

Image Processing for Artist Identification

Computerized Analysis of Vincent van Gogh’s Painting Brushstrokes

C. Richard Johnson, Jr., Ella Hendriks, Igor J. Berezhnoy, Eugene Brevdo,
Shannon M. Hughes, Ingrid Daubechies, Jia Li, Eric Postma, and James Z. Wang

As image data acquisition technology has advanced in the past decade, museums have routinely begun to assemble vast digital libraries of images of their collections. The cross-disciplinary interaction of image analysis researchers and art historians has reached a stage where technology developers can focus on image analysis tasks supportive of the art historian’s mission of painting analysis, in addition to activities in image acquisition, storage, and database search [2, 20].

In particular, the problem of artist identification seems ripe for the use of image processing tools. In making an attribution experts often use not only current knowledge of the artist’s common practices, in combination with meticulous comparisons of a variety of technical data (acquired e.g. through ultraviolet fluorescence, infrared reflectography, x-radiography, paint sampling, and/or canvas weave count); they also include a visual assessment of the presence of the artist’s “handwriting” in the brushwork. This suggests that mathematical analysis of a painting’s digital representation could assist the art expert in the process of attribution.

A survey of the literature [3, 17, 19, 22, 26, 30] reveals that image processing tools aimed at supplementing the art historian’s toolbox are currently in the earliest stages of development. In part, this is because the necessary data for research has not been made widely available. To jump-start the development of such methods, the Van Gogh and Kröller-Müller Museums in the Netherlands graciously agreed to make a dataset of 101 high-resolution grayscale scans of paintings within their collections available to groups of image processing researchers from several different universities. To our knowledge, this is the first time that a dataset of such size and quality has been available for this purpose or that such a dataset has been available simultaneously to multiple research groups, so that results of the different approaches in development can be compared directly. For more details, see [14]. This paper describes the approaches to brushwork analysis and artist identification developed by three research groups, within the framework of this dataset.

The description of their results follows below, after a first section that describes brushwork characteristics as used by art historians for authentication purposes, and a brief section describing the database as well as the image analysis tools used (wavelets, hidden Markov models).

Art Historical Brushstroke Analysis

It is important to understand what the art specialist currently considers when visually assessing brushwork in paintings. A literature search shows that few art specialists have attempted to describe a method or provide guidelines for the visible analysis of brushwork in paintings. Differences in individual perception, viewing conditions, and knowledge of the picture’s materials, state of preservation, and the painter’s common working methods are just some factors that play a role.

A first step is to establish which parts of a painting should be discounted in an evaluation of brushwork, since they may not be due to the artist’s hand. A variety of imaging techniques can be used to discriminate and document additions by restorers, including high resolution digital photographs of the painting captured at different spectral wavelengths [23]. Visual inspection can also be useful. For instance, original parts of the painting have also changed in appearance due to the deterioration of materials used. An example is color shift and fading as a consequence of van Gogh’s use of poor quality paints; this can have a substantial impact on the perceived clarity and style of brushwork (see e.g., [28]). Brushstrokes may also be disrupted by the formation of drying cracks. For example, in van Gogh’s Paris works the drying cracks are usually associated with the presence of an underlying, slow-drying oil film with zinc white pigment, applied by the artist to cover up an abandoned design [13]. Another frequent change is the flattened texture of *impasto*

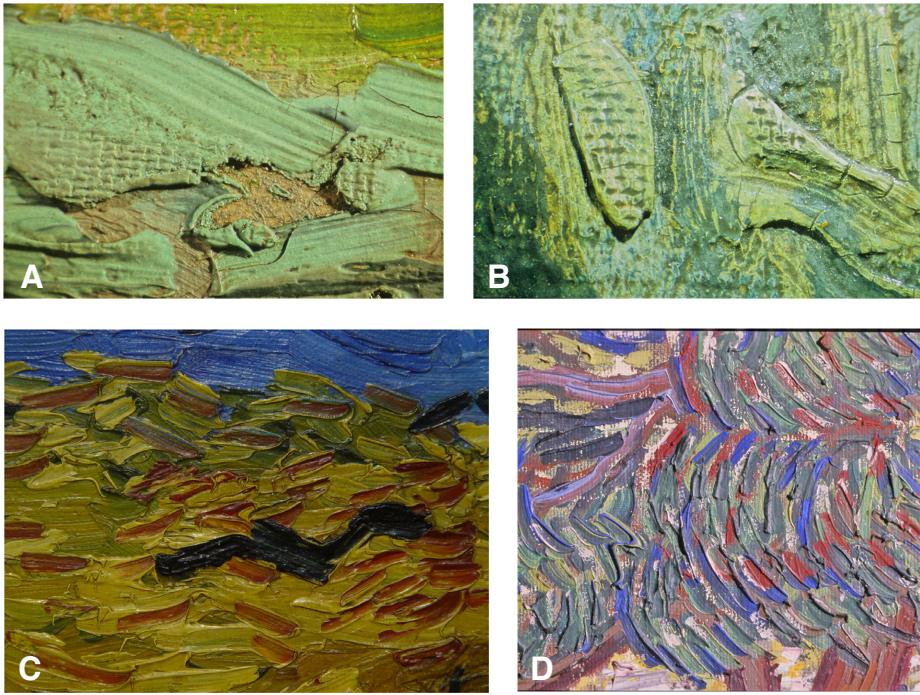


Figure 1: Portions of van Gogh paintings. From upper-left to lower-right: A. Detail of flattened impasto with weave imprint in F779 (*Crows in the Wheatfields*, July 1890). A loss occurred in the green brushstroke where fresh paint stuck and was pulled off. B. Detail of false weave imprint in a Wacker forgery, F614 (*Cypresses*). C. Raking light detail of paint in F779 (*Crows in the Wheatfields*, July 1890) showing wet-on-wet, drippy brush marks. D. Detail of wet-on-dry touches in F659 (*The Garden of Saint Paul's Hospital*, November 1889). Final red accents were applied onto dry paint, leaving the underlying impasto undisturbed.

(thickly textured) brushwork, due to later lining treatment (adhesion of a second canvas to the reverse of the original picture support using a heated iron or hot table), or to the painter's own habit of leaning, rolling or piling freshly painted canvases while the paint was still soft (Fig. 1). These fine-scale "original damages" influence the reading of brushwork detail, but may also be considered characteristic of the painter's working method.

The next step is to describe original brushwork characteristics observed across the remainder of the painting surface. One larger scale feature to look for is the recurrent use of certain brushstroke patterns, involving a series of fast rhythmic touches arranged in a particular way. Especially in his later French paintings, van Gogh seems to have developed a consistent vocabulary of graphic touches that range from "elbow-strokes," to "serial dabs," to "halo" and "brickwork" patterns, for example. It would seem useful to pinpoint key moments when these characteristic brushwork configurations were first introduced, or the periods in which they prevailed, providing a framework to assist with the dating of van Gogh's oeuvre. At present, though references to these various patterns of brushwork may be found scattered throughout the art historical literature, there is no common procedure used to gather data or descriptive terminology used. The most thorough descriptions are to be found in exhibition catalogues, such as [24].

