Empirical Mode Decomposition Analysis for Visual Stylometry

James M. Hughes, Dong Mao, Daniel N. Rockmore, Yang Wang, and Qiang Wu

Abstract—In this paper we show how the tools of empirical mode decomposition (EMD) analysis can be applied to the problem of "visual stylometry," generally defined as the development of quantitative tools for the measurement and comparisons of individual style in the visual arts. In particular we introduce a new form of EMD analysis for images and show that it is possible to use its output as the basis for the construction of effective support vector machine-based stylometric classifiers. We present the methodology and then test it on a collections of two sets of digital captures of drawings: a set of authentic and well known imitations of works attributed to the great Flemish artist Pieter Bruegel the Elder (1525 – 1569) and a set of works attributed to Dutch master Rembrandt van Rijn (1606–1669) and his pupils. Our positive results indicate that EMD-based methods may hold promise generally as a technique for visual stylometry.

Index Terms—Empirical mode decomposition, stylometry, classifier, image processing.

1 Introduction

Classification plays an important role in the evaluation of a work of art: our knowledge that a painting is by Picasso changes our entire conception of its value, artistically, art historically, culturally, and monetarily. When dealing with works of art, classification in this sense is known as "attribution" or "authentication." Traditionally, scientific investigation for the purpose of attribution of works of art has been the domain of technical art historians, who employ the tools of materials science. These tools enable an examination of primary physical evidence derived from the work. For example, studies of the media can be performed to determine if they are consistent with a known working style of the artist. The pigments and paper can be analyzed and through methods like carbon dating, the time period of the work can be determined. Tools such as infrared and x-ray analysis enable us to look at information that lies beneath the surface of the work. While mathematics plays a role in these analyses through simple applications of well known techniques such as differential equations, it is ultimately secondary to physics.

- *J.M. Hughes is with the Department of Computer Science, Dartmouth College*
  Email: hughes@cs.dartmouth.edu
- *D. Mao is a senior engineer at Brion Technology Co.*
  Email: dmao@math.msu.edu
- *D.N. Rockmore is with the Departments of Computer Science and Mathematics, Dartmouth College, and The Santa Fe Institute*
  Email: dnrockmore@gmail.com
- *Y. Wang is with the Department of Mathematics, Michigan State University.*
  Email: ywang@math.msu.edu
- *Q. Wu is with the Department of Mathematical Sciences, Middle Tennessee State University.*
  Email: qwu@mtsu.edu

Contact Author: Yang Wang (ywang@math.msu.edu)

Physical examination can accomplish only so much, after which we must turn to visual information. Historically, visual information has been assessed using the techniques of "connoisseurship," a process by which a questioned work is subjected to evaluation by a few experts who are steeped in the work and life of the artist of interest. The analyses are usually based on the experts' extensive visual experience and encyclopedic knowledge of the career of the artist in question, as well as other kinds of art historical data. In short, the connoisseur is an expert in the "style" of the artist and the period, and has acquired an ability to identify this style and distinguish it from others.

While there has long been a notion of "scientific connoisseurship," in the sense that this discipline possesses a principled underlying methodology (see e.g., [6], [39]), it is only recently, with the advent of high-resolution digitization of works of art and the efforts that many museums and scholars are making to digitize large collections (e.g., *Artstor* – www.artstor.org or Google's *Art Project* – www.googleartproject.com) as well as individual works (for the purposes of conservation), that we find ourselves at a point where it may be possible to apply a variety of new mathematical tools to the analysis of art, and in particular to the problem of the quantification of artistic style.

Although relatively new to the visual arts, the problem of style quantification is one that has a long history. Its roots lie in the search for methods for quantifying literary style, which date at least as far back as the 1854 ruminations of the mathematician Augustus De Morgan [27]. In 1897 the term *stylometry* was coined by the historian of philosophy, Wincenty Lutaslowski, as a catch-all for a collection of statistical techniques applied to questions of authorship and evolution of style in the literary arts (see e.g., [31]). Today, the quantification of writing style (perhaps more broadly considered a sub-discipline of text classification) is a central area of research in statistical learning (see e.g., [9]).

Literary stylometry benefits from the fact that there is generally some fundamental agreement on the basic atom of analysis – the word. Even some musical stylometry benefits from the use of the unit of the note (see e.g., [25] and also [34], which develops a stylometry of performance based on phrasing). However, in the case of visual media, while all analysis ultimately begins with pixel information, a much wider range of first-order analytic tools are used to extract features (and create classifiers) that begin to get at stylistic characteristics. Examples include box counting analysis, Fourier spectra, wavelet spectra, etc. (see e.g., [12], [16]). The different approaches appear to articulate different stylistic elements so that in short, there is much less agreement on the fundamental analytic elements in stylometry. As a result, the search for and development of new analytic tools for classification is ongoing.

