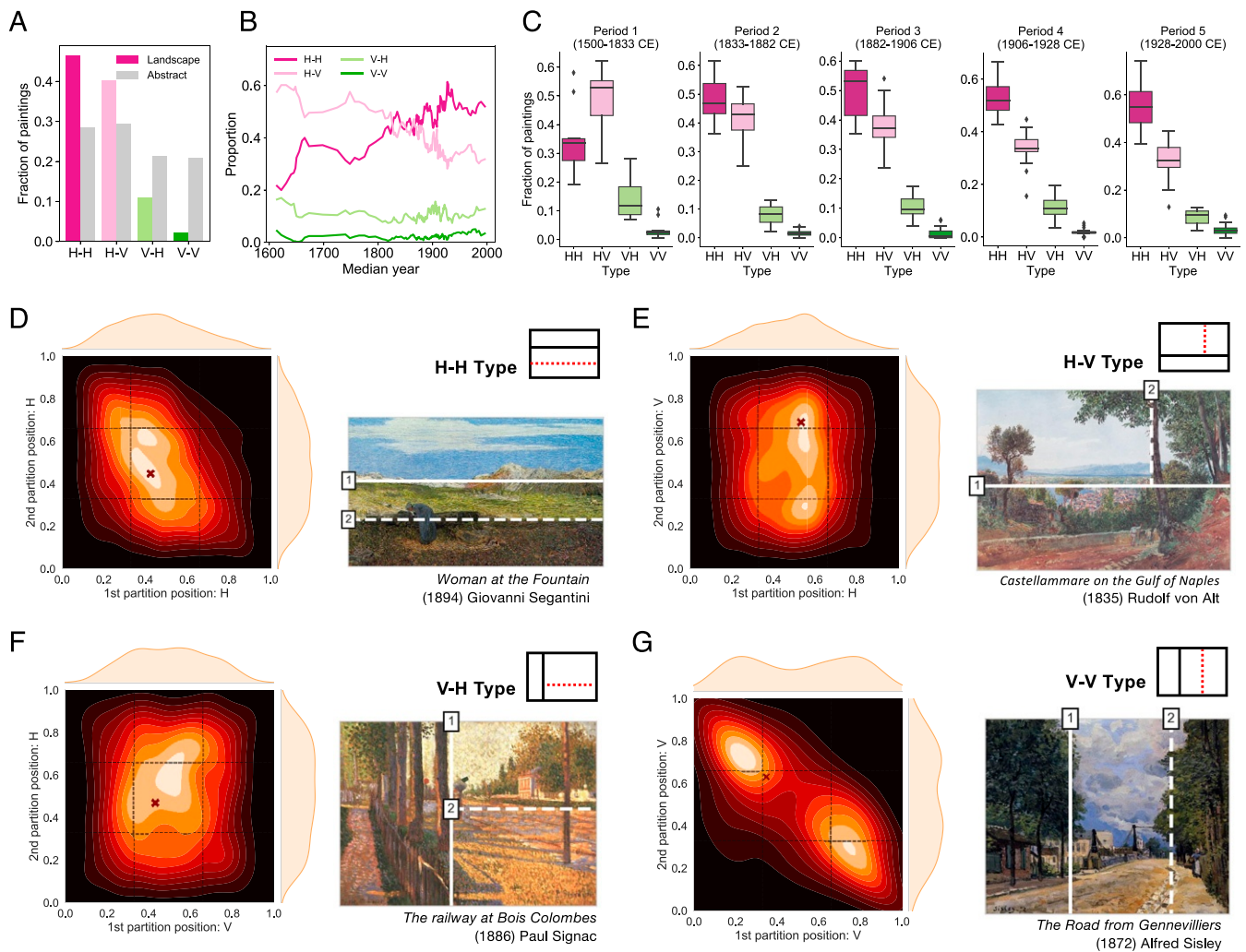
Fig. 2. Landscape painting partition preference in horizontal vs. vertical direction up to the second most dominant dissection. (*A*) For the time-aggregated dataset, the fraction of H–H-type paintings is highest, followed by H–V-, V–H-, and V–V-type paintings: H–H > H–V > V–H > V–V. Colored and gray bars indicate the proportions of partition types for landscape paintings and abstract paintings, respectively. (*B*) Gradual temporal evolution in the use of the four composition types. A moving time average was applied, where each time bin was set to include the same number of paintings (500). The fraction of H–V-type paintings gradually decreases, whereas the fraction of H–H type gradually increases to become dominant since the midnineteenth century. The fractions of V–H- and V–V-type paintings are stable over time. (*C*) The box plots summarize the distribution of fractions of each partition type across different nationalities in five consecutive time periods. The five time bins were set to include equal numbers of paintings. Behaviors of individual nationalities follow a similar temporal trend to the aggregate result (*A* and *B*). Distributions for individual nationalities are shown in *SI Appendix*, Fig. S14 in detail. (*D–G*) Four types of joint distributions of partition ratio r_c emerge from the dataset, here accompanied by typical paintings sampled near the peaks of the distributions for (*D*) H–H-, (*E*) H–V-, (*F*) V–H-, and (*G*) V–V-type paintings. The first and second partitions of sample paintings are represented by solid and dashed lines, respectively. Painting images credit: WikiArt.

horizon between earth and sky) or the most dominant chromatic difference in the associated color gradient.

