Rhythmic Brushstrokes Distinguish van Gogh from His Contemporaries: Findings via Automated Brushstroke Extraction

Jia Li, *Senior Member*, *IEEE*, Lei Yao, *Student Member*, *IEEE*, Ella Hendriks, and James Z. Wang, *Senior Member*, *IEEE*

Abstract—Art historians have long observed the highly characteristic brushstroke styles of Vincent van Gogh and have relied on discerning these styles for authenticating and dating his works. In our work, we compared van Gogh with his contemporaries by statistically analyzing a massive set of automatically extracted brushstrokes. A novel extraction method is developed by exploiting an integration of edge detection and clustering-based segmentation. Evidence substantiates that van Gogh's brushstrokes are strongly rhythmic. That is, regularly shaped brushstrokes are tightly arranged, creating a repetitive and patterned impression. We also found that the traits that distinguish van Gogh's paintings in different time periods of his development are all different from those distinguishing van Gogh from his peers. This study confirms that the combined brushwork features identified as special to van Gogh are consistently held throughout his French periods of production (1886-1890).

Index Terms—Oil paintings, van Gogh, postimpressionism, brushstroke style, statistical analysis.

1 Introduction

Art historians employ a wide range of methods for authenticating and dating works by artistic masters, for example, microchemical analysis of paint samples, canvas thread counting, documentary research, and categorizing painting styles and techniques. For the last of these approaches, art historians have become increasingly interested in computer-based analysis schemes. Some of them believe that driven by rapid advancements in digitization, computers can extract certain patterns from images more thoroughly than is possible through manual attempts, can process a much larger number of paintings, and are less subjective [12].

Research efforts based on computational techniques to study art and cultural heritages have emerged in recent years [25], [26], [30], [1], [17], [16], [7], [21], [28], [6], [2], [12], [14]. A recent comprehensive survey by Stork [29] provides more details. A rich resource on old master attribution using forensic technologies is also provided at the Web site of Veritus Ltd. [32], a company that argues forcefully for computational analysis and points out limitations of expert opinions. Previous computerized studies of the paintings by Vincent van Gogh were mostly based on color or local visual features such as texture or edges [2], [12]. Although the extraction of brushstrokes or brushstroke-related features has been investigated [5], [13], [27], [19], [3], it is not evident that these methods can be readily used to find a large number of brushstrokes for a relatively general collection of van Gogh's paintings. For instance, one particular painting of van Gogh is discussed in [27], and some manual operations are necessary to complete the process of extracting brushstrokes. In [13], to find brushstrokes, manual input is required and the method is derived for paintings drastically different from van Gogh's. In [3], the brushstroke feature is constrained to orientation because brushstrokes are not found explicitly.

In this paper, we developed a new system to extract brushstrokes from digitized paintings by van Gogh and his contemporaries. Then, we analyzed the features of these brushstrokes to provide scientific evidence of his unique brushstroke styles. We found that van Gogh's vigorous and repetitive brushstrokes constitute an eminent aspect of his distinctive style. This analysis also suggests that the traits that separate van Gogh from his peers are retained within his own paintings over different stages of his artistic development.

We based our analysis on 45 digitized oil paintings from the collections of the Van Gogh Museum and the Kröller-Müller Museum. Color large-format transparency films of the original paintings were scanned at high resolution and scaled to a uniform density of 196.3 dots per painted inch and digitized to 16 bits per channel. Fig. 1 shows some example paintings by van Gogh. The left half of each scan was provided by the museums for research. The image sizes, proportional to the physical sizes of the canvas, range from 834×319 to $6{,}356 \times 2{,}304$ pixels.

- *J. Li is with the Department of Statistics, The Pennsylvania State University, University Park, PA 16802. E-mail: jiali@stat.psu.edu.*
- *L. Yao and J.Z. Wang are with the College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802. E-mail: luy112@psu.edu, jwang@ist.psu.edu.*
- *E. Hendriks is with the Conservations Department, Van Gogh Museum, Postbus 75366, 1070 AJ Amsterdam, The Netherlands. E-mail: Hendriks@vangoghmuseum.nl.*

Manuscript received 30 Mar. 2011; revised 23 July 2011; accepted 25 Aug. 2011; published online 8 Oct. 2011.
Recommended for acceptance by D. Forsyth.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference. IEEECS Log Number TPAMI-2011-03-0186.
Digital Object Identifier no. 10.1109/TPAMI.2011.203.

(a) *Self-portrait*, Paris, Aug/Sep 1887

(b) *The Sower*, Arles, Nov 1888

(c) *A Crab on its Back*, Arles, Jan/Feb 1889

(d) *View at Auvers-sur-Oise*, May/Jun 1890

Fig. 1. Example van Gogh paintings in the Paris, Arles-Saint-Rémy, and Auvers-sur-Oise periods. Courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation).

1.1 The Problems

Two challenges were designed by art historians in order to explore the application of computational means to study brushwork. Both are related to attribution studies. The first challenge of separating van Gogh from his contemporaries (Fig. 2) is primarily aimed at coming to a more precise definition and measurement of the specific characteristics of van Gogh's style of brushwork, as distinct from other artists of his day. This is relevant to attribution studies because there are paintings by other artists, some in his close circle, that were not made as deliberate copies or forgeries, but have become mistakenly attributed to van Gogh for one reason or another. Many of the expertise paintings that come to the Van Gogh Museum fall into this category. Art historians suggested that we compare two groups of paintings described below.

- Four late paintings by van Gogh. *Portrait of a Young Girl Against a Pink Background* (painting ID number F518, Auvers, late June-early July 1890), *Chestnut Tree in Flower: White Blossoms* (F752, Auvers, May 1890), *Still Life: Vase with Rose Mallows* (F764a, Auvers, June 1890), *View at Auvers* (F799, May-June 1890).
- Four paintings by van Gogh's contemporaries. *Red Cliffs near Anthéor* (S447, by Louis Valtat, c. 1903), *Schönbrunn* (S448, by Carl Moll, c. 1910), *Garden with Hollyhock* (S457, by Ernest Quost, before 1888), and *Mills at Westzijderveld near Zaandam* (S503, by Claude Monet, 1871).

Here, the ID numbers of the van Gogh paintings (F-numbers) are based on the catalogue numbers in the revised edition of the oeuvre catalogue by de la Faille [9]. The ID numbers of the paintings by van Gogh's contemporaries (S-numbers) are

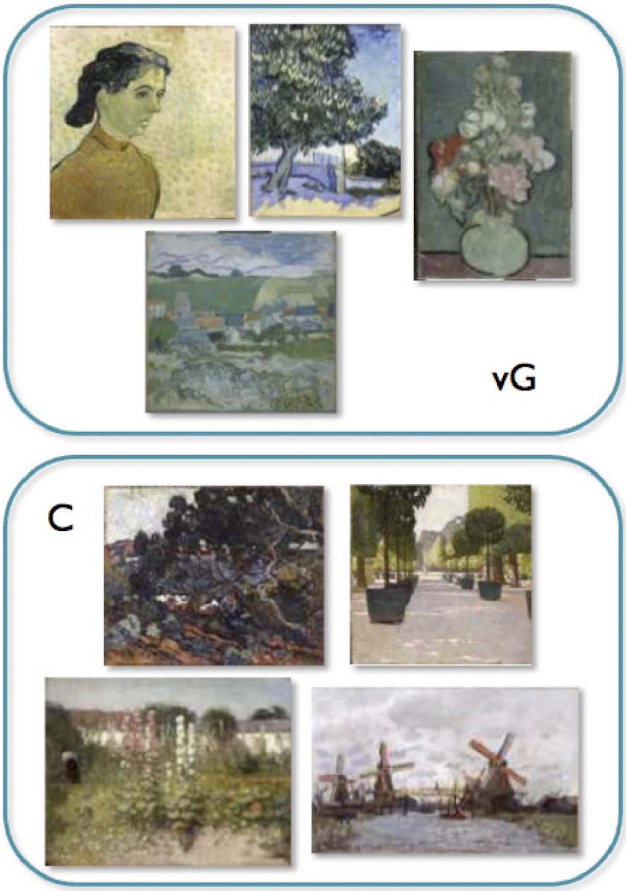


Fig. 2. The attribution challenge aims at finding features that clearly separate van Gogh from his contemporaries. vG: van Gogh. C: van Gogh's contemporaries, that is, non-vG. Painting images courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation) and the Kröller-Müller Museum.

based on the inventory numbers of the Van Gogh Museum collection. The dating of works in the Kröller-Müller Museum collection follows that given in [4]. The dating of works by van Gogh and his contemporaries in the Van Gogh Museum collection follows that given in [31].

The second art historical challenge was to divide van Gogh's paintings by dating into two periods: Paris Period versus Arles and Saint-Rémy Period (Fig. 3). This challenge addresses some real dating questions on van Gogh paintings, with examples of works that seem to share characteristics from different periods of his production and have hence been variously dated in the art historical literature.

- The first group of Paris works includes eight paintings all dating to 1887, the second year of his stay in the French capital: *A Skull* (F297, May-June 1887), *Still Life: Romans Parisiens* (F358, October-November 1887), *Still Life with Plaster Satuette, a Rose and Two Novels* (F360, late 1887), *Japonaiserie: The Flowering Plum Tree: after Hiroshige* (F371, October-November 1887), *Red Cabbage and Onions* (F374, November 1887-February 1888), *Four Cut Sunflowers* (F452, August-October 1887), *Self-Portrait with Straw Hat* (F469, August-September 1887), and *Self-Portrait with Pipe and Straw Hat* (F524, September-October 1887). The van Gogh painting F358 has formerly been dated to the Arles period, but the shift to late Paris, as well as the dating of the other Paris works in the Van Gogh Museum collection, is based on [11]. The dating of F360 follows [4].
- The second group contains eight paintings, seven of which are dated to 1888 in Arles: *Blossoming Almond Branch in a Glass* (F392, March 1888), *Wheatfield* (F411, June 1888), *Seascape at Saintes-Maries* (F415, June 1888), *The Baby Marcelle Roulin* (F441, December 1888), *The Sower* (F451, c. 25 November 1888), *The Green Vineyard* (F475, c. 3 October 1888), *Portrait of Camille Roulin* (F538, December 1888). A last painting in this group—*Leather Clogs* (F607)—was formerly dated to late 1888 in Arles, but is now considered to have been painted in late 1889 when the artist stayed in Saint-Rémy. The dating of F392, F411, and F415 follows [31]. The dating of F475 follows [4]. The dating of F441, F451, and F538 follows [8]. The revised dating of F607 is in [33].

The art historians are interested in identifying attributes that distinguish the two periods and will help to address some unresolved issues in van Gogh scholarship. In particular, we wish to examine the brushwork features in three paintings that bridge different periods in terms of style so that opinion on dating has been divided: *Still Life: Potatoes in a Yellow Dish* (F386), *Willows at Sunset* (F572), and *Crab on Its Back* (F605).