Painting analysts also attempt to identify the specific size and shape of the individual brushes that a painter used in each painting from the marks that they have left behind. X-ray images can be useful for this purpose, or examination in so-called raking light, low-angled light that skims the picture surface, enhancing its topography. Still there are many instances where the brushmarks are too indistinct to measure and describe. Some indeterminate passages are a deliberate consequence of van Gogh's technique of diffusing crisp brushstrokes that were still wet, by wiping with a cloth, scrubbing with a dry bristle brush, or dabbing with the fingertips. Even when legible, the shape of marks made by a single brush will vary somewhat. This may be due to features such as non-uniform profile of the hair bundle, variable paint loading, wrist action, and pressure exerted on the brush. Though uncommon, occasional marks indicate the use of a palette knife instead of a brush to execute select details or parts of a painting.

Importantly, brushwork examination should take into account the visual impact of the painting materials used. Observation of van Gogh's impasto generally points toward the use of a paint of fluid consistency, but with strong internal cohesion or "tenacity" that leaves the marks standing as applied [29]. This combination explains characteristic fine trails of paint formed where the brush was lifted from the canvas, perhaps carried across from one touch to the next, as well as ridges of surplus paint flowing out toward the edges of brushstrokes. This is all visible in Fig. 1.

Van Gogh used ready-made tube colors, so that the rheology of his paints was largely determined by their methods of commercial preparation. Recent attempts to replicate the manufacture of colors (and picture supports) used by van Gogh using historically accurate techniques, pinpointed the particular oil binding medium and its way of preparation as a critical factor [6]. A mixture of poppy oil and water-washed linseed oil was found to provide a lead white oil paint with a character that most closely resembled what was seen on van Gogh paintings [6]. Since van Gogh used the same types of commercial paints as his French contemporaries, it is not surprising that similar paint flow features may be seen in their pictures too. Still, artists of van Gogh's own generation could manipulate similar paints in individual ways that define a personal "handwriting," so that stylistic assessment remains of paramount importance to identify the artist's hand.

Another critical factor affecting the appearance of brushstrokes, is the surface onto which they are applied. Technical and analytical studies have disclosed the diversity of van Gogh's technique across time (e.g. [12]), ranging from his practice of oiling out the surface of ready-primed canvas to facilitate brushstrokes gliding out into the wet layer of the medium [13], to his use of a thinly primed and coarsely woven jute fabric that impeded fluent brushwork to a large extent [18]. Alternatively, there are many known examples where van Gogh simply painted a new picture on top of an existing one in order to reuse the canvas, perhaps allowing the rough texture of the underlying painting to play through to the surface of the current picture to a fair extent. Also significant is whether brushstrokes were applied onto a wet paint surface (so-called wet-on-wet technique), or onto one that was already dry (wet-on-dry) since each method of paint application leads to a quite different look, as illustrated in two of the details in Fig. 1. Van Gogh's pictures tend to show predominantly, but not exclusively, wet-on-wet brushwork, helping to create the highly sculptured surfaces and color marbling effects that we generally associate with his mature technique.

Description of the Dataset and Analysis Tools

The dataset provided by the Van Gogh and Kröller-Müller Museums consists of high resolution grayscale scans of existing Ektachrome films of 101 paintings, scaled (via bi-cubic resampling) to a uniform density of 196.3 dots per painted-inch and digitized to 16 bits per channel [14]. Of the 101 paintings, 82 have consistently been attributed to van Gogh (vG), 6 have always been known to be non-vG, and 13 have been or are currently questioned by experts.

The results presented below were obtained with several types of wavelet transform. Wavelets can capture local features in a wide range of scales. That this is potentially useful in the present context is suggested by the variety of different scales used in the art historians' visual inspection of the brushwork in a vG painting, as described above.

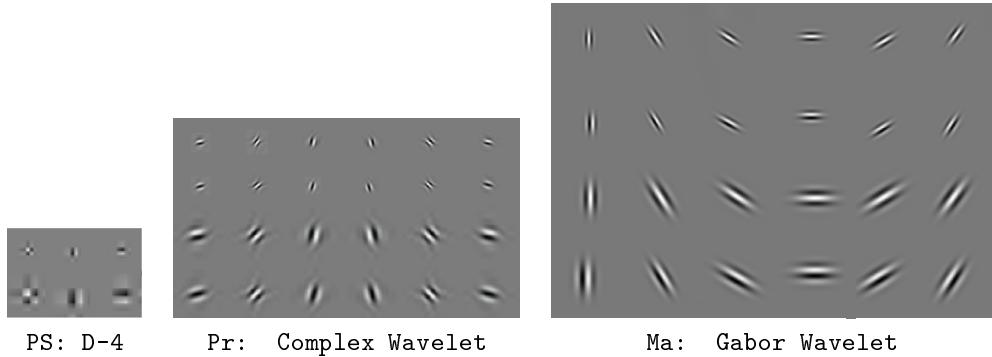


Figure 2: The three different choices for the wavelet templates (shown at the same size); in each case, wavelets are shown for two successive scales, and at all the orientations.

Wavelet transforms for 2D images use templates (*wavelets*) that exhibit different scales and orientations. Arbitrary images can be written as a superposition of such templates located on (a) grid(s), weighted appropriately. There are many ways to choose these templates and associated grids; three different choices are represented below, labeled by the home institution of the authors who have used them for this study. The first approach (Penn State or PS) uses an orthonormal wavelet basis, for which the transform is implemented via fast subband filtering and has the virtue of providing a critically sampled representation of the data; the second (Princeton or Pr) abandons critical sampling (but not subband filtering) in order to obtain a greater orientation selectivity via the use of complex wavelets [15, 27]; the third (Maastricht or Ma) opts for a family of wavelets with the same orientation selectivity as Pr, but with a (Gabor, or modulated Gaussian) wavelet that is close to physiologically suggested templates [11]. Fig. 2 illustrates these different wavelet choices.

Because the individual paintings, at the fine discretization level provided, are very large, they must be divided in patches for analysis purposes. These patches are analyzed individually, and the statistics obtained for the patches are then used to classify the paintings.

Group	Analysis	Patch Size (pixels)	Decomposition	Mathematical Model	Classifier
PS	Texture	512 × 512	D4 Wavelets	2-D HMM	Likelihood
	Geometry	512 × 512	Contours (Edges)	Clustering	Mallows distance
Pr	Style	512 × 512	Complex Wavelets	HMM	MDS
	Fluency	128 × 128		Scale medians	Thresholding
Ma	Texture	256 × 256	Gabor Wavelets	Scale Energies	SVM

Table 1: The basic preprocessing and statistical methods used by the three groups. Abbreviations: HMM: Hidden Markov Models, MDS: Multi-Dimensional Scaling, SVM: Support Vector Machines.

Classification is done in various ways. Both PS and Pr use a Hidden Markov Model (HMM) approach [7, 25]. In a Markov model, variables are correlated to their predecessors (or neighbors); typically the variables can be in several possible states (each with its corresponding probability distribution), and the behavior of the model is characterized by the matrix p , in which each entry $p_{s,s'}$ is the probability that a variable in state s is “followed by” another in state s' . In a HMM, the states are assumed to be “hidden,” meaning that the *transition probabilities* $p_{s,s'}$ and the probability distributions of the observations are unknown for each state, and have to be inferred from the data. These parameters, which characterize the model to some extent, can then be used for classification purposes. Other classification tools used in this study include support vector machines that are trained on histograms of energy values.