Fig. 3 exhibits the change of the distribution of r_c measured and encapsulated in three different conceptual categories: 1) discrete creation date time spans (20-y bins from 1500 to 2000 CE), 2) conventional artistic style periods, and 3) the oeuvre of individual artists. We analyze the top 25 and arguably least contested conventional style periods with the largest number of artworks in the dataset for the rest of our study. These conventional style periods compose 92.8% of all artworks in the dataset. Regarding individual oeuvres, we focus on 131 painters, all within the top 10% based on the number of paintings in our dataset, in two groups before and after 1800 CE, resulting in 30 and 101 painters from the two periods, respectively. We introduce this separate selection criterion as the number of recorded paintings by painter largely increases from 1800 CE. Applying a uniform selection criterion across all time would possibly cause a selec-

tion bias toward the modern era (*SI Appendix*, section V has the statistics). In Fig. 3*C*, we exhibit a random sample of 66 (50%) individual painters from the 131 representative painters further restricted for the purpose of visualization. The distributions in all three conceptual categories are sorted and colored by the median year in each group.

In all three cases, the distributions of r_c are mostly unimodal and varying smoothly over art historical time, indicating that there are characteristically dominant compositional proportions used by landscape artists (equivalent to the peak of the respective distribution), while the distribution gradually changes over time instead of showing a singular or a random preference in peak shift. The change of peaks over time periods in Fig. 3*A* shows three stages. Initially, in the midsixteenth century, a small r_c around one-third was dominant. However, the peak r_c gradually increased from the late sixteenth century and reached its highest value from 1640 to 1680 CE. This high peak r_c was
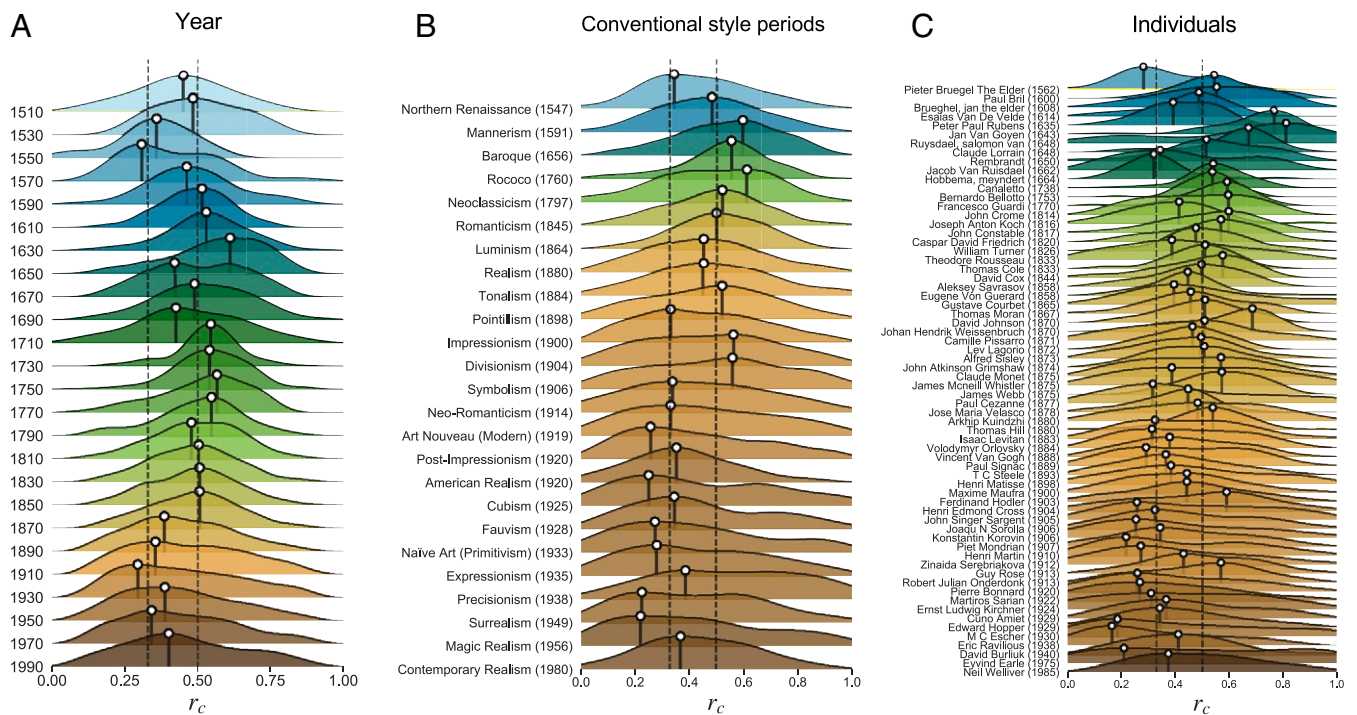
Fig. 3. A gradual systematic evolution of Western landscape painting is implied by the frequency distributions of partition ratio r_c for 20-y bins, conventional style periods, and individual artist oeuvres. White dots represent the peaks of the distributions. The same colors of the distributions were used to indicate a similar median production year of each group. Distributions of r_c (*A*) for discretely binned 20-y time windows from 1500 to 2000 CE, (*B*) for 25 different conventional style periods as assigned by curators in the original datasets, and (*C*) for 66 (50%) painters randomly sampled from the top 10% of representative individuals with the largest numbers of artworks in the dataset. Distributions in *B* and *C* are sorted by the median production time of the artworks. Two dashed lines from left to right indicate one-third and one-half, respectively. The implied gradual evolution of landscape painting may, of course, emerge from historical fact or from a systematic selection bias of art historical scholarship constituting the dataset. Nevertheless, its quantification is important as the gradual metanarrative can function as a null model that can serve as a point of reference to be confirmed and critiqued by further scholarship.