F386 was formerly considered to be one of the last works that van Gogh made in Paris, but has recently been proposed as one of his earliest works made in Arles [4]. The same catalogue discusses the problem of dating F572, which is also now considered to be an early work painted in Arles in March 1888. The question of whether the current January 1889 dating of F605 *Crab on Its Back* should be revised was raised by Pickvance's [23] assertion that the related picture of *Two Crabs* F606 should be relocated from January 1889 to the late Paris period.

1.2 Objective and Contributions of the Work

The objective of the work is to develop a rigorous numerical approach to validate the existence of distinction between two groups of paintings in terms of brushstroke characteristics. Art historians often have anecdotal accounts on which traits distinguish one class of paintings from another. For instance, it is believed that van Gogh's paintings during the Arles and Saint-Rémy period have broader brushstrokes than the Paris period, an observation made subjectively. Our numerical approach, however, extracts brushstrokes in a consistent manner by a computer program. Moreover, formal statistical tests are applied to decide whether significant difference between groups exists in an average sense and to quantify the significance level by p-values.

How art experts are expected to interact with the statistical results is a question of time. Art experts, curators, and especially conservators are already becoming increasingly familiar with the use of quantitative data in support of traditional judgment, in the analysis of materials, automatic thread counting measurements, visualization of underlying compositions based on XRF measurements, etc. Close interdisciplinary interaction is needed to decide the best

Fig. 3. The dating challenge attempts to separate van Gogh paintings into two periods, Paris Period versus Arles and Saint-Rémy Period. Three van Gogh paintings, *Still Life: Potatoes in a Yellow Dish* (F386), *Willows at Sunset* (F572), and *Crab on Its Back* (F605), are in question. Painting images courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation) and the Kröller-Müller Museum.

ways of presenting and interpreting quantitative results within an art historical context. This is expected to be a long term effort not to be accomplished by one or a few papers.

In the study of paintings, the number of available samples tends to be small, e.g., a few dozen, at least in comparison with many online collections of digital photos (several thousand or even up to millions). The availability of high-resolution digitized paintings is limited by the copyright of museums and sometimes by the sheer body of work from an artist. To avoid findings due to overfitting, we hereby adopt a statistical testing approach, profoundly different from the classification approach in existing work [12], [19]. For the two challenges described in the previous section, the set of paintings under study is so small that even if the two groups separate perfectly by a certain feature, we cannot claim reliably that high classification accuracy is achieved. The statistical hypothesis testing, on the other hand, is to validate a much weaker statement. The alternative hypothesis under test is that one group of paintings differs from another in the average value of a certain feature. We thus caution the reader in the interpretation of our results. Two groups that differ in average at a significant level may still overlap substantially depending on the within-group variation, and hence may be difficult to classify. In other words, conclusions drawn here about the comparisons between groups of paintings are only meaningful in the average sense.

Because our emphasis is on hypothesis testing rather than classification, we do not seek features that yield the best classification. We believe that the given collection of paintings is too small to support robust identification of a good set of features for classification purposes. Instead, we focus on features of brushstrokes that are easily interpretable for art historians and the aim is to aid art historians in the validation of certain statements about brushstroke characteristics.

In addition to the concern of overfitting and the fact that brushstroke characteristics are important in their own right; another reason for us not exploiting color and texture features is that such features are highly prone to variations during digitization of paintings. In the case of color, it also lacks fidelity due to aging. The effect of digitization on the computational analysis of paintings is investigated in great depth [22]. It is found that, in some recent studies, the difference found between forgeries/copies and the authentic van Gogh paintings correlates strongly with the extent of blur in the digitized paintings. When the sharpness of the digitized paintings is adjusted to the same level, the wavelet-type texture features can no longer detect the forgeries or copies. Another limitation with texture features is that they are to some extent "black box" features, highly localized and hard to interpret. This also explains why the strong bias drawn from the analysis of texture features as a result of the digitization process was not spotted quickly by experts. In contrast, our features are derived from high-level visual elements, and hence directly reflect the visual appearance of the paintings. Some features such as orientation computed from brushstrokes or by texture analysis can be greatly correlated, while some other features such as brushstroke length, size, broadness homogeneity (BH) are almost irrelevant to texture.

(a) (b) (c)

Fig. 4. The intermingling nature of brushstrokes and the low contrast in some areas create a challenge for computerized extraction of individual brushstrokes. Portions of van Gogh's paintings are shown. (a) *Still Life: Potatoes in a Yellow Dish*. (b) *Willows at Sunset*. (c) *Crab on Its Back*. Painting images courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation) and the Kröller-Müller Museum.

The main contributions of our work include.

- A statistical framework for assessing the level of distinction between categories of paintings and for identifying attributes that differ significantly in average.
- Extracting individual brushstrokes automatically is a challenging problem partly due to the intermingling nature of brushstrokes in oil paintings and the low contrast in some painted areas (Fig. 4). A novel brushstroke extraction algorithm is developed by integrating edge detection and segmentation.
- A set of numerical features is proposed to characterize brushstrokes.
- New insights are gained for two important questions raised by art historians. Computational techniques combining pattern recognition and statistical analysis have rarely been exploited by art historians. The current work exemplifies the potentially fruitful collaboration between the pattern analysis and art communities.

1.3 Outline of the Paper

The remainder of the paper is organized as follows: In Section 2, we provide details on our brushstroke extraction algorithm. The brushstroke features are described in Section 3. The statistical methods used in the comparative study and the findings are summarized in Section 4. Results are presented in Section 5. We conclude in Section 6.

2 Brushstroke Extraction

To extract the brushstrokes automatically, we exploit an integration of edge detection and clustering-based segmentation. Edge detection-based extraction is effective if the edge lines around a brushstroke can be completely identified. In comparison to segmentation-based extraction, the edge-based approach is relatively robust to slight color variation within a brushstroke. On the other hand, the edge line around a brushstroke may not be completely sharp and is thus broken in detection, causing the failure of subsequent brushstroke extraction. To address this issue, we develop morphological operations to complete nearly enclosed edges. In addition, we concatenate a step of image segmentation to acquire brushstrokes missed by edge detection.

2.1 Algorithm

We now explain the step-by-step procedure for extracting brushstrokes. A summarizing flowchart of the algorithm is shown in Fig. 5.

1. The EDISON edge detection algorithm by Meer and Georgescu [20] is applied to the image to identify edge pixels. The edges found are thinned to a single pixel wide.
2. The edge linking algorithm by Kovesi [15] is applied to the detected edges to a) remove short noisy edges and b) trace every legitimate edge and record the coordinates of the pixels on the edge in the tracing order.
3. Perform enclosing operation on the edges. The edge around a brushstroke may not be completely detected due to lack of sharpness. The enclosing operation aims at closing the missing gaps between the broken edge segments. For every end pixel of a detected edge, if there exist other edge pixels which are disconnected from the end pixel within its neighborhood of size 31×31, a straight line linking the end pixel and its closest neighboring edge pixel is added.
4. A brushstroke fully enclosed by an edge is spatially isolated from other nonedge pixels and forms a connected component. We thus extract all the connected components from the image by setting the edge pixels as background and the nonedge pixels as foreground.

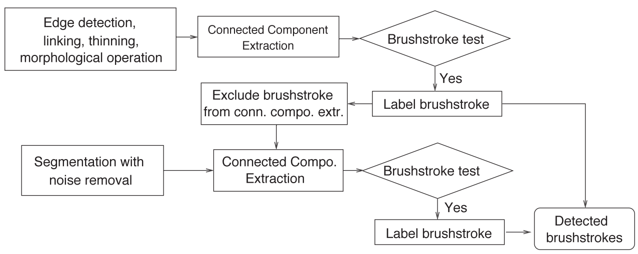


Fig. 5. The flow of the brushstroke extraction algorithm.

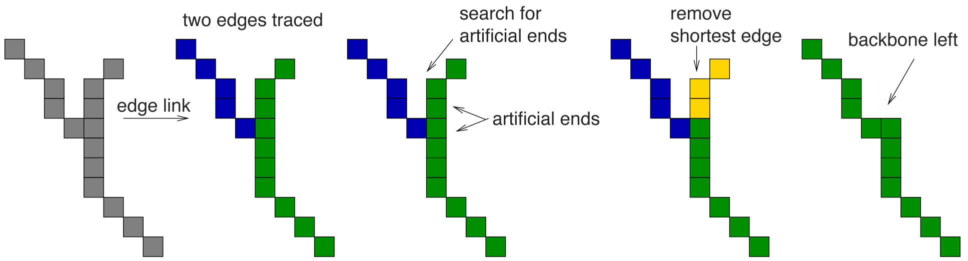


Fig. 6. The process to find the backbone of a branched skeleton.

5. A connected component extracted from the previous step becomes a candidate for a brushstroke if its size (number of pixels contained) is between two preselected thresholds, specifically 100 and 800 in our case. The heuristic is that a connected component that is too small or too large is unlikely to be a brushstroke. The particular thresholds used may be altered depending on the digitization resolution. Skeletons for candidate brushstrokes are obtained by the thinning operation.
6. A candidate brushstroke is labeled as a brushstroke if it satisfies the following conditions. The intuition underlying these conditions is that a brushstroke should appear roughly as a strip, which may be curved. Webbed or salt-and-pepper noisy areas should be excluded.