Table 1 gives an overview of the different approaches described below: patch sizes, the wavelets used, the highlighted features in the paintings, the models and the classifiers used.

Similarity Assessment via Texture and Brushstroke Geometry Modeling — The work of the Penn State group

For the PS development of statistical models of wavelet-based texture and brushstroke geometry for vG’s paintings, art historians first selected, in the 101-painting dataset, 23 works that are unquestionably by vG, and that represent different periods of his art life and different painting techniques. They are F25 (Coming Out of Church in Neunen), F61v (Self Portrait), F76 (Still Life: Satin Flowers and a Bowl with Leaves and Flowers), F82 (The Potato Eaters), F160 (Peasant Woman in a Red Bonnet), F206 (Head of a Woman: Nearly Full Face), F208 (Self Portrait with Pipe: Three Quarters to the Right), F215 (Nude Study: Little Girl Seated), F234 (Still Life: One-Eared Vase with Aster and Phlox), F256 (Mussels and Shrimp), F260 (Houses in Antwerp), F261 (View of Paris), F266a (Factories Seen from a Hillside in Moonlight), F292 (Boulevard de Clichy), F293 (The Banks of the Seine), F309a (Undergrowth), F316 (View from Montmarte), F344 (Self Portrait in a Gray Felt Hat: Three Quarters to the Left), F371 (Japonaiserie: The Flowering Plum Tree), F423 (A Bugler of the Zouave Regiment), F522 (Self Portrait in Front of The Easel), F632 (The Plow and the Harrow), and F779 (Crows in the Wheatfields). The painting ID numbers are based on the original attempt, published in 1928, at compilation of all of van Gogh’s works by J-B. de la Faille, revised in 1970 in *The Works of Vincent van Gogh: His Paintings and Drawings* (Reynal and Company). Thumbnails of the 101 paintings are available in [14].

These 23 paintings serve as training set for the PS analysis; every other painting is then compared with each of the 23 training paintings. A low average distance (in a sense defined below) indicates a measure of stylistic proximity to the work of vG.

Note that only vG examples are used in this training process. This differs from standard classification methods, where distinction is learned from both positive and negative examples. Not using negative (non-vG) examples in the training process makes this a more highly stringent assessment of the proposed methods.

Next, the methods to compare paintings, specifically, to compute the distances, are described.

Feature Extraction

The paintings are divided into patches of about 512×512 pixels, corresponding to roughly $2.5'' \times 2.5''$ on the canvas. In order to ensure that patches from the same painting are of the same size, the number of pixels on a side of the patch is set to vary between 400 and 600 pixels. These patches are the basic elements for comparing paintings. Distance (or dissimilarity) measures are defined between patches using both texture-based and stroke-based features. Once the patch-wise distances are obtained, the distance between two paintings or between a painting and a collection of paintings as a whole is computed by aggregating the patch-wise distances. The aggregation will be described in the next subsection.

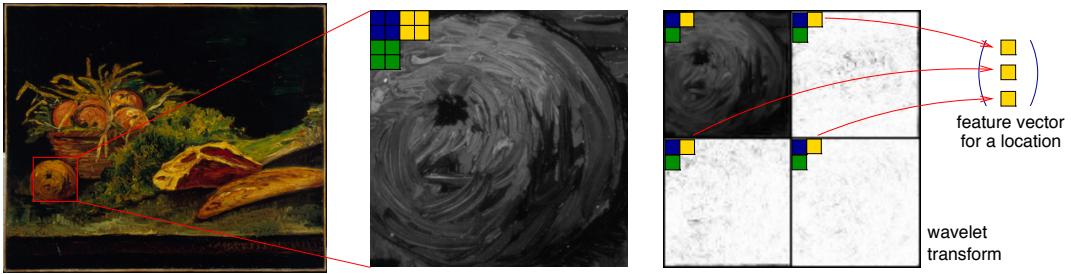


Figure 3: The formation of the texture feature vector based on the wavelet coefficients. The vG painting F219 (Still Life: Basket with Apples, Meat and a Breadroll) is shown.

To compare painting patches, the PS analysis extracts two types of features for every patch. Wavelet-based texture features are extracted from the D4 orthonormal wavelet transform of the pixel intensities [10]. The coefficients resulting from a wavelet transform reflect different orientations and abruptness of variation in an image, and are widely used as low-level texture features [9]. For the PS texture features, only the wavelet coefficients at the finest scale were used; coefficients corresponding to the three different orientations (see Fig. 2) at each grid location are grouped into one 3D-vector. Note that there is a spatial correspondence between the wavelet coefficients and the original pixels. As shown in Fig. 3, a 2×2 block of pixels $[(2i, 2j), (2i+1, 2j), (2i, 2j+1), (2i+1, 2j+1)]$ in the original image is roughly captured by four wavelet coefficients each located at position (i, j) in one of the four frequency bands. Three of the four coefficients, which locate in the three fine scale frequency bands, form the texture feature vector for their corresponding block in the original image. These feature vectors are statistically dependent, especially among nearby grid nodes. In addition to the marginal distributions of the texture feature vectors at individual grid positions, this spatial dependence among feature vectors can contribute to distinguishing textures, and is therefore taken into account in the PS probabilistic models for the texture feature vectors.

The second type of features includes geometric characteristics of strokes. Strokes are higher level characteristics of a painting and are more directly perceived by viewers than the texture features. It is, however, extremely challenging to locate strokes accurately from grayscale images in a fully automated manner (Fig. 4). Using an edge detection based method developed to trace the contours of strokes, several geometric features are computed, such as the length, orientation, and average curvature for each stroke contour line. Without loss of generality, consider one specific feature and denote it by x_i , where $i = 1, \dots, k$ is the index for extracted brushstroke contour lines. A probabilistic model is built based on each feature set $\{x_i, i = 1, \dots, k\}$, where x_i 's are treated as independent samples from the same distribution.

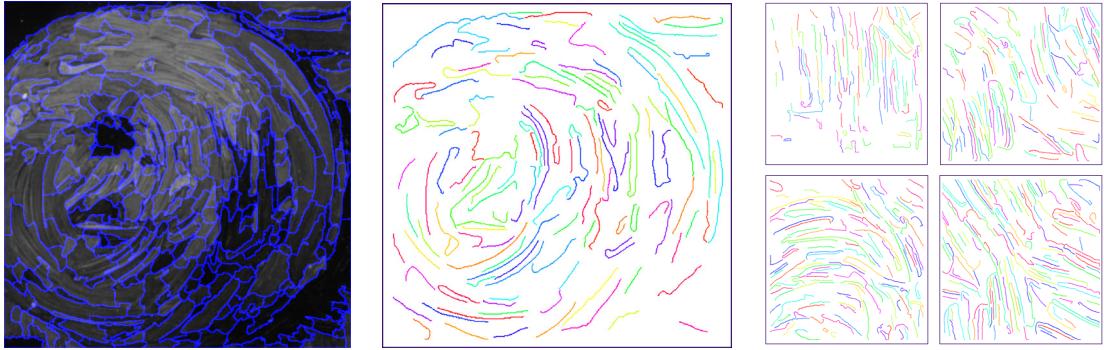


Figure 4: Segmenting brushstrokes in a grayscale image is highly challenging due to the intermingling nature of the brushstrokes. Left: A conventional region segmentation algorithm, based on the K-Means clustering, was unsuccessful in locating individual brushstrokes on this portion of a vG painting. Middle: We developed an edge detection based algorithm to trace the individual brushstrokes. Right: The patterns of the segmented brushstrokes can potentially be analyzed using fluid dynamics principles.