maintained until the midnineteenth century, when the peak r_c gradually decreased toward near the one-third position until the late nineteenth century. This indicates that in the sixteenth century, the dominant horizontal partitioning position was usually above the middle of the canvas, gradually moving to the lower middle until the late seventeenth century. Artists continued to frequently produce high r_c paintings (low horizon) until the midnineteenth century when the dominant dissection position on the canvas starts to move upward again.

The trend of r_c essentially recapitulates the historiography of landscape painting yet is still surprising as it appears systematic at scale beyond anecdotal examples. The overall small values of r_c in the sixteenth century intuitively fit artists who started to produce "pure" landscape paintings, such as Joachim Patinir in the Netherlands, who produced a style of panoramic landscapes with small figures from a high aerial view point known as "world landscape." A similar composition was also created by Pieter Bruegel the Elder. The high aerial view with a large portion of ground results in images partitioned by a high horizon. From the seventeenth century, paintings begin to exhibit increasingly high r_c. This trend is exemplified by a group of Dutch landscape painters, such as Jan van Goyen (1595 to 1656) and Salomon van Ruysdael (1602 to 1670), as these Dutch painters depicted landscapes where the sky occupies a larger portion of the canvas (e.g., the median r_c for van Goyen's paintings is 0.69). The trend of high r_c continued up to the early and midnineteenth century, where painters such as John Constable (1776 to 1837), Caspar David Friedrich (1774 to 1840), and William Turner (1775 to 1851) also expressed their mastery with a focus on the sky. From the midnineteenth century, the peak proportion starts to

decrease and moves toward the one-third position, implying that the painting is partitioned around the upper third line of the canvas (Fig. 3 *A* and *B*). A possible interpretation is that a relatively small number of paintings utilized a large portion of the canvas to depict the sky. We note that the distribution of r_c also becomes broader in the nineteenth century, indicating a more diverse use of compositional proportion.

The midnineteenth and twentieth centuries are commonly assumed to be a period when a diversity of conventional artistic styles flourished, leading to supposedly completely different artistic styles coexisting in parallel. Grouping paintings by these conventional style periods, we can see if the common assumption of stylistic difference can be put into question by measurement (Fig. 3*B*) (*SI Appendix*, section I has the percentage of each style period in the dataset). Overall, the trend of the peaks and distributions follows a similar pattern to Fig. 3*A*. Focusing on the most well-known conventional style periods, we find that baroque painters frequently used large r_c (peak at $r_c \approx 0.57$), while in the rococo and Romantic periods, we commonly find a proportion near one-half (peaks at $r_c \approx 0.54$ and $r_c \approx 0.52$, respectively). In realism, impressionism, and later styles, the dominant proportion then started to decline, reaching a position near one-third in the upper half of the canvas. Postimpressionism and expressionism are examples with a peak r_c close to one-third (Fig. 3*B*). An interesting observation is that despite the well-known high diversity in painting styles and topics in the modern era, the distributions of r_c of various artistic styles in the twentieth century are similar with the peak r_c located near the one-third position from the top. In other words, landscapes in different twentieth century "styles" or "isms" seem to share a common mode of

construction. We further study the distributions of r_c of individual painters. Fig. 3C shows the distributions of compositional proportions for the random 50% sample of individual artists with the largest numbers of paintings in our dataset. While naturally subject to larger fluctuations due to the smaller number of images per artist, the peaks for individual artists still demonstrate a similar trend in r_c. Consequently, we are able to investigate temporal clusters of individuals in detail based on similar use of compositional proportions in the following section.