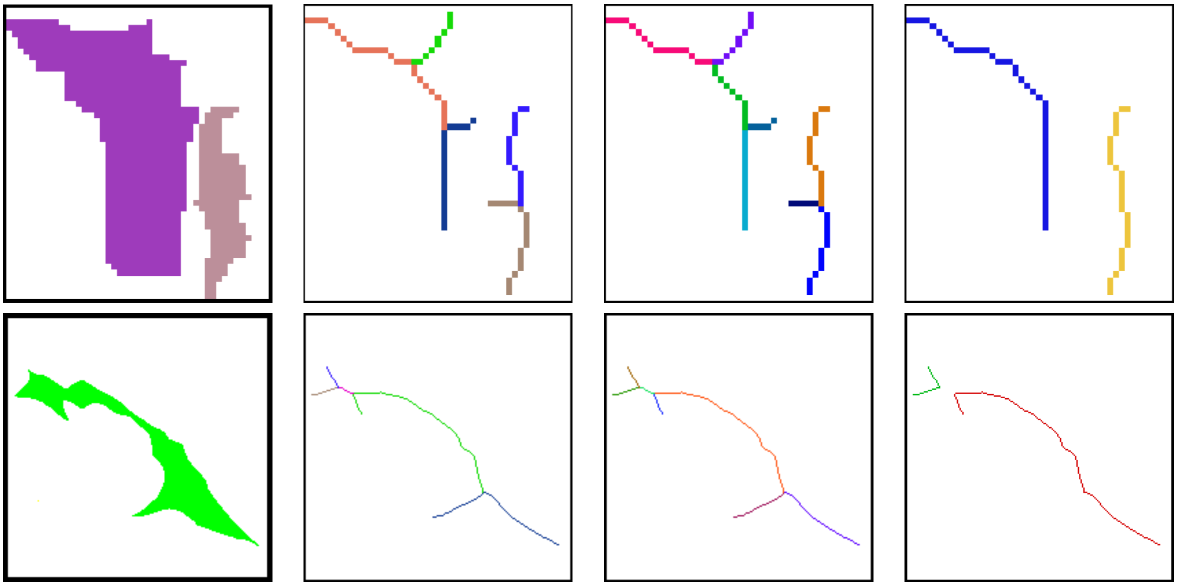
   a. The skeleton is not severely branched. The definition of being severely branched and the method for detection will be described shortly.
   b. The ratio of broadness to length is in the range $[0.05, 1.0]$.
   c. The ratio $\frac{\text{the size of the brushstroke}}{2\times\text{length}\times\text{width span}}$ is within the range $[0.5, 2]$. The width span refers to the maximum distance from a boundary point to the skeleton.
7. Perform image segmentation by clustering all the pixel-level features in the image. The feature vector includes the red, green, and blue (RGB) color components and the gradients of the intensity in horizontal and vertical directions at each pixel. The clustering algorithm applies k-means multiple times, with a gradually decreasing threshold for the average within-cluster distance. After each complete execution of the k-means, connected components are extracted, and those with sufficiently small sizes are not subjected to further clustering. This strategy allows us to more effectively balance the chance of obtaining a fine segmentation and the resistance to noisy (that is, tiny salt-and-pepper-like) connected components. A noise removal procedure is applied after the clustering.
8. The brushstrokes already extracted based on edge detection are set as background. For the pixels not included in the brushstrokes, extract connected components based on the segmentation result. The connected components are passed through Steps 5 and 6 to decide whether each is a brushstroke.
9. Combine brushstrokes obtained by both edge detection and segmentation. The above extraction procedure ensures no overlap between the brushstrokes extracted by the two approaches.

We now describe the method for detecting severely branched skeletons and the method for finding the *backbone* of a not severely branched skeleton. The *backbone* refers to the part of the skeleton remaining after removing the shortest branches, and is itself not branched.

For a given skeleton, the edge link algorithm will trace it and produce a set of edges that are not branched. Each edge is recorded as a sequence of the pixel coordinates positioned one by one from one end of the edge to the other. If the skeleton has no branches, only one edge will be found. Otherwise, multiple edges will be recorded. As shown in Fig. 6, a "Y" shape skeleton is usually divided into two edges. At the branching position of the "Y," we would like to compare the three arms fanning out from this position in order to remove the shortest arm. The two arms left will form the backbone. Because the edge link algorithm usually will not partition a "Y" skeleton as three arms, two artificial end points will be inserted at the branching position of the edge which covers two arms. This edge is then divided into two, each corresponding to one arm. To illustrate the process, Fig. 6 shows every edge in a different color at any stage. It is then straightforward to examine the lengths of the three edges and remove the shortest one. The two longer edges left are merged into one that is the backbone. We can extend the idea to more sophisticated branching patterns. The procedure, referred to as *backbone seeking*, includes the following steps:

1. Find artificial end points along every edge e if they exist by the following process.

   a. Visit the pixels along edge e one by one, starting from one end of the edge. A pixel becomes an artificial end point if i) the pixel is not an end point of edge e; ii) the pixel is not a neighbor (8-connected) of either end point of edge e; iii) the pixel is a neighbor of an end point of another edge.
   b. If a pixel becomes an artificial end point of e, the next pixel to be visited along e becomes the other artificial end point.
   c. Divide edge e into two new edges using the artificial end points. Each new edge will start from one end point of e and stop at one of the artificial end points.
2. Repeat Step 1 until no artificial end point exists in any edge. At this moment, the edges fanning out from any branching position have all been registered separately.



(a) (b) (c) (d)

Fig. 7. Examples for detecting severely branched skeletons and for finding backbones of skeletons. First row: An example of "not severely branched" skeletons. Second row: An example of a "severely branched" skeleton. At any stage, different edges are marked by different colors and every edge is a connected curve. (a) Candidate brushstroke segments. (b) The skeletons and the edges formed by edge link. (c) Edges formed after inserting artificial end points. (d) Edges left after backbone seeking.

3. If an end point of a certain edge has two neighboring end points from two other edges, this end point is marked as a branching position. Remove the shortest edge emitting from this branching position, and merge the two longer edges into one. Then, the three end points at this branching position either are removed from the skeleton or become internal points of the merged edge. Proceed to find and process the next branching position on the skeleton.
4. After processing all the branching positions, if two or more edges are left, this skeleton is declared "severely branched." Otherwise, it is declared "not severely branched."

Note that "not severely branched" is only in the topological sense. Whether the corresponding candidate brushstroke will be accepted depends eventually on its geometry. The conditions specified in Steps 6b) and c) in the brushstroke extraction algorithm have to be satisfied.

Fig. 7 illustrates the backbone seeking process using real candidate brushstroke segments and their skeletons appearing in some paintings. In the example shown in the top row, a region contains two skeletons that are found to be "not severely branched." In the bottom row, a region contains a "severely branched" skeleton, which is reduced to two disconnected edges after backbone seeking.

2.2 Results and Evaluation of Brushstroke Extraction

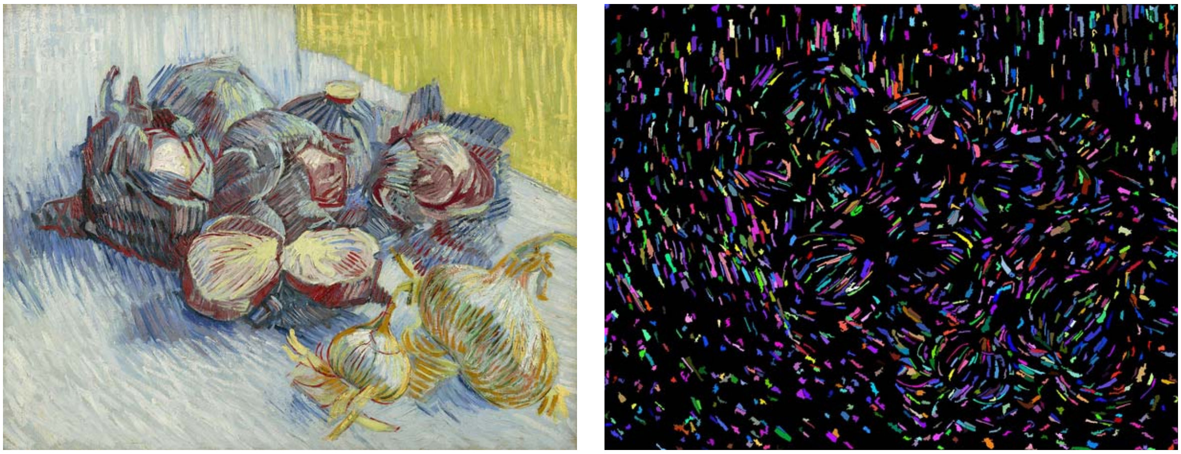
The Van Gogh Museum provided us with several images of the full paintings for illustration purpose in this paper. The brushstroke results obtained for some of these full paintings are shown in Fig. 8.

To numerically evaluate our brushstroke extraction algorithm, we generated manually marked brushstrokes using 10 example regions from paintings in the collection. We did not collect manual data for the whole paintings because they can each contain thousands of brushstrokes. In the example regions, there are, on average, 120 manually marked brushstrokes, which are already laborious to acquire. As shown by examples in Fig. 9, one has to trace the irregular boundaries of the dense brushstrokes by hand.

For every example region, we applied to the manually marked brushstrokes the brushstroke selection criteria specified in Steps 5 and 6 in the brushstroke extraction algorithm. The percentage of manual brushstrokes passing through the selection, referred to as *sensitivity*, is computed. The values of sensitivity for the 10 regions are provided in Table 1. The selection process achieves a high average sensitivity of 95 percent.

We now compare the manual brushstrokes with the ones extracted using our algorithm. For a region in consideration, suppose n brushstrokes, B_i, $i = 1, \ldots, n$, are found by the extraction algorithm, and m brushstrokes, B^*_j, $j = 1, \ldots, m$, are marked manually. B_i (or B^*_j) is the set of pixel coordinates in the brushstroke. $B_i \cap B^*_j$ is the set of overlapped pixels in the two brushstrokes. We say B_i is *validly covered* (or, for brevity, *covered* in the sequel) by B^*_j if the overlap between the two accounts for more than 80 percent of pixels in B_i, that is, $|B_i \cap B^*_j|/|B_i| > 80\%$. We use $C_{i,j} = 1$ to indicate that B_i is covered by B^*_j and $C_{i,j} = 0$ otherwise. Obviously, B_i can be covered by at most one manual brushstroke. Define $C_{i,\cdot} = \sum_{j=1}^{m} C_{i,j}$, which indicates whether B_i is covered by any manual brushstroke at all. $C_{i,\cdot} \in \{0,1\}$. If $C_{i,\cdot} = 1$, we say B_i is *valid*. Define $C_{\cdot,j} = \sum_{i=1}^{n} C_{i,j}$, which is the number of automatically extracted brushstrokes that are covered by manual brushstroke B^*_j. $C_{\cdot,j} \in \{0, 1, \ldots, n\}$. We say that B^*_j is *detected* if $C_{\cdot,j} \geq 1$.