Stochastic Model Based Comparison

As mentioned earlier, statistical models are constructed for both texture-based and stroke-based features. The model for texture features is a spatial model, specifically, a variation of the 2-D Hidden Markov Model (HMM) [16]; for stroke features i.i.d. (independent and identically distributed) models are used. The 2-D HMM models an array of vectors $\{u_{i,j}, i = 1, \dots, n, j = 1, \dots, m\}$ by assuming an underlying spatial state process $\{s_{i,j}, i = 1, \dots, n, j = 1, \dots, m\}$. Because it is difficult to flexibly model spatial dependence among continuous random variables, the states are introduced to “discretize” the dependence. The state process is assumed to follow a Markov mesh (a discrete causal Markov random field) [1], an extension of the Markov chain from 1-D (a sequence) to 2-D (a grid). Since the model is ultimately for continuous vectors $u_{i,j}$, a Gaussian distribution is assumed for $u_{i,j}$ when $s_{i,j}$ is given. Moreover, to retain a tractable model, $u_{i,j}$ is assumed to be conditionally independent from other $u_{i',j'}$ given $s_{i,j}$.

Spatial relationship is less of a concern for the stroke features because they capture relatively large-scale characteristics in an image. Consider one stroke feature, e.g., average curvature. Let the set of features for a patch be $\{x_i, i = 1, \dots, k\}$. We form a discrete distribution for the patch by quantizing the x_i 's. Here, we used the Lloyd algorithm or K-Means clustering to partition the feature space and calculate the centroids of each partition. Suppose after quantization, the x_i 's are divided into K groups, each with an average feature value z_j . The proportion of x_i 's in the j th group is p_j , $j = 1, \dots, K$. We then form a discrete distribution $\{(z_1, p_1), (z_2, p_2), \dots, (z_K, p_K)\}$, where z_j 's are the support points of the distribution and p_j 's are the corresponding probabilities. We adopt a simple discrete distribution here because the number of stroke lines k varies dramatically among patches. The number of support values, K , is dynamically determined by the closeness between the x_i 's. We use the Mallows distance [21] for the discrete distributions.

To aggregate the patch-wise distance into image-wise distance, we employ the following scheme. Let I_1 be the test image and I_2 be the reference image. For every patch P in I_1 , the patch in I_2 that is closest to it is found, and the associated distance is recorded for P . The average of these distances across all the patches in I_1 is taken as the distance from I_2 to I_1 . A similar technique for merging patch-wise distances is used when comparing one image with several images as an entity. When the number of patches in the reference images is large, for any patch P in I_1 , one can also find the k closest patches to it and compute the average of the k distances as the recorded value for P .

Experimental Results

After removing the 23 paintings of the training set, 78 other paintings remain in the dataset, including all the paintings that have at one time been, but are no longer, attributed to vG, as well as the 6 paintings that have always been identified as produced by vG's peers. The distance between each of the 78 images and the training set as an entity is then computed; the 78 paintings are ranked according to these distances.

The PS group has developed an online system that can compare any painting against any subset of

training paintings, using any or all types of the features extracted¹. For instance, when the vG paintings F632 (The Plow and the Harrow) and F779 (Crows in the Wheatfields) are used as training, the system indicated that the test paintings F304 (Pont de la Grande Jatte), F418 (The Sea at Saintes-Maries, a Wacker forgery)² and F511 (Orchard in Blossom) all resemble the styles of the two training paintings according to the standard deviations of the brushstroke curvatures. However, when compared using the lower level texture features, F418 is shown to be unlike the two training paintings, while F511 is shown to be similar to the two. As another example, when a pointillism painting F309a (Undergrowth) is used as the only training painting, the system identified F341 (View from Vincent's Room in the Rue Lepic), upper portion of F734 (The Garden of Saint Paul's Hospital), F314 (Voyer-D'Argenson Park at Asnières), and F342 (Interior of a Restaurant) as paintings of a similar pointillism style based on brushstroke lengths. In contrast, F482 (Vincent's Bedroom in Arles) and F260 (Houses in Antwerp) were selected as those differentiating the most from pointillism. Space limitations prevent a more extensive description of the results obtained.

Some of the results, based only on the texture modeling, are illustrated in Fig. 5, which displays the 5 paintings (among the 78) that are closest to the 23-painting training set, as well as the 5 paintings furthest from the training set. Among the five least similar ones, three are non-vG; the other two, F267 (Self Portrait: Three Quarters to the Left) and F572 (Willows at Sunset), are deemed authentic; among the most similar ones, S225V (Portrait of Van Gogh Painting the Sunflowers) is by Gauguin, while the other four are true vGs. The 23 vG paintings, which serve as the sole source of knowledge in the computerized learning process, may not be sufficiently representative for covering the range of brushstroke styles of vG. Further, these inconsistencies within the sets of the most and least similar images underscore the need for refinement of our comparison methods.



Figure 5: Paintings ranked as most similar (top row) and least similar (bottom row) to vG's brushstroke styles learned from the 23 training images using wavelet-based texture features. F314 (Voyer-D'Argenson Park at Asnières), S225V (Portrait of Van Gogh Painting the Sunflowers), F511 (Orchard in Blossom), F464 (Vincent's House on the Place Lamartine, Arles), F555 (The Pink Orchard), F572 (Willows at Sunset), S218V (View from Montmartre), F253 (Still Life: A Bottle, Two Glasses and a Plate of Bread), F267 (Self Portrait: Three Quarters to the Left), S251V (Vase with Flowers).

¹The URL of the online system is available to researchers upon request.

²A Wacker forgery is a fake van Gogh painting commissioned or sold by the German art dealer Otto Wacker (1898-1970).

Characterizing Scales at which Telling Details Emerge — The work of the Princeton group

The Pr analysis focuses on 76 paintings from the dataset: 65 from vG’s Paris period onward, the 6 consistently attributed to other artists, and 5 (of the 13 questioned ones) formerly considered vGs but since definitively deattributed. The 17 remaining vGs (early period) are excluded because they are very dark; their digitization is concentrated in too small a portion of the full range of grayscale values to provide useful information for the Pr analysis. Fig. 6 illustrates the Pr decomposition of a vG painting, using wavelets with 6 orientations.