The Similarity Network of Individual Artists and Conventional Styles Is Subject to Meaningful Clustering and Confusion. The fact that each compositional proportion distribution for time periods, artists, and style concepts has distinguishing characteristics naturally suggests consideration of clustering between the distributions to characterize separation and confusion. Which artists and styles are similarly characterized with regard to their typical r_c distribution? Which are far apart from each other? Can we find communities of artists and styles sharing a similar use of dominant compositional proportion? To answer these questions, we compare the distributions of r_c within and between individual artist oeuvres and styles periods. Doing so, we effectively treat both "artists" and "style attributions" as groups of artworks, much like they may form latently in the mind of an art historian or a trained neural network. We define a proportion similarity measure J between two distributions of r_c between P_i and P_j using an information-theoretic distance, the JSD:

$$
\begin{aligned}
J &= 1 - \mathrm{JSD}(P_i, P_j) \\
&= 1 - [H(\frac{P_i + P_j}{2}) - \frac{1}{2}(H(P_i) + H(P_j))],
\end{aligned}
$$
[7]

where H is the Shannon entropy with base 2 and the proportion similarity J is bounded between zero and one.

We investigate how individual artists cross-correlate with multiple styles in terms of horizon choice. We first create a bipartite or two-mode network composed of landscape artist oeuvres and conventional style periods, respectively, to detect groups of artists and style periods with similar habits of composition in paintings. We measure the distributions of r_c of the same 131 representative individual painters and 25 conventional style periods as used in the previous subsection. Then, we set the weight of links between pairs of artists and styles by the proportion similarity between their r_c distributions. From the originally fully connected bipartite weighted network, we then compute a significance criterion for each link following the procedure described in ref. 42. We only keep significant links at the level of $\alpha = 0.366$, which maximizes $N_b/N_0 - L_b/L_0$, where N_b (L_b) and N_0 (L_0) are the numbers of nodes (links) in the filtered network and the original network, respectively. Then, we apply the Louvain community detection algorithm (43, 44) on the filtered network, which is a fast hierarchical community detection algorithm widely used in network science, to find hidden community structure.

In the final similarity matrix in Fig. 4A, we observe that the compositional similarity network has a clear community structure with three groups of artists and style periods based on their distribution of using particular compositional proportion. Within each community, the link density is high, and the weights of links tend to be much larger than between communities off the diagonal. Each community has two distinguishing features from the others (Fig. 4B). The first feature is the shape of the aggregated r_c distribution of paintings in each community. Fig. 4B shows the distribution of r_c and production year of paintings by all artists in each community. The distribution of dissection proportion r_c in each community has clear difference from the others. Community 1 has the largest peak dissection position (peak $r_c = 0.531$), and the distribution is sharper with smaller

SD ($\sigma = 0.15$). Community 3 has smaller peak position (peak $r_c = 0.303$), and the distribution is much wider ($\sigma = 0.21$). Community 2 has intermediate characteristics between communities 1 and 3 (peak $r_c = 0.393$, $\sigma = 0.17$). Another remarkable feature is that without including any explicit metadata regarding information on the activity time of artists or their stylistic classification, the artists and styles are meaningfully grouped in terms of time period (Fig. 4 A and B). At a minimum, this means the metadata are coherent with the visual features of the images, pointing to a coherent classification by individual qualitative art historians. Of course, critical eyes may surmise that the emerging metanarrative seems almost too coherent to be true. This would mean the classification is coherent yet leaves out everything that does not fit into the all to smooth story of Western art.