To assess the level of agreement between the manual and automatically extracted brushstrokes, we define the following measures:



(a) *Cabbages and Onions*, Paris, 1887-1888

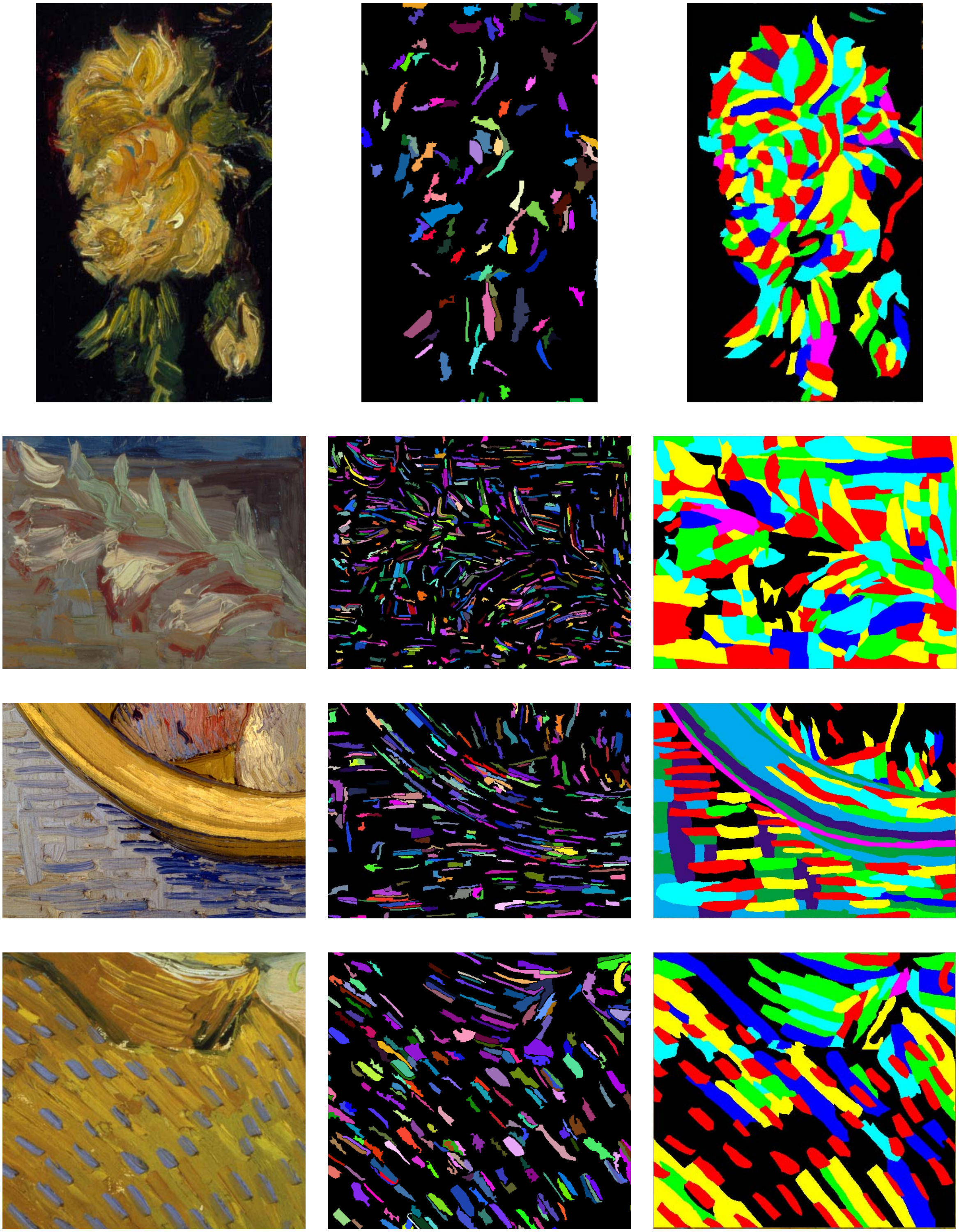
(b) *A Pair of Leather Clogs*, Arles, 1889

(c) *Wheatfield*, Arles, June 1888

Fig. 8. Brushstroke extraction results for van Gogh paintings. Painting images courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation).

- *Valid rate.* $r_v = \sum_{i=1}^{n} C_{i,\cdot}/n$, the percentage of valid automatically extracted brushstrokes.
- *Detection rate.* $r_d = \sum_{j=1}^{m} I(C_{\cdot,j} \geq 1)/m$, the percentage of detected manual brushstrokes.

Table 1 lists the sensitivity, valid rate, and detection rate for the 10 example regions. Fig. 9 shows example regions in paintings by van Gogh, the manual brushstrokes, and the brushstrokes extracted using our method. In these examples, the automatically extracted brushstrokes may not correspond precisely to the physical brushstrokes seemingly laid down by the artist. For instance, when the physical brushstrokes are cross hatched, the brushstrokes detected by the computer are often the inner layer of paint showing through the hatched ones. In brushstrokes



(a) (b) (c)

Fig. 9. Comparison of automatically extracted and manually marked brushstrokes in example regions. (a) Original image. (b) Automatically extracted brushstrokes. (c) Manually marked brushstrokes. Painting IDs from top to bottom: F218, F248a, F386, F518.

executed in thick paint, the paint is often pushed outward, forming ridges on one or two sides, and allowing under paint to show through in the middle. The ridges and the under paint from one manually marked brushstroke may be picked as multiple brushstrokes. Sometimes, only a portion of a manually marked brushstrokes is extracted automatically. Despite the disparity with physical brushstrokes, which we usually think of as individual loads of paint, the brushstrokes extracted by the computer appear to capture well the characteristics of the textured patterns created by the paint, for example, orientation and richness of color. One might argue that such textured patterns are more relevant to our visual impression of the paintings than the physical procedure taken by the artist to reach the end result. Indeed, it is

TABLE 1
Comparing Automatically Extracted and Manually Marked Brushstrokes in 10 Example Regions Cropped from the Specified Paintings

Painting ID	Sensitivity (%)	Valid Rate (%)	Detection Rate (%)
F218	97.4	42.7	21.6
F248a	83.6	75.4	78.8
F297	95.6	57.9	52.0
F374	96.0	58.2	63.2
F386	97.8	73.7	68.4
F415	90.9	46.9	60.0
F518	97.2	60.7	75.2
F538	98.0	49.0	44.9
F572	96.6	83.9	65.6
F652	95.9	50.0	72.5
Average	94.9	59.8	60.2

difficult even for art historians to decipher the painting process down to a brushstroke level for a medium as complex as oil paint, applicable in so many different ways, and let alone for a vigorous artist such as van Gogh, who did not belong to any school and varied his techniques widely over a sadly short period of life as a painter.

3 Brushstroke Features

For each painting, the aforementioned algorithm is applied to extract the brushstrokes. Each brushstroke is recorded digitally as the collection of pixels it contains. Various characteristics, referred to as features, of each brushstroke are computed (Fig. 10). We categorize the features as interactive versus individual. The interactive features depend on the arrangement of other brushstrokes around the one in consideration, which include the number of brushstrokes in a neighborhood, the number of brushstrokes with similar orientations in the neighborhood (NBS-SO), and the amount of variation (measured by standard deviation) in the orientations of brushstrokes in the neighborhood. The size of the neighborhood area is fixed across the paintings. The individual features capture the geometric appearances of the brushstroke itself. There are seven such features: length, width (or broadness), size, broadness homogeneity, elongatedness, straightness, and orientation.

The choice of these features is strongly influenced by the opinions of art historians. For the dating challenge, art historians have noted that the Arles and Saint-Rémy period shows broader brushstrokes. The interactive features are motivated by the fact that van Gogh's paintings are rich in brushstrokes and they appear to be well organized.

We now define the features of the brushstrokes. Suppose the skeleton of a brushstroke has been obtained using the thinning and branch cleaning operation. Denote the center coordinates of a brushstroke i by $(\bar{u}_i, \bar{v}_i)$, where $\bar{u}_i$ is the average vertical position of all the pixels in the brushstroke and $\bar{v}_i$ is the average horizontal position. The coordinate of a pixel in the digitized painting is (u, v), where $u = 0, 1, \ldots, R-1$ and $v = 0, 1, \ldots, C-1$. R is the total number of rows and C is the total number of columns in the image. The brushstroke features are:

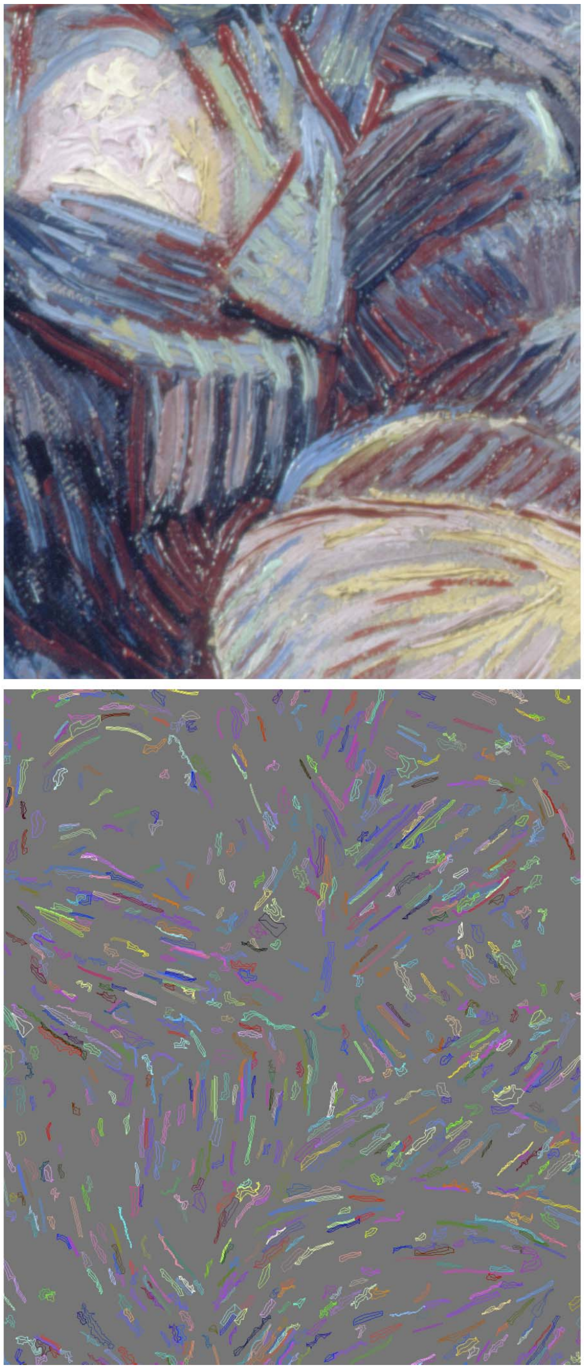


Fig. 10. After the extraction of brushstrokes, numerical features can be computed. The skeletons and boundaries of the brushstrokes are shown.

1. Number of brushstrokes in the neighborhood (NBS-NB). A brushstroke j is a neighbor of brushstroke i if $|\bar{u}_i - \bar{u}_j| < s$ and $|\bar{v}_i - \bar{v}_j| < s$, where s is a threshold set to 200 in our experiments. NBS-NB is obtained simply by counting the number of brushstrokes that are neighbors of i.
2. Number of brushstrokes with similar orientations in the neighborhood. A brushstroke j is considered to have similar orientations as i if the difference between their orientations is below a threshold, set to 0.35 in our experiments.