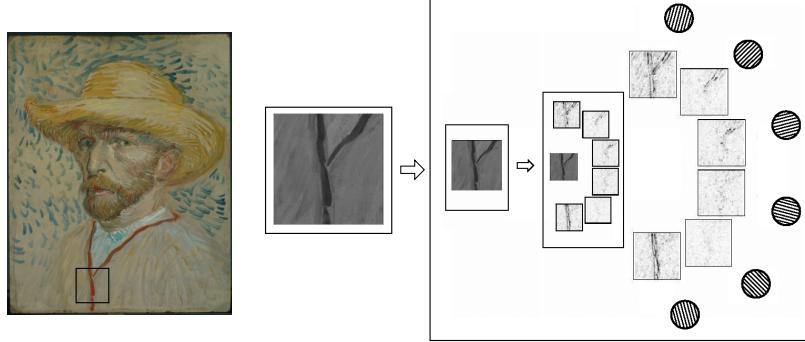


Figure 6: One patch in F469 (Self Portrait with Straw Hat) by vG, and two layers of its wavelet transform, along the six orientations, and its subsampled low-pass version. (Wavelet coefficients with larger/smaller magnitudes in darker/lighter gray.) For each scale s , there are 4 times more wavelet coefficients than for the next coarser scale ($s - 1$).

Data Analysis and Feature Extraction

To analyze each 512×512 patch p , its wavelet coefficients are modeled by a Hidden Markov Tree (HMT, a special kind of HMM). In this model, each coefficient is associated with a hidden state, taking one of 2 values, “edge” or “non-edge,” indicating whether the coefficient’s wavelet template overlaps an edge in the image. All the coefficients of scale s and orientation α with hidden state “edge” (resp. “non-edge”) are then modeled by a zero-mean Gaussian distribution, with a variance $\sigma_{s,\alpha}^{\text{edge}}$ (resp. $\sigma_{s,\alpha}^{\text{non-edge}}$) that will likely be large (resp. small). Dependencies between coefficients with the (scale,orientation) pairs (s, α) and $(s - 1, \alpha)$ at the same location are modeled by the 2×2 matrix $\epsilon_{s,s-1}^\alpha$ of transition probabilities between hidden states. Transitions between the two states occur often; for example, a smooth gradient between solid regions corresponds to an “edge” state at coarse scales and a “non-edge” at finer scales. HMT models of this type were used successfully for the characterization and separation of texture in images [25]. For each patch p , the 4 model parameters (2 transition probabilities, 2 variances) for each (s, α) (108 in total) are learned from p , and combined into a feature vector $\mathbf{v}^{[p]} \in \mathbb{R}^{108}$.

Examination of the Brushstroke Scale: Van Gogh’s Style Emerges?

The 108 features are ranked and then renumbered according to their effectiveness in distinguishing vG and non-vG patches. The features that dominate in this ranking (and thus come first after renumbering) are systematically transition probabilities from non-edge to edge states (from coarser to finer scales), identifying the scales at which detail information “emerges,” as one gradually zooms in, in vG more so than in non-vG paintings. These characteristic scales turn out to be different for different orientations; their identity seems characteristic for vG’s style.

Setting the “ m -feature vector” $\mathbf{v}^{[p|m]} \in \mathbb{R}^m$ of a patch p to be the truncation of the feature vector $\mathbf{v}^{[p]}$ to its first m (renumbered) entries, an “ m -similarity distance” between paintings is defined by adding, for all pairings of a patch p of one painting with a patch p' of the other, the weighted distances $\mathbf{d}_m(p, p') =$

$\left[\sum_{\ell=1}^m w_\ell \left| \mathbf{v}_\ell^{[p|m]} - \mathbf{v}_\ell^{[p'|m]} \right|^2 \right]^{1/2}$ between their m -feature vectors, where the weight w_ℓ of the ℓ -th feature is proportional to its effectiveness in the vG versus non-vG classification process.

A multidimensional scaling algorithm is used to find the arrangement of 76 points in 3D space (with every point representing one painting) most in accordance with the pairwise m -similarity distances between the 76 paintings; in this representation, the center C_{vg} of the vG cluster is determined (all vGs weighted equally). Even for small values of m , this 3D representation gives a visualization with good separation between vGs and non-vGs in the dataset; see Fig. 7. This visualization is an effective tool in the communication between the technical team and art historians.³ The separation is made quantitative by using a radius classifier: for

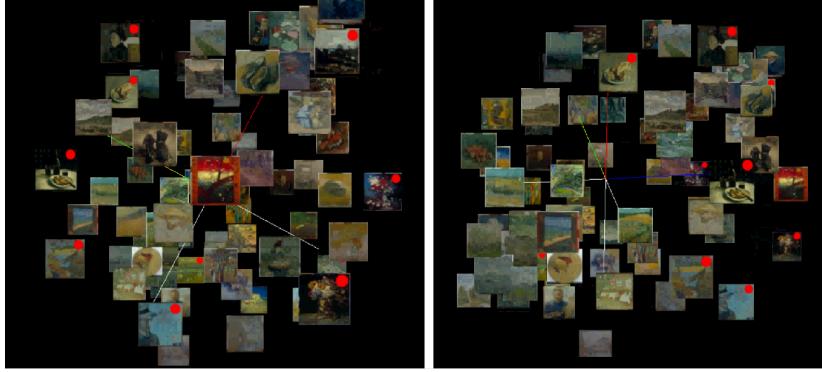


Figure 7: Two views (different angles) of the 3D representation of the 76 paintings obtained for $m=1$. The 11 non-vG paintings (among the 76 paintings considered by Pr) have red dots; they are systematically further from C_{vg} than most vGs.

an appropriate radius r , points closer to C_{vg} than r overwhelmingly represent vGs; points further from C_{vg} than r more likely correspond to non-vGs. Additionally, stylistically similar vG paintings tend to cluster in this analysis, with stylistically less typical vG paintings tending toward non-vG regions; the results of this analysis can thus be interpreted as a characterization of a painting's style. Testing this classifier through leave-one-out cross-validation (i.e. classifying each painting according to C_{vg} and r calculated from the other 75 paintings) shows that it generalizes well to new examples. In a different classification scheme, boosting a distance classifier of the 16-feature vectors leads to even better results under cross validation. (See Table 2 for all numerical results).

	radius classifier		generalization success	
	case $m = 1$	case $m = 2$	radius cl. $m = 1$	boosted distance cl. $m = 16$
vG closer to C_{vg} than r (out of 65)	51	55	50	54
non-vG further from C_{vg} than r (out of 11)	9	9	8	8

Table 2: Results for the Pr classifiers: radius classifier for a 3D realization constructed from the pairwise m -similarity distances, for $m = 1, 2$; success at generalization (under leave-one-out cross-validation) for radius classifier and a boosted distance classifier based on the 16-feature vectors.

Examination of Extremely Fine Scale: Copies Exposed

To pinpoint paintings, such as copies or forgeries of true vGs, that are stylistically similar but are not authentic, much finer scales in the wavelet decomposition are used, on patches of 128×128 pixels. The Ma group had observed earlier [3] that the infamous Wacker forgeries of vG paintings had many more large-valued wavelet coefficients than true vG paintings. Wavelet transforms typically use fewer coefficients at

³A movie of this spatial arrangement can be found in <http://www.digitalpaintinganalysis.org/Princeton/demos.htm>.

coarse than at fine scales; in particular, there are, in the Pr wavelet transform, 4 times as many coefficients in each scale as in the next coarser scale (see Fig. 5). It follows that the two finest scales contribute a fraction of $15/16 = 93.75\%$ of the total number of all the wavelet coefficients (for all scales combined) for each patch; a surfeit of large-valued wavelet coefficients is thus possible only if there is an over-abundance of large coefficients *at the finest scales*. This can (tentatively) be attributed to more hesitant brushstrokes, caused by a reduction in motion fluidity when copying another painting or another painter’s manner.