Looking into details, community 1 in Fig. 4A is characterized by high r_c paintings and contains most of the artists before 1850 CE. Conventional styles up to the midnineteenth centuries in the dataset fall into this group. Community 2 exhibits a peak at $r_c = 0.393$, which is above the middle of the canvas and contains many artists who worked during the late nineteenth century and some in the early twentieth century. Realism, impressionism, and related painters characterize this group in an exemplary way. As we have seen in Fig. 3, from the periods of realism and impressionism onward, the dominant partitioning proportion began to decrease. As such, this period bridges the classically preferred composition to the newly preferred composition in the modern period. This result is in line with the established studies that realism and impressionism served as the transition toward the modern era. Lastly, artists in community 3 show a distribution of r_c with the smallest modal value among the three groups (Fig. 4B) with a less pronounced peak r_c found at the near–one-third position ($r_c = 0.303$). Looking at temporal distribution, the paintings in community 3 are mainly concentrated in the early twentieth century, with some in the late nineteenth century. The characteristic shape of the r_c distribution is broader than in the other two communities, representing a more diversified use of proportional dissection. However, contrary to the common notion that the styles or isms of the modern era are distinct stylistic expressions, their distributions of compositional proportions are similar and clustered into a same group: the peak r_c of aggregated paintings (Fig. 4B, third column) and each style's peaks (Fig. 3B) were all found near $r_c = 1/3$. In other words, "anything goes" not only across modern isms but also, within each single ism, while "anything" is bounded with a bias to $r_c = 1/3$. As such, it is not individual isms that correspond to conventional style periods such as renaissance or baroque. Instead, one could argue that the sum of isms is the actual style period of isms, which all behave the same with respect to proportional dissection while differing in more limited ways only, particularly in their textual phenotype. One could say all isms share the same generating function regarding the dissection of landscapes.

We also investigate the proportion similarity matrices between all pairs of artists and conventional style periods (Fig. 4 C and D). To compare the clustering of artists and style periods within each concept with the joint artist–style clusters, we newly grouped individual artists and style periods into three communities using the same community detection algorithm. Clustering of artists and styles (Fig. 4 C and D, respectively) correlates meaningfully with the artist–style clusters, again exhibiting clear clustering structure in time. At the same time, the pure artist and pure style clusters deviate from the joint artist–style clusters (Fig. 4A) in detail and consequently, also from the respective one-mode projections (*SI Appendix*, Figs. S19 and S20), which means we can identify exceptions to the respective superera mainstream in individual artist oeuvres and style concepts.

This effect is in line with Fig. 3 B and C, where the systematic evolution of r_c appears more clear for conventional style
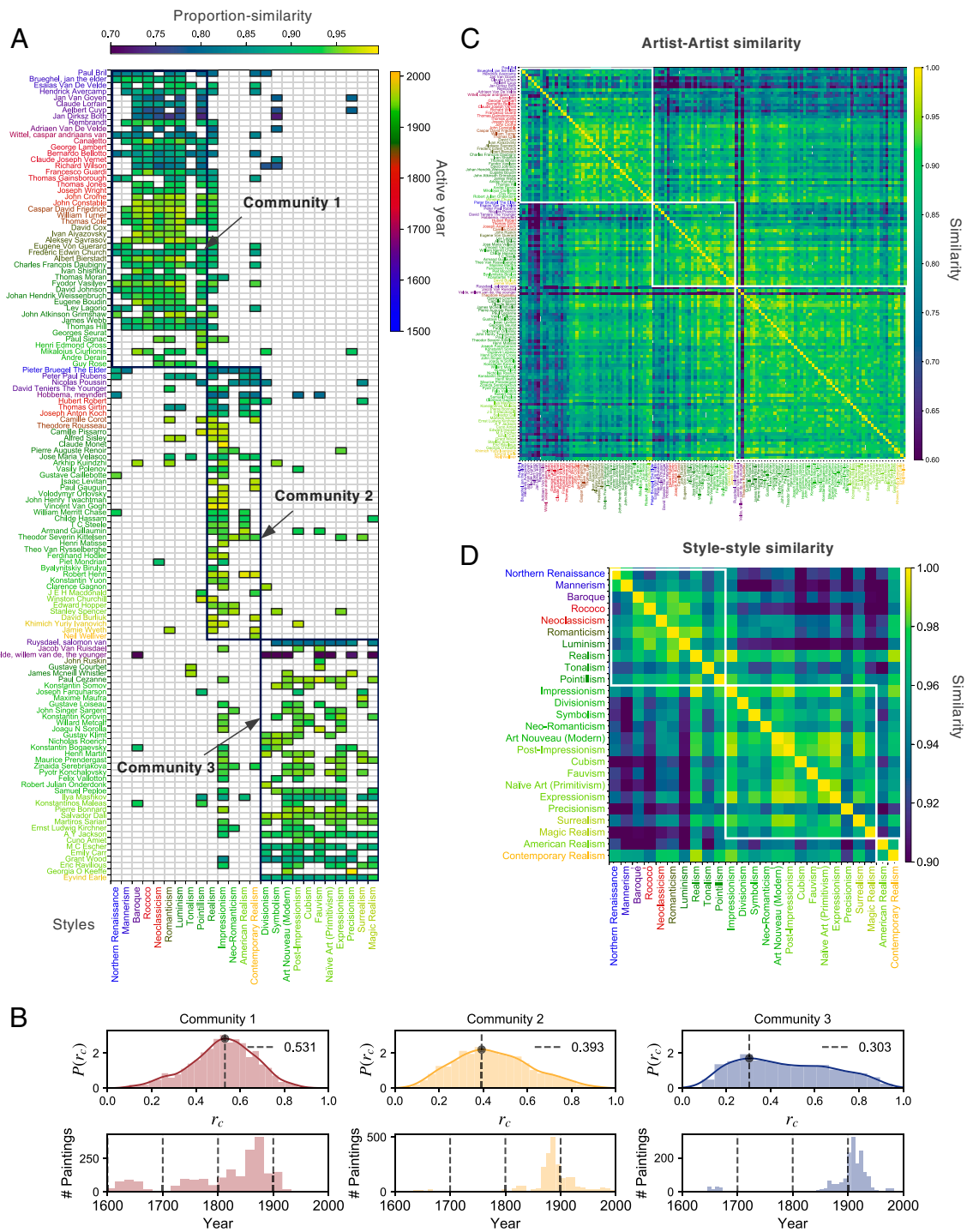
Fig. 4. Clustering structure in three similarity networks of dominant proportion: the bipartite network of individual artist oeuvres and conventional style periods, the artist–artist network, and the style–style network. (*A*) The matrix elements represent the proportion similarity between individual artists and style periods. Statistically significant links based on the disparity filter (42) are kept in the bipartite network and are colored in the matrix. Boxes indicate communities found by the Louvain community detection algorithm (43). Matrix elements within each community are sorted by the median production year of landscape paintings within the individual artist oeuvre or style period. Median years are also reflected in the colors of the row and column labels. Although the clustering algorithm does not use any information regarding implicit time, the artists and style periods tend to be clustered in time, which suggests the existence of supergroups of artists and styles rooted in the evolution of horizontal proportional dissection in landscape painting. (*B*) The distributions of r_c and production dates of paintings by all artists in each artist–style community in *A*. The distribution of dissection proportion r_c differs clearly across communities. The peak dissection position is largest in community 1 (peak $r_c = 0.531$), and the distribution is sharper with smaller SD ($\sigma = 0.15$), whereas community 3 has a smaller peak (peak $r_c = 0.304$) and the distribution is much wider ($\sigma = 0.21$). The distribution of production dates for all paintings in each community also differs from another, implying that the individual artists are also grouped in time. (*C* and *D*) Proportion similarity matrix between all pairs of individual artists (*C*) and style periods (*D*). The tendency that community elements are clustered in time in the bipartite cases (*A*) is still preserved in both similarity matrices as reflected in the colors of the row and column labels. Enlarged versions of *A*, *C*, and *D* can be found in *SI Appendix*, section V.