3. Orientation standard deviation for brushstrokes in a neighborhood (OSD-NB). For any brushstroke i, compute the standard deviation for the orientations of the brushstrokes in the neighborhood of i.
4. Size. The size of a brushstroke is the number of pixels in the brushstroke.
5. Length. The length of a brushstroke is the number of pixels along the skeleton of the brushstroke.
6. Broadness. The broadness of a brushstroke is the average euclidean distance on the image plane from a boundary pixel in the brushstroke to the skeleton of the brushstroke. The distance between a boundary pixel and the skeleton is the minimum distance between the coordinate of the boundary pixel and the coordinate of any pixel on the skeleton.
7. Broadness homogeneity. For every boundary pixel in the brushstroke, find its distance to the skeleton of the brushstroke. The standard deviation of these distances normalized by the broadness of the brushstroke is defined as BH. The smaller the value, the greater the homogeneity.
8. Straightness. To measure the straightness of the brushstroke, we compute the absolute value of the linear correlation coefficient between the horizontal and vertical coordinates of the pixels located on the skeleton of the brushstroke. If the skeleton is a perfect straight line, the correlation coefficient has absolute value of one. Otherwise, if the skeleton is curved, the absolute value of the coefficient will be smaller than one. Suppose a brushstroke contains N pixels with coordinates (u_i, v_i), $i = 1, \ldots, N$. The straightness is defined by $|S_{uv}|/(S_u \cdot S_v)$, where

$$S_{uv} = N\sum_{i=1}^{N} u_i v_i - \sum_{i=1}^{N} u_i \sum_{i=1}^{N} v_i \ ,$$

$$S_u = \sqrt{N\sum_{i=1}^{N} u_i^2 - \left(\sum_{i=1}^{N} u_i\right)^2} \ ,$$

$$S_v = \sqrt{N\sum_{i=1}^{N} v_i^2 - \left(\sum_{i=1}^{N} v_i\right)^2} \ .$$

9. Elongatedness. The measure for elongatedness is defined as the ratio between the length and the broadness.
10. Orientation. The definition given by Russ [24] is used. The orientation of an area is essentially that of its principal axis. Again, suppose a brushstroke contains N pixels with coordinates (u_i, v_i). Let

$$m_u = \sum_{i=1}^{N} u_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} u_i\right)^2,$$

$$m_v = \sum_{i=1}^{N} v_i^2 - \frac{1}{N}\left(\sum_{i=1}^{N} v_i\right)^2,$$

$$m_{u,v} = \sum_{i=1}^{N} u_i v_i - \frac{1}{N}\sum_{i=1}^{N} u_i \sum_{i=1}^{N} v_i.$$

Compute the feature *orientation* as follows:

$$\begin{cases} \frac{\pi}{2}, & \text{if } m_{u,v} = 0; \\ \arctan \frac{m_u - m_v + \sqrt{(m_u - m_v)^2 + 4m_{u,v}^2}}{2m_{u,v}}, & \text{otherwise.} \end{cases}$$

4 Statistical Comparison

In the comparative study of two groups of paintings, we employ a unified statistical paradigm. To compare the overall paintings, the features of the brushstrokes in one painting are summarized by some statistics that then serve as the attributes for the entire painting. For every brushstroke-level feature except orientation, the average across all the brushstrokes in one painting is used as a painting-level attribute. For orientation, the standard deviation instead of average is taken as a painting-level attribute. We observe that van Gogh's brushstrokes do not stay in similar orientations across paintings. Hence, the average orientation is not a meaningful feature. However, the amount of variation in orientation as measured by standard deviation is related to whether a painting conveys a unified or organized look. We normalize the average brushstroke length and broadness by the square root of the painting size (i.e., total number of pixels), while we normalize the average brushstroke size by the painting size. We also include the total number of brushstrokes (TNBs) in a painting as an attribute. In summary, every painting has 11 attributes: total number of brushstrokes, number of brushstrokes in the neighborhood, number of brushstrokes with similar orientations in the neighborhood, orientation standard deviation in the neighborhood (OSD-NB), broadness homogeneity, elongatedness, straightness, length, broadness, size, and orientation standard deviation (OSD).

To test whether two groups of paintings differ significantly in terms of a certain attribute, we set the null hypothesis that the two groups have the same average in this attribute. We then use a two-sided permutation test to compute the p-value. Consider a particular attribute for two groups. Let the attribute values for the first group be $\{x_1, x_2, \ldots, x_n\}$ and those of the second group be $\{x_{n+1}, x_{n+2}, \ldots, x_{n+m}\}$. The two-sided permutation test is performed by randomly shuffling $x_1, \ldots, x_{n+m}$. Assign the first n of the shuffled values, say $\{x_{(1)}, \ldots, x_{(n)}\}$, to the first permuted group and the other m values, $\{x_{(n+1)}, \ldots, x_{(n+m)}\}$ to the second permuted group. Compute the difference between the two groups. For the original two groups, we have

$$\delta_o = \left|\frac{1}{n}\sum_{i=1}^{n} x_i - \frac{1}{m}\sum_{i=1}^{m} x_{n+i}\right|.$$

For any two permuted groups,

$$\delta_p = \left|\frac{1}{n}\sum_{i=1}^{n} x_{(i)} - \frac{1}{m}\sum_{i=1}^{m} x_{(n+i)}\right|.$$

Repeat the random shuffling and count the number of times $\delta_p \geq \delta_o$. The proportion of times $\delta_p \geq \delta_o$ is the p-value. Smaller p-values indicate stronger evidence against the null hypothesis. The rationale for using the permutation test is that if the two groups of paintings have no real difference in a certain attribute, their attributes in either group can be considered as a random assignment from the given pool of

TABLE 2
P-Values for the 11 Attributes under Two Studies

Painting-level Attributes	van Gogh vs. non van Gogh	VG landscape vs. non van Gogh	All VG Non-VG
TNBS	0.457	**0.086**	**0.030**
NBS-NB	**0.029**	**0.029**	**0.002**
NBS-SO	0.114	0.171	**0.000**
OSD-NB	0.114	0.543	**0.002**
elongatedness	**0.029**	**0.029**	**0.000**
straightness	**0.029**	**0.029**	**0.000**
BH	**0.057**	**0.057**	**0.000**
length	0.857	0.343	0.668
size	0.571	0.257	0.661
broadness	0.343	0.171	0.122
OSD	0.114	0.314	**0.006**

Those attributes with a high significance level for each study, as indicated by the p-values, are highlighted.

TABLE 3
P-Values for the 11 Attributes under Two Studies

Painting-level Attributes	Paris vs. Arles and St.-Rémy	Paris vs. Arles/St.-Rémy still life/portrait	Arles/St.-Rémy landscape vs. still life/portrait
TNBS	0.561	**0.059**	0.200
NBS-NB	0.852	0.537	0.343
NBS-SO	0.449	0.566	1.000
OSD-NB	**0.069**	0.588	**0.057**
elongatedness	0.208	**0.018**	**0.086**
straightness	0.837	0.921	0.743
BH	0.763	0.671	0.229
length	**0.097**	**0.014**	0.114
size	**0.090**	**0.014**	0.143
broadness	**0.080**	**0.016**	0.200
OSD	**0.062**	0.576	**0.086**

Those attributes with a high significance level for each study, as indicated by the p-values, are highlighted.

values. The chance of δ_o being extreme under a random assignment should be small. Each permutation corresponds to a random assignment of values into two groups. For a general introduction to permutation test, we refer to [10]. If the values in the original two groups differ significantly in average, δ_o will be more extreme (in this case, larger) than δ_p for a high percentage of permutations, hence leading to a small p-value. At a given threshold α, we recognize that an attribute differs between the two groups at significance level α if the p-value is below α.

To further quantify the separation of two groups, we also introduce the *separation statistic*. Consider two groups with measurements $\{x_{1,1}, x_{1,2}, \ldots, x_{1,n}\}$ and $\{x_{2,1}, x_{2,2}, \ldots, x_{2,m}\}$. Let the q percentile of group i, $i = 1, 2$, be $x_i^{(q)}$. The medians of the two groups are $x_1^{(0.5)}$ and $x_2^{(0.5)}$. Without loss of generality, suppose $x_1^{(0.5)} \geq x_2^{(0.5)}$. The separation statistic is defined as $S = \arg\max_q x_2^{(q)} \leq x_1^{(1-q)}$, which is a ratio less than one. That is, S portion of points in group 1 are larger or equal to S portion of points in group 2, and S is the maximum of the values such that the statement holds. If $x_1^{(0.5)} < x_2^{(0.5)}$, then the separation statistic is $S = \arg\max_q x_1^{(q)} \leq x_2^{(1-q)}$. It can be shown that S is in the range $[0.5, 1]$. When $S = 1$, the two groups are completely separated, that is, $\min_j x_{1,j} \geq \max_l x_{2,l}$ or $\min_j x_{2,j} \geq \max_l x_{1,l}$. If $x_1^{(0.5)} = x_2^{(0.5)}$, then $S = 0.5$.

5 Results

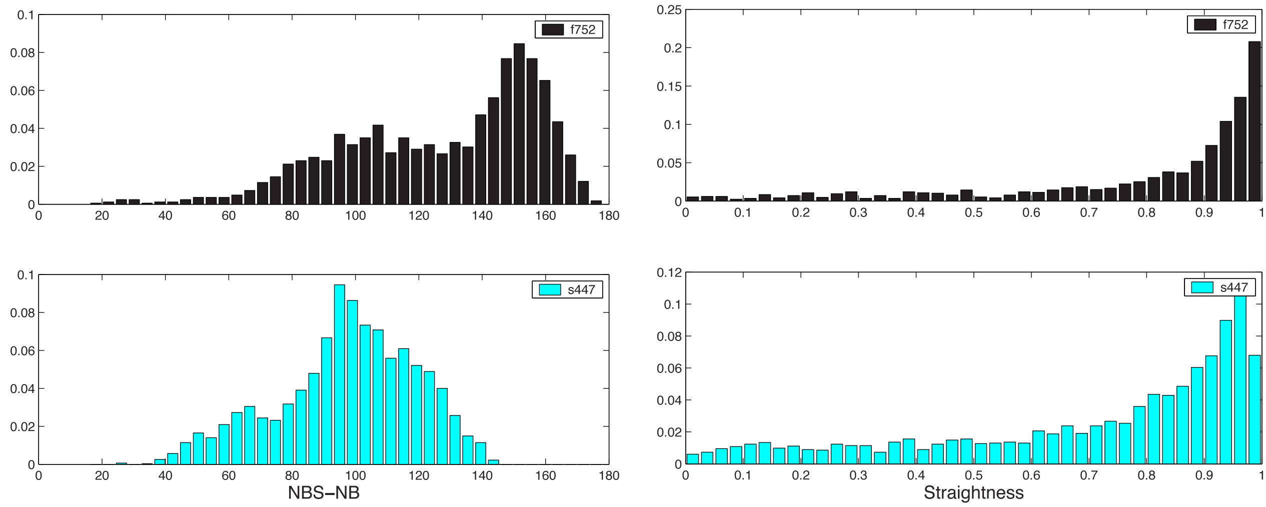
We evaluate the p-values for two comparative studies. The first study is to compare van Gogh's paintings with his contemporaries using the aforementioned eight paintings suggested by the art historians, four of which, referred to as vG, are by van Gogh and the other four, referred to as non-vG, are by his peers. The second study addresses the dating challenge raised by the art historians. As described previously, the two groups each contain eight paintings of van Gogh from the Paris and Arles-Saint-Rémy periods, respectively, referred to as vG-Paris and vG-Arles. The p-values of the permutation test for the 11 attributes under each of two studies are listed in Tables 2 and 3, respectively.