It is in the spirit of this continuing effort that we introduce a new mathematical tool, *empirical mode decomposition* (EMD), for visual stylometry. EMD is a highly adaptive scheme that serves as a powerful complement to Fourier and wavelet transforms. The technique was originally introduced in a one-dimensional (time series) setting for the purpose of analyzing the *instantaneous frequency* of signals [10]. Since then EMD has found many successful applications in a diverse collection of subjects, ranging from biological and medical sciences to geology, astronomy, engineering, and others ([3], [5], [10], [11], [22], [32]). Of particular relevance to this paper are applications to texture analysis and Chinese handwriting recognition ([41], [42], [43], [44]), indications of the suitability of EMD for the analysis of a medium like drawings.

In these varied examples EMD shows itself to be an effective tool for analysis in situations in which distinguishing features cannot be captured fully by spectral or wavelet techniques. EMD shares with wavelets the quality of being a multiscale technique, but one in which information at different scales is captured by so-called *intrinsic mode functions* (IMFs) that are constructed in an adaptive (i.e., data-driven) manner. The adaptive multiscale nature of EMD analysis serves as a strong motivation for finding a way to apply this framework to two-dimensional data such as images.

As previously used, EMD seems intrinsically one-dimensional, depending as it does on the analysis of the so-called "instantaneous frequencies" or *Hilbert-Huang transforms* [10] of the IMFs, a concept that does not lend itself to a natural higher dimensional generalization. However, a formulation of EMD via *iterative filtering* solves this problem [21]. Iterative filtering EMD offers several advantages over classical EMD in the context of visual stylometry: it is easily applicable to higher dimensional data, it is stable under small perturbations, it is completely flexible in degrees of localization which may vary from artist to artist, and it is easily implemented to capture directional features (in arbitrary directions). The last two properties are crucial for analyzing strokes (or "marks" as they are called in art conservation and analysis).

The multiscale nature of iterative filtering EMD connects it to several other techniques that have proved to be useful for visual stylometry. This includes the fractal-based authentication and multifractal analysis of the works of Jackson Pollock (see e.g., [38], [37], [36], [17], [18] as well as some of the critiques). Wavelet-based techniques have been successfully applied to authenticating drawings of Pieter Bruegel the Elder [19], [23] and to many of the approaches used in the "Van Gogh Project," an effort directed toward the quantification of the style of Vincent van Gogh (see [16] for a survey of the techniques used in the project). Recent successful uses of adaptive methods in stylometry include the technique of sparse coding [13].

Here we demonstrate that the EMD framework, which is a hybrid of multiscale and adaptive techniques, is useful in the context of visual stylometry. We describe the methodology behind EMD for two-dimensional data and then show how it may be used as the basis of a stylometric classifier by applying it to two visual art datasets: one a well-known set of digitizations of drawings, all of which were at one time attributed to the great Flemish artist Pieter Bruegel the Elder, and the other a set of pen and ink drawings attributed to Rembrandt van Rijn and his pupils. The latter is analyzed by digital means for the first time in the present work. Our results indicate that this relatively simple analytic tool produces a structured decomposition of the images that provides useful summary data for stylistic classification.

While we use the language of attribution (i.e., *authentic* vs. *imitation*) in our classification results, the goal in stylometry research generally is to uncover analytic tools that can be used to quantify stylistic similarity. In comparison, in the best case, sincere judgements of authenticity in general are multi-faceted, taking into account a wide spectrum of evidence, with the "proof" of authenticity resembling much more a careful courtroom argument than either a mathematical proof or even purely statistical argument (see e.g., [6]). It is worth noting that attribution, or at least the process of ascribing some probability to a work of art that it was executed by a particular artist is but one use of the extraction of quantitative characteristics of style. Other potential uses include the analysis of the evolution of an artist's style or a process for a quantitative form of comparison of the working styles of different authors. The successful construction of a classifier built from the EMD output (or any analysis mechanism) is a positive first step for demonstrating that the derived features might be useful for these more nuanced investigations.

The paper is organized as follows. In Section 2 we introduce the iterative filtering formulation of EMD. This has a natural multidimensional extension whose application to visual stylometry we address in Section 3. Section 4 contains the Bruegel analysis, using the data set examined in [13] and [23] as well as some newly included drawings related to the Bruegel corpus. In Section 5 we apply these techniques to a set of drawings generally attributed to Rembrandt van Rijn and his pupils. We close with a summary and some ideas for future directions of this work.

2 Empirical Mode Decomposition – The iterative filtering approach

The original *empirical mode decomposition* (EMD) for one-dimensional data (e.g., time series data) is a highly adaptive method that produces a multiscale decomposition of the data. It serves as an alternative to the more traditional Fourier series and wavelet decomposition and was motivated primarily by the need for an effective way to analyze the *instantaneous frequency* of signals. Using the so-called *sifting algorithm* EMD decomposes the input signal into a trend function plus a finite, often small, number of components called *intrinsic mode functions* (IMFs) that oscillate about 0. For complete details on EMD, IMFs and the sifting algorithm see [10].