periods than for individual artists due to the existing heterogeneity on the level of individual. This result points to the fact that the grand narrative of Western art history, as expressed in the coarse-grained conventional style periods, is too smooth and somewhat artificial, failing to capture the rich nonlinearity of art production throughout time on an individual level. Put in other words, our quantification indicates that a proper grand narrative of art history requires a multiplicity of perspectives, both qualitative and quantitative, as a great amount of nonlinear detail gets lost in an overly conventional mainstream narrative.

Discussion

For millennia, artists have discussed and developed various compositional techniques to represent their own creative ideas and purposes. The results explored in this study provide a macroscopic picture of how landscape artists have used proportional dissection in their compositions according to the latent consensus of mostly Western art historiography, as expressed in the widely available and broadly used WA and WGA datasets. Analyzing 14,912 landscape paintings covering more than 500 y, we have retraced the evolution of compositional proportion in landscape paintings.

First, our results show that the dominant modes of landscape composition based on partition direction take gradual changes over time while transcending nationality. Before the midnineteenth century, the vast majority of paintings were frequently characterized by H–V-type composition followed by H–H, V–H, and V–V types. However, the use of H–V type continuously decreased, whereas composition with double horizontal partitions (H–H type) gradually became the most frequent composition type. Meanwhile, the fractions of V–H and V–V types were kept almost stable over entire period. Because of the fact that this pattern transcends national groups, the long-standing division of landscape art history by nation must be put into question, in favor of a broader comparative perspective. Concretely, our result confirms the stance of scholarly literature that has not emphasized national concepts of landscape painting for decades, while suggesting that library classification, metadata in visual resource collections, and the categorization of encyclopedia entries should follow suit.