5.1 van Gogh versus Non-van Gogh

In the study of vG versus non-vG, at level $\alpha = 0.1$, four attributes are shown to differ significantly: NBS-NB, elongatedness, straightness, and BH, which are aptly called the *marking attributes*. The first one is an interactive attribute in that it depends on how a brushstroke relates to other brushstrokes, and the last three are individual attributes which characterize the geometric characteristics of the brushstrokes. On average, vG paintings have more brushstrokes in the neighborhood; however, the total number of brushstrokes is not significantly more in paintings belonging to vG than to non-vG. This suggests that the tight arrangement of brushstrokes plays a remarkable role in distinguishing van Gogh from his peers. In terms of geometric appearance, van Gogh's brushstrokes are straighter, more elongated, and more homogeneous in broadness.

As an example, we compare, in detail, one vG painting, *Chestnut Tree in Flower: White Blossoms* (F752, by van Gogh, May 1890), and one non-vG painting, *Red Cliffs near Anthéor* (S447, by Louis Valtat, c. 1903). In Fig. 11, the histograms of brushstroke features are shown for the two paintings. The number of brushstrokes extracted for F752 is 1,655 and for S447 is 3,150. The histograms are normalized for comparison. From Fig. 11a, we see that F752 has much more brushstrokes that have high values for the attribute NBS-NB, although S447 has more brushstrokes in total, indicating that the brushstrokes in F752 appear more closely packed; from Fig. 11b, we see that a much higher percentage of brushstrokes in F752 yield large values for the measure of straightness.

The four non-vG paintings selected for comparison by art historians are all landscape paintings, while, among the four vG paintings, one is a still life and another a portrait. One may suspect that the difference in the painting subjects contributes predominantly to the observed distinctions rather than the difference in painting styles. We thus form the following set of van Gogh landscape paintings to compare with the four paintings of his contemporaries: F411 (*Wheatfield*), F415 (*Seascape at Saintes-Maries*), F475 (*The Green Vineyard*), and F799 (*View at Auvers*). The p-values obtained are listed in Table 2. All the marking attributes identified using the previous four vG paintings remain as



(a) (b)

Fig. 11. The normalized histograms of the brushstroke features of paintings F752 and S447. (a) NBS-NB. (b) Straightness.

marking attributes. This demonstrates that the painting style, not subject matter, is the key factor in the distinction between van Gogh and his contemporaries. Actually, when vG landscape paintings are used for comparison, the total number of brushstrokes also becomes a marking attribute, which shows that the art historian's selection of paintings is steered toward reducing the difference between the vG and non-vG groups.

We also conducted the permutation test between all 31 van Gogh paintings in the collection versus 14 paintings by others. The same three individual marking attributes yield the smallest p-value. For the interactive attributes, the smallest p-value is achieved by NBS-SO. NBS-NB, the marking attribute identified previously using the eight paintings, achieves the second smallest p-value. We also computed the separation statistic for the five marking attributes that distinguish van Gogh from his contemporaries using the 45 paintings. The values are 85.7 percent for NBS-NB, 78.6 percent for NBS-SO, 85.7 percent for elongatedness, 85.7 percent for straightness, and 87.10 percent for BH. For the 31 van Gogh paintings, the average number of brushstrokes extracted in a painting is 3,532, while the average number for the 14 paintings by his contemporaries is 1,207. Fig. 12 compares the boxplots of the numbers of brushstrokes for van Gogh paintings and those of his contemporaries. Apparently, the van Gogh paintings, on average, have more brushstrokes. And several of the van Gogh paintings possess much more brushstrokes than all the paintings of his contemporaries. However, as aforementioned, based on the eight paintings selected thoughtfully by the art historians, the total number of brushstrokes is not a marking attribute between the vG and non-vG groups.

In Fig. 13, scatter plots of NBS-SO versus BH and straightness versus elongatedness are shown for all 45 paintings. The dash dot lines mark the threshold for each single attribute that best separates the two groups. Fig. 13a shows that, on average, vG paintings have brushstrokes that are more homogeneous in broadness (smaller BH value) and are surrounded by more brushstrokes with similar orientations in the neighborhood. The only painting yielding both attributes on the wrong side of the threshold is *Schönbrunn* (S448, by Carl Moll, 1910). In terms of BH alone, three paintings are on the wrong side: *Two Children* (S506, by Cuno Amiet, 1907) and *Vase with Roses* (S286v, by L. Van Rijsel, 1907) are on the side of vG, while *Roses and Peonies* (F249, by van Gogh, June 1886) is on the side of non-vG. It is pointed out by art historians that during van Gogh's first year in Paris (1886), he painted floral still life under the influence of Monticelli's impasto brushwork. It is believed that there are strong similarities between the floral paintings of the two artists. In terms of BH, F249 is close to two floral paintings (ID: 56, 21834) by Monticelli.

Fig. 13b shows that, on average, vG paintings have brushstrokes that are more elongated and straighter. Only the painting *Two Children* (S506, by Cuno Amiet, 1907) is on the vG side of the threshold for both attributes. The painting S506 (Fig. 14) is a known copy of van Gogh painting F784 (June 1890). In terms of the individual brushstroke attributes (straightness, elongatedness, and BH), S506 is on the vG side. For the interactive attribute NBS-SO, S506 is on the non-vG side. This indicates that the copied painting

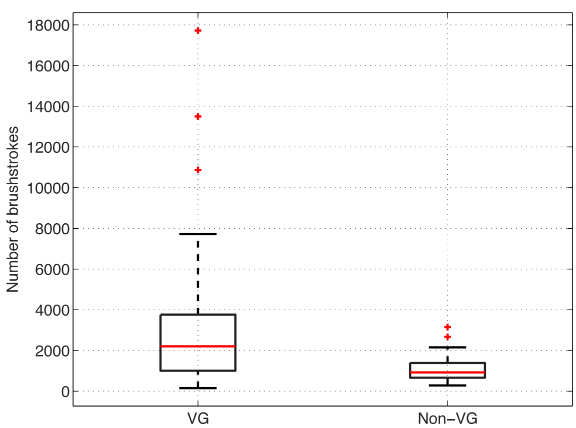
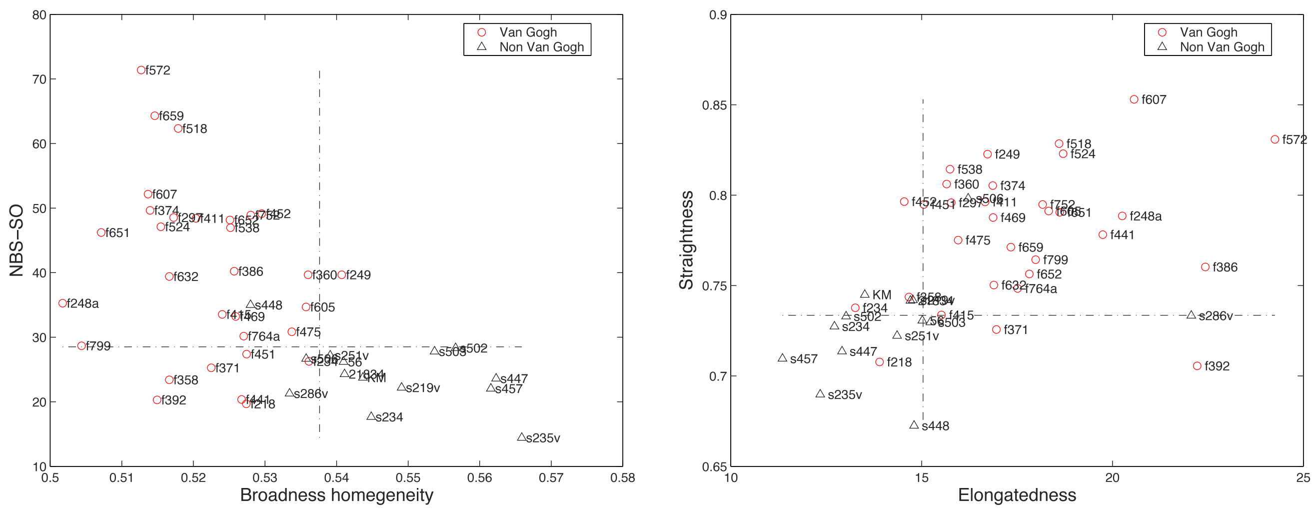


Fig. 12. Boxplots for the number of brushstrokes extracted from van Gogh paintings and those of his contemporaries.



(a) (b)

Fig. 13. The scatter plots of vG versus non-vG paintings. The dash dot lines are the optimal thresholds between the two groups based on each individual attribute. (a) NBS-SO versus BH. (b) Straightness versus elongatedness.

is less effective at mimicking the relationships between brushstrokes in van Gogh's work than the individual look of the brushstrokes.

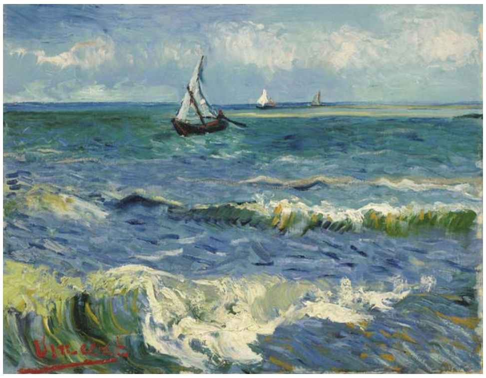
The Wacker forgery painting (F418), shown in Fig. 15, was attributed to van Gogh by art historians for a long period. We compared its values of the four marking attributes with those of the four vG paintings and four non-vG paintings selected by the art historians. It is found that except for broadness homogeneity, the attributes of F418 are clearly closer to the averages of the vG paintings than the non-vG paintings. As for the value of broadness homogeneity, F418 is slightly closer to the vG paintings. This shows that the marking attributes identified by comparing with the paintings of van Gogh's contemporaries cannot point to the subtle difference between van Gogh's paintings and the forgery which once passed the scrutiny of art historians. We note that the paintings of the contemporaries are not good training examples for classifying forgeries because they are not forgeries themselves and have not been misattributed.