To quantify this, the Pr analysis measures the median wavelet coefficient strength at each of the two finest scales; this does indeed bring out copies and forgeries in the dataset from most of the authentic, original vGs: they have an excess of features of width 0.25-0.5mm (2-4 pixels only), at the very limit of the spatial resolution in the dataset. Out of all 76 paintings, 6 have their wavelet coefficient medians in the top 8 for both the finest scale S and the second-finest ($S - 1$).

In particular, the one known deliberate forgery in the dataset (F418 - The Sea at Saintes-Maries) places 2nd at scale S and 4th at scale ($S - 1$). More intriguingly, the other 5 (of the 6) include 1 copy by vG himself (F634 - Sheep Shearers, after Millet), and 2 paintings in which, experimenting with technique, he had traced over his own brushstrokes again after the paint had dried (F652 - Fir Woods at the Fall of Day and F734 - The Garden of Saint Paul’s Hospital); the lack of fluency has therefore a natural explanation for them as well. This observed feature of copies of (and by) vG warrants further study. The remaining 2 outliers are F742 (A Clump of Trees in the Garden of Saint Paul’s Hospital), which shows a very strong canvas artifact at the extra-fine scale, and F377 (Still Life: Sunflower), a small study with atypical brushwork, in which vG was studying color more closely than form.

Biologically Inspired Painting Analysis — The work of the Maastricht group

Painting Analysis with Gabor Wavelets

The purpose of the digital analysis is to compare (parts of) paintings in terms of perceptually meaningful similarities. Although perceptual meaningfulness is a subjective concept, it can be measured with a computer by using techniques derived from biological and psychological models of the human visual system. The Ma analysis is guided by the following three principles: (1) contours are important, (2) images must be analyzed at multiple scales, and (3) similarities between paintings are reflected in the local texture (i.e., patterns of brushstrokes).

These principles are implemented by convolving the paintings with multi-scale oriented Gabor wavelet filters [11] and by histogramming the resulting coefficients. The Gabor wavelet filters come in pairs; $G_{even}(x, y, \sigma, \alpha, \omega)$ and $G_{odd}(x, y, \sigma, \alpha, \omega)$ are, respectively, the real and imaginary parts of the function $e^{2\pi i\omega(x\sin\alpha+y\cos\alpha)} e^{-\frac{x^2+y^2}{2\sigma^2}}$, where x and y are the spatial coordinates assuming the filters are centered at the origin, α sets the spatial orientation and ω the spatial frequency of the filter pair. Each pair of filters “responds” to the presence of an intensity transition in a particular orientation-scale combination. We apply a set of filters that covers a range of orientations and scales: six orientations ($\alpha = \frac{k}{6}\pi$, for $k \in \{0, 1, 2, 3, 4, 5\}$), and four scales (ω and σ set to values so that the smallest and largest filters roughly match the smallest and largest brushstrokes). Convoluting these filter pairs with an image yields a decomposition of the image structure into “energy values,” one energy value for each pixel, orientation, and scale. The Gabor wavelet energy value of an image patch is defined as the sum of the squared values obtained by convolving both components with an image patch.

Fig. 8 is an example of an image patch (left) and the responses of the filters (right; dark to bright indicate low to high energy). Each row represents a scale, from fine (top row) to coarse (bottom), each column an orientation from horizontal (left column) with steps of 60 degrees counterclockwise to almost horizontal (right). The total image energy of each orientation-scale filter, obtained by summing the energy values for the entire patch, is given on top of each sub-image. The largest image energy is associated with the horizontal orientation that matches the main orientation of the strokes in the image patch (first column). The finest scale of analysis as displayed in the top row corresponds to strokes or other visual structure with a diameter as small as 2-3 pixels (0.3-0.4 mm).

Intuitively speaking, measuring the total energy by simply summing the energy values of all constituent patches of a painting (one of the simplest forms of analysis) corresponds to counting the contours (light-dark

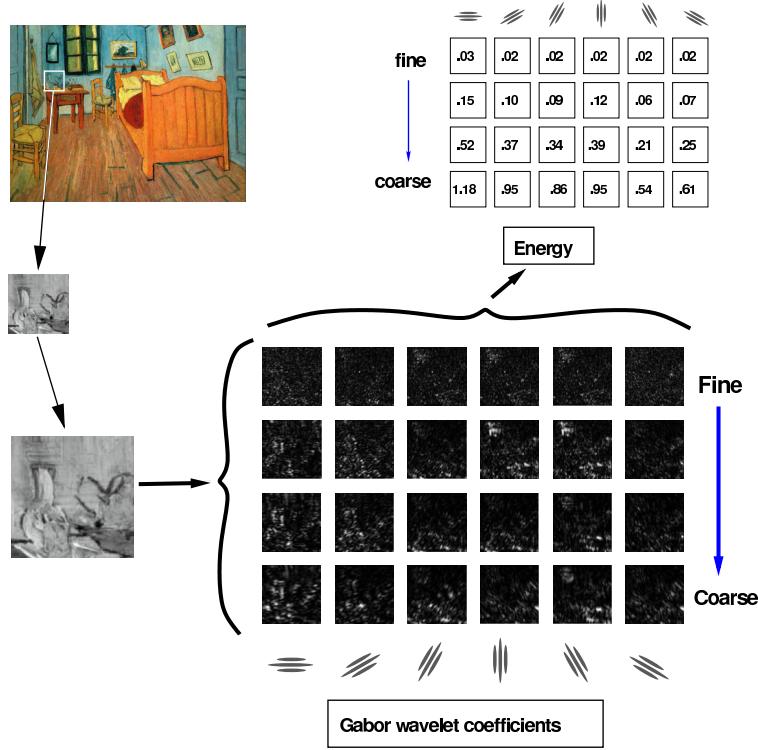


Figure 8: The image patch on the left, taken from F482 (Vincent’s Bedroom in Arles), is decomposed by the Gabor wavelet transform into 4 scales and 6 orientations visualized on the right, where the rows represent scales (top to bottom: fine to coarse) and the columns orientations (left to right: counterclockwise from horizontal to almost horizontal). The numbers above each scale-orientation square give the energy value in each orientation-scale combination.

transitions) in the painting. By selecting the appropriate scale of analysis, i.e., the very fine scale at which the Gabor wavelet filters highlight (parts of) brushstrokes, we can make a computational analysis of the brushstrokes in a painting. More visible brushstrokes give rise to more contours, leading to large energy values. Even this straightforward analysis technique makes the well known Wacker forgery F418 (The Sea at Saintes-Maries) pop out in the analysis, as illustrated in Fig. 9, showing the energy values for 5 paintings, including F418. Large energy values, possibly indicative of more hesitant brushstrokes, are characteristic of copies and forgeries of vG paintings; as in the Pr analysis, a few true vGs are found to stand out in the Ma analysis as well.