Second, the dominant proportions of paintings characterized by the distribution of r_c gradually varied over time in a suspiciously smooth way, which could be rooted in art history itself or more likely, artificially in the selection bias of dataset curators or of course, the underlying art historical literature. Reaching the highest modal ratio r_c during the seventeenth century, including the baroque period, the favored ratio r_c began to decrease gradually over time and moved toward the ear one-third position from the top of the painting. As qualitative art historians in recent decades have progressed over the notion of a smooth consensus story of Western art and have emphasized the anecdotally evident multiplicity in the evolution of artistic expression, this study serves as a reminder that the available large-scale datasets might be rooted in older literature, likely perpetuating potentially severe biases.

Third, beyond this gradual change in compositional proportion, we employed network analyses to detect hidden community structure in the proportion similarity network between groups of images representing the oeuvre of individual artists and conventional style periods. The analyses revealed that artists in the same community of the proportion similarity network were actually grouped into temporally similar periods even though no time information from the metadata was included in the analysis. In other words, the dissection profiles of artist oeuvres and style periods do latently encode date information, again either rooted in art history or at least the consensus story of art in which our dataset is rooted.

Fourth, the distributions of compositional proportions of diverse styles or isms of the modern era were found to be similar and clustered into the same group, which counters the common intuition that the diverse isms function much like style periods.

Possible limitations of our study and future work include the following. Although the dataset used in this study includes some Japanese and Chinese landscape paintings, our dataset mainly focuses on paintings by European artists (i.e., the so-called canon of Western European art). As such, the dataset represents a conventional notion of art history that is biased not only geographically but also, toward artists of male gender and very likely streamlined in relation to a more complex reality. That said, our study is nevertheless foundational, as it reveals the streamlined nature of canonic art history, which is a valuable and necessary step preceding all future research that is interested in putting this biased canon into question. We also believe our study can function as a proper starting point to quantitatively investigate principles of composition covering broader conceptual groupings (beyond artists and conventional style periods) and a broader set of cultures and regions. Our methodology is readily applicable to paintings and any other two-dimensional representational forms after an appropriate dataset is prepared. In further analyses, comparing the characteristics of photographs as uploaded into online social media, randomly selected from street view panoramas, or procedurally generated could be intriguing avenues of study (45). An algorithmic limitation of our current partitioning methodology is that it only considers horizontal and vertical divisions. Some artists, for example, might dominantly apply diagonal composition in painting. Therefore, developing algorithms that can detect more complex compositions would be a challenging but likely rewarding issue. Our study mainly focused on the trend in the first horizontal dissection within paintings. In addition, one could explore patterns of composition based on higher-order dissections. Going beyond painting, we believe that our framework can also find use in understanding composition and geometric proportions in other art forms, including in particular typography, film, architecture, etc.

1. H. Wölfflin, *Principles of Art History* (Courier Corporation, 2012).
2. A. Riegl, *Historical Grammar of the Visual Arts* (Zone Book, New York, NY, 2004).
3. A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, M. Mazzone, "The shape of art history in the eyes of the machine" in *Thirty-Second AAAI Conference on Artificial Intelligence* (AAAI, 2018).
4. H. Y. Sigaki, M. Perc, H. V. Ribeiro, History of art paintings through the lens of entropy and complexity. *Proc. Natl. Acad. Sci. U.S.A.* 115, E8585–E8594 (2018).
5. M. Yazdani, J. Chow, L. Manovich, Quantifying the development of user-generated art during 2001–2010. *PLoS One* 12, e0175350 (2017).
6. D. Kim, S. W. Son, H. Jeong, Large-scale quantitative analysis of painting arts. *Sci. Rep.* 4, 7370 (2014).
7. B. Lee, D. Kim, H. Jeong, S. Sun, J. Park, Heterogeneity in chromatic distance in images and characterization of massive painting data set. *PLoS One* 13, e0204430 (2018).
8. R. Arnheim, *Art and Visual Perception* (University of California Press, 1974).
9. E. B. Feldman, *Varieties of Visual Experience; Art as Image and Idea* (Prentice Hall, New York, NY, 1971).
10. M. Livio, *The Story of Phi, the World's Most Astonishing Number* (Broadway Books, 2008).
11. M. Andrews, *Landscape and Western Art* (Oxford University Press, 1999).