5.2 Paris versus Arles and Saint-Rémy

When we compare vG-Paris versus vG-Arles, the marking attributes are length, size, broadness, OSD-NB, and OSD, as shown in Table 3. None of the four marking attributes identified for vG and non-vG shows significant changes across the two periods. On average, in the Arles and Saint-Rémy period, van Gogh's paintings have longer, broader, and larger brushstrokes, and in addition, the orientations of the brushstrokes vary more (higher standard deviation)

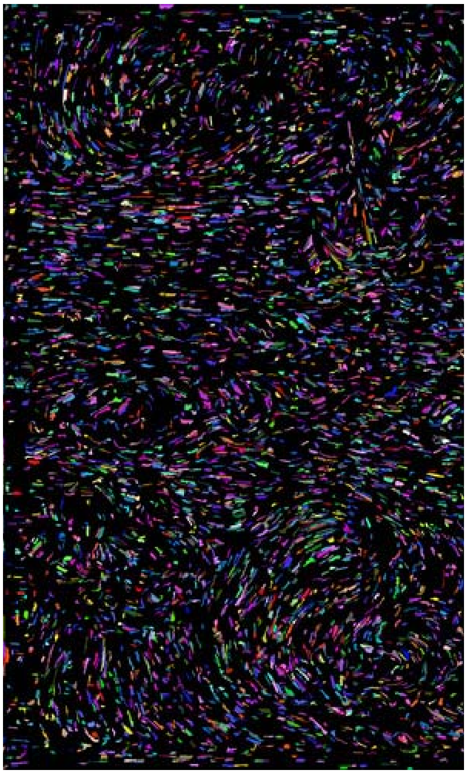
(a) by Cuno Amiet (b) by Vincent van Gogh

Fig. 14. The painting *Two Children* (S506, 1907) is a known copy by Cuno Amiet of an original van Gogh painting (F784, Auvers-sur-Oise, June 1890). Copyright of the Amiet photograph courtesy of Lumiere Technology [18]. Copyright of the van Gogh photograph courtesy of Musée d'Orsay.



(a) *Seascape near Les Saintes-Maries-de-la-Mer* by Vincent van Gogh, Arles, June 1888

(b) the Wacker forgery



(c) brushstrokes extracted from the left half of the forgery

Fig. 15. Brushstrokes extracted for the left half of the The Wacker forgery painting provided in high resolution for analysis. Painting images courtesy of the Van Gogh Museum Amsterdam (Vincent van Gogh Foundation) and the Kröller-Müller Museum.

either across the entire painting or within neighborhoods of the brushstrokes. However, we find that if the length, broadness, and size of a brushstroke are not normalized with respect to the size of the painting, the p-values for these three attributes become 0.152, 0.694, and 0.037, respectively, which indicates that there is no significant difference between the two periods in terms of absolute length and broadness.

Again, there is potential concern about the effect of the subject matter on the observed differences between the two painting periods. For the Paris period, all the paintings are either portrait or still life, while for the Arles/Saint-Rémy period, half of the paintings are landscape. We thus formed a smaller set of paintings from the Arles/Saint-Rémy period containing only portrait or still life: F392 (*Blossoming Almond Branch in a Glass*), F441 (*The Baby Marcelle Roulin*), F538 (*Portrait of Camille Roulin*), F607 (*Leather Clogs*). This set of paintings is compared with the Paris period; the p-values are listed in Table 3. It is shown that, when the subjects are similar, length, size, and broadness remain as marking attributes for the two periods. The two attributes reflecting variation in orientation, i.e., OSD and OSD-NB, are no longer marking attributes. On the other hand, significant difference is shown in terms of elongatedness and the total number of brushstrokes. On average, comparing with portrait or still life paintings in the Arles/Saint-Rémy period, paintings in the Paris period have much more brushstrokes, and the brushstrokes are less elongated.

We also compare the four landscape paintings in the Arles/Saint-Rémy period with the four portrait/still life paintings of the same period and obtain the p-values shown in Table 3. For these two groups that only divide by subjects, three attributes are shown to differ significantly: elongatedness, OSD-NB, and OSD.

The above two tests, one across periods but restricted to the same subject and the other across subjects but restricted to the same period, both confirm that the significant difference in OSD or OSD-NB for the Paris and Arles/Saint-Rémy periods results from subject-wise disparity rather than painting styles. The landscape paintings tend to have higher variation in brushstroke orientation than portrait and still life. However, in terms of brushstroke length, size, and broadness, the difference appears to be caused by styles rather than subjects.

More information can be drawn out on individual paintings from the quantitative results. The one Saint-Rémy painting in this group, F607, was shifted from late Arles to late 1889 in Saint-Rémy, based partly on the “rhythmic repetitions of the short brushstrokes,” a feature observed in other paintings of this specific date. The quantitative data confirm the short yet rhythmic brushstrokes in this painting. Also striking are the relatively long brushstrokes present in F392 *Almond Branch in Glass*, as measured in the original lengths, suggesting that van Gogh did not necessarily turn to shorter brushstrokes when working on a small scale.

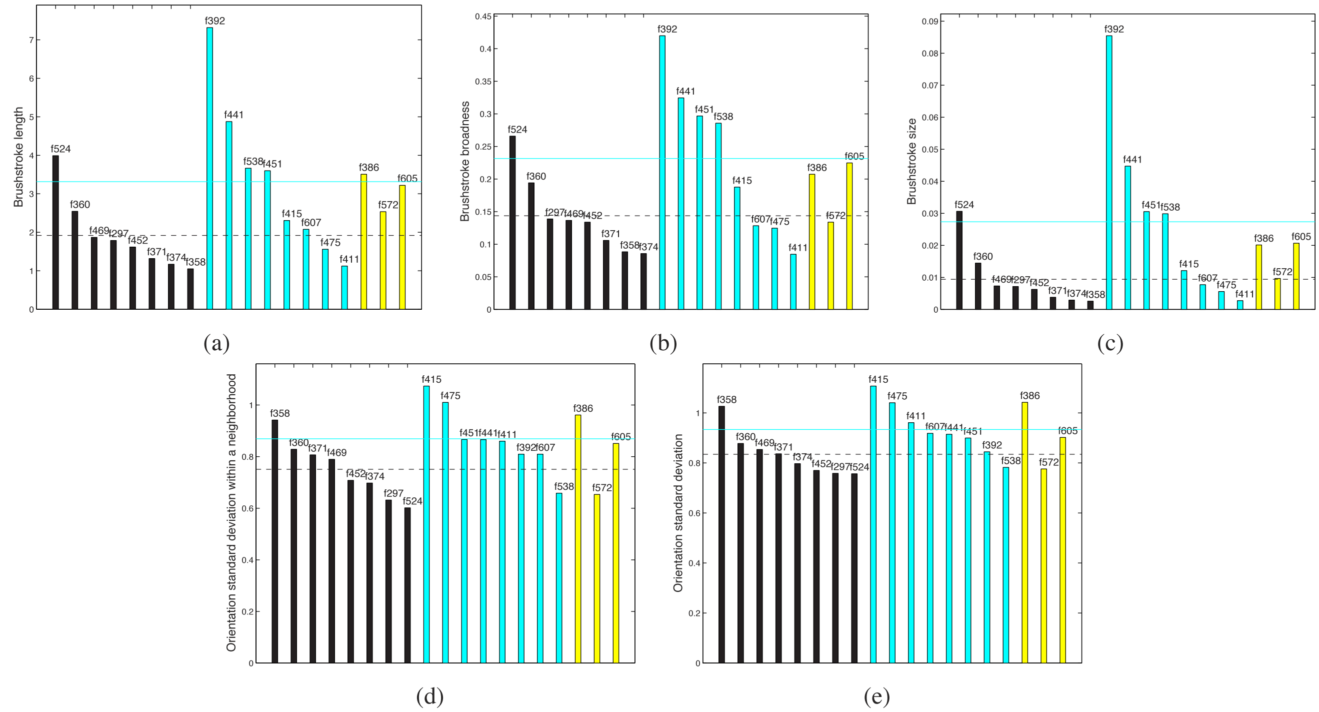


Fig. 16. Compare five attributes of paintings in the Paris (first eight in the bar plot) and Arles (next eight) groups. These five attributes are tested to be significantly different in average between the two groups. The corresponding values of the attributes for the three paintings to be dated are shown at the right end. The two horizontal lines indicate the averages of the two groups, respectively. (a) Brushstroke length. (b) Broadness. (c) Size. (d) Orientation standard deviation within a neighborhood (OSD-NB). (e) Orientation standard deviation.

In order to date the three paintings put in question by the art historians, we compare the values of the five marking attributes with the average values of the paintings in the two periods, respectively. As shown in Fig. 16, paintings F386 and F605 are substantially closer to the Arles and Saint-Rémy period according to every marking attribute. This would seem to agree with art historical observation that both paintings show the bolder, more stylized and graphic touch associated with van Gogh's later French works. However, F572 is substantially closer to the Paris period according to brushstroke broadness, size, OSD, and OSD-NB, and marginally closer to the Paris period according to brushstroke length. The relatively small, narrow, and straighter brushstrokes detected in this work can partly be explained by the subject matter of grass and reeds depicted. Indeed this has been the reason for art historians to associate the landscape with certain late Paris works where comparable subject matter is handled in a similar way.

6 Summary of Findings and Conclusions

With the capability of extracting individual brushstrokes automatically, there are endless possibilities for in-depth analysis of oil paintings. Our work represents only a first step in a promising direction.

The primary findings of our studies are:

1. The marking attributes of van Gogh versus non-van Gogh do not overlap with those distinguishing van Gogh's Paris and Arles/Saint-Rémy periods.
2. The four marking attributes of van Gogh versus non-van Gogh are: NBS-NB, elongatedness, straightness, and BH.
3. The five marking attributes of Paris versus Arles/Saint-Rémy are: length, size, broadness, OSD-NB, and OSD. The last two attributes result from subject-wise difference in the two periods, while the first three are caused by styles.
4. Although the copy of *Two Children* (S506, by Cuno Amiet, 1907) is similar to van Gogh in terms of individual brushstroke attributes, elongatedness, and straightness, it is closer to non-van Gogh in terms of the interactive brushstroke attribute NBS-SO.
5. Based on the marking attributes of van Gogh versus non-van Gogh, we cannot detect the Wacker forgery painting (F418).
6. The paintings F386 (*Still Life: Potatoes in a Yellow Dish*) and F605 (*Crab on Its Back*) are dated to the Arles and Saint-Rémy period, while the painting F572 (*Willows at Sunset*) is dated to the Paris period.

The results from computer-based analysis presented here clearly suggest that rhythmic brushstrokes in van Gogh's paintings distinguish his work from those of his contemporaries, which aligns with long-held art historical opinion on van Gogh's unique style of painting. For the first time though, the information is presented in a quantitative way, providing more refined and accurate data to substantiate the art historians' opinion. Furthermore, these new techniques were applied to compare brushwork characteristics in three paintings that have proven hard to date by art historians using traditional means: *Still Life: Potatoes in a Yellow Dish*, *Willows at Sunset*, and *Crab on Its Back*. This provided new quantitative evidence to separate the works into two distinct periods of production based on the different characteristics of their brushwork, demonstrating the usefulness of computer-based analysis as an added tool to help shed light on some standing debates among scholars.