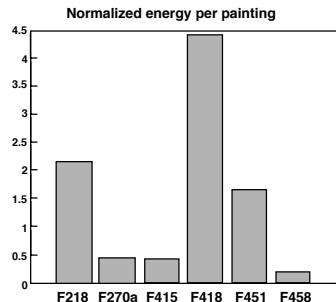


Figure 9: Normalized energy for six paintings, one of which is a forgery, measured with Gabor wavelet filters. The Wacker forgery F418 (The Sea at Saintes-Maries) clearly stands out in terms of energy.

A more informative representation of the paintings is obtained by creating a multidimensional histogram that captures information on the configuration of spatial frequencies and orientations within a patch. In such a histogram, the $24N^2$ energy values obtained for a patch of $N \times N$ pixels are aggregated in 4×6 bins, one for each scale-orientation combination.

Using such multi-dimensional histogram representations of paintings, the **Ma** analysis induces a model by performing automatic classification experiments with a Support Vector Machine or SVM (an automatic learning technique for the binary classification of input vectors [8]). The SVM requires a training set consisting of input vectors and labels indicating their class. In this case, the input vectors are the 24-dimensional-vector histograms; the labels represent the authorship of the painting from which the histograms were extracted, i.e., vG and non-vG . The SVM learns to map the input vectors from the training set onto their appropriate classes. After training, the SVM is used on test paintings, i.e., paintings outside the training set.

The leaving-one-out validation procedure is employed for testing the generalization performance of the trained SVM, which trains the classifier on the histograms of all but one painting, and subsequently tests the SVM on the histograms of the one remaining painting; the training and testing are repeated for all combinations and the test results are averaged. This means that for 101-painting dataset, 101 experiments are performed in which the SVM is trained on 100 paintings and tested on a single painting. The SVM generates a label for each patch of the test painting; the overall label for the test painting is then the most frequently returned label.

Results

Using the multidimensional histograms in combination with SVMs on the full data set, four out of the six non-vG paintings were detected, at the cost of wrongly classifying two vG paintings.

These results suggest that the **Ma** method can detect dissimilarities in the brushstroke texture of paintings, and could therefore support art experts in their assessment of the authenticity of paintings. More subtle differences require more advanced approaches. The **Ma** group has begun to apply its techniques to color reproductions of vG's paintings [4], the automatic determination and validation of brushstroke orientation [5], and is developing and testing advanced approaches that analyze configurations of brushstrokes.

Concluding Remarks

Sophisticated image processing in painting analysis is becoming possible, as high resolution and richer data are becoming available. The signal processing algorithms performing such analysis are in the early stages of development; this paper summarizes the results obtained by several groups using wavelet decompositions of the same dataset, consisting of 16-bit grayscale digitized representations of very high spatial resolution (196.3 dpi) of 101 paintings, mostly by vG. All groups obtained encouraging but not perfect results. Using a wider range of signal analysis tools, starting from a richer representations of the paintings (including color or multispectral information), and more nuanced mathematical models (possibly reflecting the more subtle visual assessment described in the first section), better results can certainly be achieved. The number of interested researchers, targeted conference sessions, and specialist workshops devoted to development and evaluation of signal processing algorithms for painting analysis, in particular in support of artist identification, an enduring task of art historians and conservation specialists, is growing. It is an exciting time for this emerging area.

Acknowledgments

The authors thank the Van Gogh and Kröller-Müller Museums for granting access to such a rich dataset, and for their continuing support of this cross-disciplinary interaction. They acknowledge the constructive comments of the anonymous reviewers. J. Li and J. Wang would like to thank their former student Weina Ge for assistance in implementing the geometric analysis and conducting the related experiments. National Science Foundation Grants IIS0347148 and EIA0202007 provided partial funding for their research. S. M. Hughes and E. Brevdo thank Peter Ramadge for many interesting discussions and suggestions. I. Daubechies gratefully acknowledges partial support of NSF grants DMS0245566 and DMS0354464. The research of I. Berezhnoy and E. Postma was carried out within the Netherlands Organization for Scientific Research (NWO) ToKeN project Authentic (grant 634.000.015). C. R. Johnson, Jr. was supported in part by a Fulbright Fellowship (2005) and a Stephen H. Weiss Presidential Fellowship (2005-2009).

Authors

C. Richard Johnson, Jr. (johnson@ece.cornell.edu) was born in Macon, GA in 1950. He received a PhD in Electrical Engineering from Stanford University, along with the first PhD minor in Art History granted by Stanford, in 1977. He joined the Cornell University faculty in 1981, where he is a Stephen H. Weiss Presidential Fellow and Professor of Electrical and Computer Engineering. His previous principal research interests have been adaptive feedback systems theory (1977-1991) and blind equalization algorithm analysis and creation (1991-2005). At the start of 2007, Johnson accepted a 5-year appointment as an Adjunct Research Fellow of the Van Gogh Museum in Amsterdam to facilitate the interaction of art historians and conservation specialists with algorithm-building signal processors.

Ella Hendriks (Hendriks@vangoghmuseum.nl) graduated in Art History at the University of Manchester, U.K., in 1982. In 1986 she completed the post-graduate training in the Conservation of Easel Paintings at the Hamilton Kerr Institute, University of Cambridge, U.K. From 1988 to 1999 she was Head of Conservation at the Frans Halsmuseum in Haarlem, The Netherlands. From 1999 up to the present day she is Head of Conservation at the Van Gogh Museum in Amsterdam, and in 2006 gained a PhD 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.

Igor J. Berezinoy (Igor.Berezinoy@authentiq.com) has an MSc degree in Applied Mathematics and Machine Learning. During the past four years he has been working on a PhD thesis on artist identification. Currently, he is employed as a Research Scientist at Philips Research Europe, in Eindhoven, The Netherlands.

Eugene Brevdo (ebrevdo@princeton.edu) is a graduate student in Electrical Engineering at Princeton University. His research interests include signal processing and machine learning, with applications to computer vision and inverse problems in medical imaging. His work is supported by the NDSEG and Gordon Wu fellowships.

Shannon M. Hughes (smhughes@princeton.edu) is a PhD student in Electrical Engineering at Princeton University. Her thesis research focuses on the development of mathematical methods for the analysis of complex, and possibly very high-dimensional, data, with applications to problems in both neuroscience and art history. She has been a recipient of an NSF Graduate Fellowship, a Princeton University Gordon S. Wu Fellowship, and an NIH Ruth L. Kirchstein National Research Service Award.

Ingrid Daubechies (icd@princeton.edu) is the William R. Kenan, Jr. Professor of Mathematics at Princeton University, where she has been since 1994. Her research interests include applications of time-frequency analysis in a wide range of fields, many related to electrical engineering. She is an IEEE Fellow and a member of the National Academy of Sciences. She has been the recipient of the Eduard Rhein Foundation Basic Research Award, a MacArthur Fellowship and the National Academy of Sciences Medal in Mathematics.

Jia Li (jiali@stat.psu.edu) is an Associate Professor of Statistics and by courtesy appointment in Computer Science and Engineering at The Pennsylvania State University, University Park. She received the MSc degree in Electrical Engineering, the MSc degree in Statistics, and the PhD degree in Electrical Engineering, all from Stanford University. She worked as a Visiting Scientist at Google Labs in Pittsburgh from 2007 to 2008, a Research Associate in the Computer Science Department at 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.