12. E. H. Gombrich, *Art and Illusion* (Phaidon Press, London, United Kingdom, 1960).
13. WikiArt, Visual Art Encyclopedia. https://www.wikiart.org. Accessed 17 July 2018.
14. E. Krén, D. Marx, Web Gallery of Art. https://www.wga.hu/. Accessed 17 July 2018.
15. N. Proctor, The Google art project: A new generation of museums on the web? *Curator* **54**, 215–221 (2015).
16. M. Perc, Beauty in artistic expressions through the eyes of networks and physics. *J. R. Soc. Interface* **17**, 20190686 (2020).
17. R. P. Taylor, A. P. Micolich, D. Jonas, Fractal analysis of Pollock's drip painting. *Nature* **399**, 422 (1999).
18. R. Taylor, Pollock, Mondrian and the nature: Recent scientific investigations. *Chaos Complex. Lett.* **1**, 29 (2004).
19. P. Pedram, G. R. Jafari, M. Lisa, The stochastic view and fractality in color space. *Int. J. Mod. Phys. C* **19**, 855–866 (2008).
20. M. Koch, J. Denzler, C. Redies, $1/f^2$ Characteristics and isotropy in the Fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. *PLoS One* **5**, e12268 (2010).
21. C. Montagner, J. M. M. Linhares, M. Vilarigues, S. M. C. Nascimento, Statistics of colors in paintings and natural scenes. *J. Opt. Soc. Am. A* **33**, A170–A177 (2016).
22. J. De La Rosa, A. Suarez, Quantitative approach to beauty. Perceived attractiveness of human faces in world painting. *Int. J. Digital Art Hist.* **1**, 112–129 (2015).
23. D. J. Graham, J. D. Friedenberg, D. N. Rockmore, D. J. Field, Mapping the similarity space of paintings: Image statistics and visual perception. *Vis. Cogn.* **18**, 559–573 (2010).
24. S. Karayev *et al.*, Recognizing image style. arXiv:1311.3715 (15 November 2013).
25. B. Saleh, K. Abe, R. S. Arora, A. Elgammal, Toward automated discovery of artistic influence. *Multimed. Tool. Appl.* **75**, 3565–3591 (2016).
26. R. P. Taylor, Authenticating Pollock paintings using fractal geometry. *Pattern Recogn. Lett.* **28**, 695–702 (2007).
27. S. Lyu, D. Rockmore, H. Farid, A digital technique for art authentication. *Proc. Nat. Acad. Sci. U.S.A.* **101**, 17006–17010 (2004).
28. J. M. Hughes, D. J. Graham, D. N. Rockmore, Quantification of artistic style through sparse coding analysis in the drawings of Pieter Bruegel the Elder. *Proc. Nat. Acad. Sci. U.S.A.* **107**, 1279–1283 (2010).
29. L. A. Gatys, A. S. Ecker, M. Bethge, "Image style transfer using convolutional neural networks" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (CPVR, 2016), pp. 2414–2423.
30. A. Elgammal, B. Saleh, "Quantifying creativity in art networks" in *Sixth International Conference on Computational Creativity (ICCC)* (ICCC, 2015), pp. 39–46.
31. S. P. Fraiberger, R. Sinatra, M. Resch, C. Riedl, A. L. Barabási, Quantifying reputation and success in art. *Science* **362**, 825–829 (2018).
32. M. Schich *et al.*, A network framework of cultural history. *Science* **345**, 558–562 (2014).
33. L. Liu *et al.*, Hot streaks in artistic, cultural, and scientific careers. *Nature* **559**, 396–399 (2018).
34. F. Nake, Information aesthetics: An heroic experiment. *J. Math. Arts* **6**, 65–75 (2012).
35. G. D. Birkhoff, *Aesthetic Measure* (Harvard University Press, Cambridge, MA, 1933).
36. M. Kosheleva, V. Kreinovich, Y. Yam, Towards the use of aesthetics in decision making: Kolmogorov complexity formalizes birkhoff's idea. *Bull. EATCS* **66**, 166–170 (1998).
37. M. Bense, *Einführung in Die Informationstheoretische Ästhetik Grundlegung Und Anwendung in der Texttheorie* (Rowohlt Taschenbuch Verlag, 1969).
38. C. E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
39. J. Rigau, M. Feixas, M. Sbert, Informational aesthetics measures. *IEEE Comput. Graph. Appl.* **28**, 24–34 (2008).
40. I. S. Shin, S. K. Han, K. M. Lee, S. B. Lee, H. W. Jung, Bi-partitioning algorithm for information-theoretic measures of aesthetics. *New Phys. Sae Mulli* **63**, 655–660 (2013).
41. F. Aurenhammer, R. Klein, D. T. Lee, *Voronoi Diagrams and Delaunay Triangulations* (World Scientific Publishing Company, 2013).
42. M. Á. Serrano, M. Boguná, A. Vespignani, Extracting the multiscale backbone of complex weighted networks. *Proc Natl Acad Sci. U.S.A.* **106**, 6483–6488 (2009).
43. V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theor. Exp.* **2008**, P10008 (2008).
44. M. Newman, *Networks* (Oxford University Press, 2018).
45. N. Naik, S. D. Kominers, R. Raskar, E. L. Glaeser, C. A. Hidalgo, Computer vision uncovers predictors of physical urban change. *Proc. Nat. Acad. Sci. U.S.A.* **114**, 7571–7576 (2017).