Acknowledgments

During the four-year period of this research, J. Li was supported by the US National Science Foundation (NSF) under Grant Nos. 0936948 and 0705210, and The Pennsylvania State University. L. Yao and J.Z. Wang were supported under NSF Grant No. 0347148 and The Pennsylvania State University. The computational infrastructure was provided under NSF Grant No. 0821527. The Van Gogh and the Kröller-Müller Museums provided the photographs of the original paintings for this research. C.R. Johnson Jr., E. Hendriks, and L. van Tilborgh prepared the challenges before the Second International Workshop on Image Processing for Artist Identification, Van Gogh Museum, October 2008. The authors are also indebted to Eric Postma, Igor Berezhnoy, Eugene Brevdo, Shannon Hughes, Ingrid Daubechies, and David Stork for assistance with image data preparation and for useful discussions. They thank John Schleicher at Penn State University for producing the manually marked brushstrokes. J. Li would like to thank C.R. Rao and Robert M. Gray for encouragement, and her toddler daughter Justina for enduring her occasional late working hours. Portions of the work were done when J.Z. Wang was a visiting professor of robotics at Carnegie Mellon University in 2008, and when he was a program manager at the NSF in 2011. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Foundation. He would like to thank Takeo Kanade, Gio Wiederhold, Dennis A. Hejhal, Maria Zemankova, and Stephen Griffin for encouragement. The authors thank the reviewers and the associate editor for constructive comments.

References

[1] K. Barnard, P. Duygulu, and D. Forsyth, "Clustering Art," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 434-439, 2001.

[2] I.J. Berezhnoy, E.O. Postma, and H.J. van den Herik, "Computer Analysis of van Gogh's Complementary Colours," *Pattern Recognition Letters,* vol. 28, no. 6, pp. 703-709, 2007.

[3] I.J. Berezhnoy, E.O. Postma, and H.J. van den Herik, "Automatic Extraction of Brushstroke Orientation from Paintings," *Machine Vision and Applications,* vol. 20, no. 1, pp. 1-9, 2009.

[4] J.T. Berge et al., *The Paintings of Vincent van Gogh in the Collection of the Kröller-Müller Museum.* Kröller-Müller Museum, 2003.

[5] T.J. Bright, "Brush Mark Analysis Method for Painting Authentication," US Patent 6,017218, Washington, D.C.: Patent and Trademark Office, 2000.

[6] M. Bulacu and L.R.B. Schomaker, "Text-Independent Writer Identification and Verification Using Textural and Allographic Features," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 29, no. 4, pp. 701-717, Apr. 2007.

[7] A. Criminisi and D.G. Stork, "Did the Great Masters Use Optical Projections While Painting? Perspective Comparison of Paintings and Photographs of Renaissance Chandeliers," *Proc. Int'l Conf. Pattern Recognition,* vol. 4, pp. 645-648, 2004.

[8] D. Druick and P.K. Zegers, *Exhibition Cat Van Gogh and Gauguin: The Studio of the South,* The Art Inst. of Chicago and Van Gogh Museum, 2001-2002.

[9] J.-B. de la Faille, *The Works of Vincent van Gogh: His Paintings and Drawings,* Reynal, 1970.

[10] P.I. Good, *Permutation, Parametric, and Bootstrap Tests of Hypothesis,* third ed. Springer, 2009.

[11] E. Hendriks and L. van Tilborgh, *Vincent van Gogh, Paintings, Antwerp and Paris (1885-1888),* M. van Eikema Hommes and M. Hageman, vol. 2. Van Gogh Museum, June 2011.

[12] C.R. Johnson Jr, E. Hendriks, I.J. Berezhnoy, E. Brevdo, S.M. Hughes, I. Daubechies, J. Li, E. Postma, and J.Z. Wang, "Image Processing for Artist Identification—Computerized Analysis of Vincent van Gogh's Painting Brushstrokes," *IEEE Signal Processing Magazine,* vol. 25, no. 4, pp. 37-48, July 2008.

[13] P. Kammerer and R. Glantz, "Segmentation of Brush Strokes by Saliency Preserving Dual Graph Contraction," *Pattern Recognition Letters,* vol. 24, no. 8, pp. 1043-1050, 2003.

[14] A.G. Klein, D.H. Johnson, W.A. Sethares, H. Lee, C.R. Johnson, and E. Hendriks, "Algorithms for Old Master Painting Canvas Thread Counting from X-Rays," *Proc. 42nd Asilomar Conf. Signals, Systems, and Computers,* pp. 1229-1233, 2008.

[15] P. Kovesi, School of Computer Science and Software Eng., The Univ. of Western Australia, http://www.csse.uwa.edu.au, 2001.

[16] J. Li and J.Z. Wang, "Studying Digital Imagery of Ancient Paintings by Mixtures of Stochastic Models," *IEEE Trans. Image Processing,* vol. 13, no. 3, pp. 340-353, Mar. 2004.

[17] J.C.A. van der Lubbe, E.P. van Someren, and M.J.T. Reinders, "Dating and Authentication of Rembrandt's Etchings with the Help of Computational Intelligence," *Proc. Int'l Cultural Heritage Informatics Meeting,* pp. 485-492, Sept. 2001.

[18] Lumiere Technology, http://www.lumiere-technology.com, 2011.

[19] L.J.P. van der Maaten and E.O. Postma, "Identifying the Real van Gogh with Brushstroke Textons," Technical Report TiCC TR 2009-001, Tilburg Univ., 2009.

[20] P. Meer and B. Georgescu, "Edge Detection with Embedded Confidence," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 12, pp. 1351-1365, Dec. 2001.

[21] C. Papaodysseus, D.K. Fragoulis, M. Panagopoulos, T. Panagopoulos, P. Rousopoulos, M. Exarhos, and A. Skembris, "Determination of the Method of Construction of 1650 B.C. Wall Paintings," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 28, no. 9, pp. 1361-1371, Sept. 2006.

[22] G. Polatkan, S. Jafarpour, A. Brasoveanu, S. Hughes, and I. Daubechies, "Detection of Forgery in Paintings Using Supervised Learning," *Proc. IEEE Int'l Conf. Image Processing,* pp. 2921-2924, 2009.

[23] R. Pickvance, "Van Gogh," *The Burlington Magazine,* pp. 500-502, 2006.

[24] J.C. Russ, *The Image Processing Handbook,* pp. 553-554, CRC Press, 2006.

[25] R. Sablatnig, P. Kammerer, and E. Zolda, "Hierarchical Classification of Paintings Using Face and Brush Stroke Models," *Proc. Int'l Conf. Pattern Recognition,* pp. 172-174, 1998.

[26] R. Sablatnig, P. Kammerer, and E. Zolda, "Structural Analysis of Paintings Based on Brush Strokes," *Fakebusters: Scientific Detection of Fakery in Art,* W.C. McCrone and R.J. Weiss, eds., pp. 222-244, Hansen, 1999.

[27] M. Shahram, D.G. Stork, and D. Donoho, "Recovering Layers of Brush Strokes through Statistical Analysis of Color and Shape: An Application to van Gogh's Self Portrait in a Grey Felt Hat," *Proc. SPIE Electronic Imaging: Computer Image Analysis in the Study of Art,* vol. 6810, pp. 68100D1-12, 2008.

[28] M. van Staalduinen, J.C.A. van der Lubbe, G. Dietz, F. Laurentius, and T. Laurentius, "Comparing X-Ray and Backlight Imaging for Paper Structure Visualization," *Proc. Electronic Imaging and Visual Arts Conf.,* pp. 108-113, Apr. 2006.

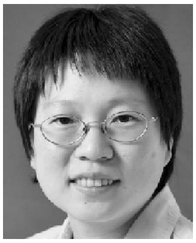
[29] D.G. Stork, "Computer Vision and Computer Graphics Analysis of Paintings and Drawings: An Introduction to the Literature," *Proc. Int'l Conf. Computer Analysis of Images and Patterns,* pp. 9-24, 2009.

[30] R.P. Taylor, A.P. Micolich, and D. Jonas, "Fractal Analysis of Pollock's Drip Paintings," *Nature,* vol. 399, p. 422, 1999.

[31] *The Rijksmuseum Vincent Van Gogh,* E. van Uitert and M. Hoyle, eds., Meulenhoff/Landshoff, 1987.

[32] Veritus Ltd., http://www.veritusltd.com, 2011.

[33] *Exhibition Catalogue: Vincent van Gogh and His Time. Still Lifes from the Van Gogh Museum and the H.W. Mesdag Museum,* Seji Togo Memorial Yasuda Kasai Musem of Art, Tokyo, 1996.



Jia Li received the MSc degree in electrical engineering, the MSc degree in statistics, and the PhD degree in electrical engineering, all from Stanford University. Currently, she is an associate professor of statistics and, by courtesy appointment, in computer science and engineering, The Pennsylvania State University, University Park. She worked as a visiting scientist at Google Labs in Pittsburgh from 2007 to 2008, a research associate in the Computer Science Department, Stanford University, in 1999, and a researcher at the Xerox Palo Alto Research Center from 1999 to 2000. Her research interests include statistical modeling and learning, data mining, computational biology, image processing, and image annotation and retrieval. She is a senior member of the IEEE.

Lei Yao received the BS and MS degrees in computer science from Zhejiang University, China. Currently, she is working toward the PhD degree and is a research assistant in the College of Information Sciences and Technology, The Pennsylvania State University. In 2005, she served as a research intern at the Institute of Telematics, University of Lübeck, Germany. She worked as a research intern at SONY US research center in the summer of 2011. She has been awarded a Toshiba fellowship and a Marykay fellowship. Her research interests include computational modeling of aesthetics and computerized analysis of paintings. She is a student member of the IEEE.

Ella Hendriks received the graduation degree in art history from the University of Manchester, United Kingdom, postgraduate training at the Conservation of Easel Paintings of the Hamilton Kerr Institute, University of Cambridge, United Kingdom, in 1982 and 1986, respectively, and the PhD degree in art history for her studies on the working method of van Gogh as represented by the Antwerp and Paris period paintings (1886-1888) in the collection of the Van Gogh Museum, in 2006. From 1988 to 1999, she was head of conservation at the Frans Halsmuseum, Haarlem, The Netherlands. Since 1999, she has been the head of conservation at the Van Gogh Museum, Amsterdam.

James Z. Wang received the bachelor's degree summa cum laude in mathematics and computer science from the University of Minnesota, the MS degree in mathematics and the MS degree in computer science, both from Stanford University, and the PhD degree in medical information sciences from Stanford University. Currently, he is a professor in the College of Information Sciences and Technology, the Department of Computer Science and Engineering, and the integrative biosciences program at The Pennsylvania State University. He is also a program manager at the US National Science Foundation. He has been a visiting professor of the Robotics Institute at Carnegie Mellon University (2007-2008). He has also held visiting positions at SRI International, IBM Almaden Research Center, NEC Computer and Communications Research Lab, and Chinese Academy of Sciences. In 2008, he served as the lead guest editor of the *IEEE Transactions on Pattern Analysis and Machine Intelligence* special section on real-world image annotation and retrieval. His main research interests include automatic image tagging, image retrieval, computational aesthetics, and computerized analysis of paintings. He has been a recipient of an NSF Career award and an endowed PNC Professorship. He is a senior member of the IEEE.