Eric Postma (Postma@MICC.unimaas.nl) is a Professor in Artificial Intelligence at Maastricht University, the Netherlands. His main research interests are machine learning and biologically-inspired image processing. Together with his group he works on adaptive active vision, manifold learning, and cognitive models of visual perception. His application-oriented work focuses on the development of techniques to support experts in the domains of the cultural heritage and medicine.

James Z. Wang (jwang@ist.psu.edu) is a Visiting Professor at the Robotics Institute, Carnegie Mellon University. He has been on faculty at the College of Information Sciences and Technology of Penn State since 2000. His research interests are image database retrieval, computational aesthetics, image tagging, and biomedical informatics. He has been a recipient of an NSF Career award and the endowed PNC Technologies Career Development Professorship. Wang holds a summa cum laude Bachelor's degree from University of Minnesota, and an MSc in Mathematics, an MSc in Computer Science, and a PhD in Medical Information Sciences, all from Stanford University.

References

- [1] K. Abend, T. Harley, and L. Kanal, "Classification of binary random patterns," *IEEE Trans. on Information Theory*, vol. 11, no. 4, pp. 538-544, 1965.
- [2] M. Barni, A. Pelagotti, and A. Piva, "Image processing for the analysis and conservation of paintings: Opportunities and challenges," *IEEE Signal Processing Magazine*, vol. 22, no. 5, pp. 141-144, September 2005.
- [3] I. E. Berezhnoy, E. O. Postma, and H. J. van den Herik, "Computerized visual analysis of paintings," In *Proc. Int. Conf. Assoc. for History and Computing*, pp. 28-32, September 2005.
- [4] I. E. 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.
- [5] I. E. Berezhnoy, E. O. Postma, and H. J. van den Herik, "Automatic extraction of brushstroke orientation from paintings," *Machine Vision and Applications Journal*, 2008, to appear.
- [6] L. Carlyle, "HART (historically accurate paint reconstruction techniques) final report: 2002-2005," Chapter 3, conducted within the De Mayerne Programme, supported by NWO (Netherlands Organisation for Scientific Research), 2005.
- [7] H. Choi, J. Romberg, R. Baraniuk, and N. G. Kingsbury, "Hidden Markov tree modelling of complex wavelet transforms," In *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Istanbul, Turkey, 2000.
- [8] N. Cristianini and B. Schölkopf, "Support vector machines and Kernel methods: the new generation of learning machines," *AI Magazine*, vol. 23, no. 3, pp. 31-41, 2002.
- [9] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, 2008, to appear.
- [10] I. Daubechies, *Ten Lectures on Wavelets*, SIAM: Society for Industrial and Applied Mathematics, June 1992.
- [11] J. G. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Signal Processing*, vol. 36, no. 7, pp. 1169-1179, July 1988.
- [12] E. Hendriks and M. Geldof, "Van Gogh's Antwerp and Paris picture supports (1885-1888): Reconstructing choices," *Art Matters: Netherlands Technical Studies in Art History*, vol. 2, pp. 39-75, 2005.
- [13] E. Hendriks and L. van Tilborgh, *New Views on Van Gogh's Development in Antwerp and Paris: An Integrated Art Historical and Technical Study of His Paintings in the Van Gogh Museum*, PhD thesis, Faculty of Humanities, University of Amsterdam, vol. 1, p. 132, and vol. 2, p. 62, cat. 72, November 2006.
- [14] C. R. Johnson, Jr., ed., *Proc. 1st Int. Workshop on Image Processing for Artist Identification*, Amsterdam, the Netherlands, May 2007.
<http://digitalpaintinganalysis.org/workshop/ProceedingsIP4AI-1.pdf>
- [15] N. G. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *Applied and Computational Harmonic Analysis*, vol. 10, no 3, pp. 234-253, 2001.
- [16] J. Li, A. Najmi, and R. M. Gray, "Image classification by a two dimensional hidden Markov model," *IEEE Trans. on Signal Processing*, vol. 48, no. 2, pp. 517-33, February 2000.
- [17] J. Li and J. Z. Wang, "Studying digital imagery of ancient paintings by mixtures of stochastic models," *IEEE Trans. on Image Processing*, vol. 13, no. 3, pp. 340-353, March 2004.
- [18] K. H. Lister, C. Peres, and I. Fiedler, "Tracing an interaction: Supporting evidence, experimental grounds," In D. W. Druick and P. K. Zegers, *Van Gogh and Gauguin: The Studio of the South*, Thames and Hudson, pp. 354-369, 2001.

- [19] S. Lyu, D. Rockmore, and H. Farid, “A digital technique for art authentication,” *Proc. National Academy of Sciences*, vol. 101, no. 49, pp. 17006-17010, December 7, 2004.
- [20] H. Maitre, F. Schmitt, and C. Lahanier, “15 years of image processing and the fine arts,” In *Proc. IEEE Int. Conf. on Image Processing*, vol. 1, pp. 557-561, October 2001.
- [21] C. L. Mallows, “A note on asymptotic joint normality,” *Annals of Mathematical Statistics*, vol. 43, no. 2, pp. 508–515, 1972.
- [22] T. Melzer, P. Kammerer, and E. Zolda, “Stroke detection of brush strokes in portrait miniatures using a semi-parametric and a model based approach,” In *Proc. Int. Conf. on Pattern Recognition*, pp. 172-174, August 1998.
- [23] J.-P. Mohen, M. Menu, and B. Mottin, *Mona Lisa: Inside the Painting*, Abrams, New York, 2006.
- [24] R. Pickvance, *Van Gogh in Arles*, Metropolitan Museum of Art, New York, 1984.
- [25] J. Romberg, H. Choi, and R. Baraniuk, “Bayesian tree-structured image modeling using wavelet-domain hidden Markov models,” *IEEE Trans. on Image Processing*, vol. 10, no. 7, pp. 1056-1068, July 2001.
- [26] R. Sablatnig, P. Kammerer, and E. Zolda, “Hierarchical classification of paintings using face- and brush stroke models,” In *Proc. Int. Conf. on Pattern Recognition*, pp. 172-174, August 1998.
- [27] I. W. Selesnick, “A new complex-directional wavelet transform and its application to image denoising,” In *Proc. International Conference of Image Processing*, vol. 3, pp. 573-576, 2002.
- [28] M. van Bommel, M. Geldof, and E. Hendriks, “An investigation of organic red pigments used in paintings by Vincent van Gogh (November 1885–February 1888),” *Art Matters: Netherlands Technical Studies in Art*, vol. 3, p. 123 (figs. 8-9), 2005.
- [29] M. M. Van Dantzig, *Vincent? A New Method of Identifying the Artist and His Work and of Unmasking the Forger and His Products*, pp. 24-25, Amsterdam, 1952.
- [30] M. Yelizaveta, C. Tat-Seng, and R. Jain, “Semi-supervised annotation of brushwork in paintings domain using serial combinations of multiple experts,” In *Proc. ACM Multimedia*, pp. 529-538, October 2006.