The ability of EMD to create adaptive multiscale decompositions suggests that a two-dimensional form of it might be useful for image classification generally, and for visual stylometry in particular. This is made possible via a formulation of EMD via *iterative filtering* [21], which admits a natural multidimensional generalization. The approach differs from the classical EMD in its choice of the sifting algorithm used to obtain the IMFs. In iterative filtering we begin with a convolution kernel A so that for an input signal $X(t)$ (for $t \in \mathbb{R}^d$ or $\mathbb{Z}^d$) we define the associated convolution operator

$$\mathcal{L}_A(X) = A * X.$$

We generally pick A so that $\mathcal{L}_A(X)(t)$ represents some kind of moving average of $X(t)$. The new "sifting" operator is now defined as

$$\mathcal{T}_A(X) = X - \mathcal{L}_A(X) \tag{2.1}$$

and the associated sifting algorithm computes the limit of iterating $\mathcal{T}_A$

$$\lim_{n\to\infty} \mathcal{T}_A^n(X).$$

For suitable choice of the kernel A the iteration will converge (see below). To obtain the IMFs for X using the sifting algorithm, we choose a suitable A_1 apply the sifting algorithm with $\mathcal{T}_{A_1}$ to obtain the first IMF

$$I_1 = \lim_{n\to\infty} \mathcal{T}_{A_1}^n(X).$$

Subsequent IMFs are then obtained by an iterative process: having derived the first $k-1$ IMFs $(I_1, \ldots, I_{k-1})$, we remove $I_1, \ldots, I_{k-1}$ from the data and then choose a new convolution kernel A_k in order to produce

$$I_k = \lim_{n\to\infty} \mathcal{T}_{A_k}^n(X - I_1 - \cdots - I_{k-1}).$$

The process stops when $Y = X - I_1 - \cdots - I_m$ has at most one local maximum or local minimum. The essential difference between the classical EMD and the iterative filtering EMD lies with the choice for the "averaging" operator. The iterative filtering EMD uses a filter $\mathcal{L}_A$ while the classical EMD uses the average of two cubic splines that envelop $X(t)$. The filter approach provides us with far more flexibility and it is clearly not constrained by the dimensionality of the data X.

Choice of kernel (mask). The multiscale nature of the iterative filtering EMD derives from a judicious choice of kernels. The A_k generally bear some relation to one another, effectively defining the degree of "localization" of the derived IMF I_k. We call these kernels the *masks* or *footprints* of the EMD. Intuitively, if we choose A_1 to have very small support then only the finest details will be captured by the IMF I_1. With larger support for A_1 we may capture the lower frequency components of the signal in addition to the high frequency components. The same goes for other masks and their associated IMFs. In this way, range of support of the masks for the EMD algorithm leads to a multiscale decomposition of the signal X.

The ability to choose the masks, particularly in terms of their sizes and "shapes," (i.e., the extent of their localization or support) is one of the main advantages of the iterative filtering EMD for the analysis of visual stylometry. The classical EMD has no such flexibility. With iterative filtering EMD, if two signals are decomposed using identical masks, we can expect the k-th IMFs for the two signals to capture information in the same frequency range, allowing us to compare "apples to apples."

Convergence and the use of double averaging. As mentioned, it is possible that the iteration of $\mathcal{T}_A$ may not converge. For example, let Ω be a centrally symmetric set in $\mathbb{R}^d$ or $\mathbb{Z}^d$ and let $B(t) = \frac{1}{|\Omega|}$ for $t \in \Omega$ and $B(t) = 0$ otherwise. Then $B * X$ represents the average filter over a neighborhood in the shape of Ω. If we simply choose $\mathcal{L}(X) = B * X$ and $\mathcal{T}(X) = X - \mathcal{L}(X)$ the iterative filtering does not converge. However, if we take $\mathcal{L}(X) = B * B * X$, i.e., the *double average filter* of B, then the iteration of $\mathcal{T}(X) = X - \mathcal{L}(X)$ does converge. This is true in general (see [21] as well as [40]).

Thus, in order to avoid questions of convergence, for visual stylometry analysis we will use double average filters exclusively. Let Ω be any centrally symmetric set in $\mathbb{R}^d$ or $\mathbb{Z}^d$. We shall use D_Ω to denote the double average kernel. Note that for digital image analysis the domain is $\mathbb{Z}^2$.

3 EMD for Stylometry Analysis in Art

As applied to visual stylometry, we see EMD providing an effective tool for a structured multiscale decomposition of a given artwork whose output can be used effectively for classification. In the "EMD methodology" we perform several independent EMDs using masks of different shapes to capture mark information pertaining to such characteristics as strength, orientation, and spatial frequency. The resulting information provides the basic data for the construction of feature vectors useful for analysis and classification.

In particular we perform five different EMDs (corresponding to different filter shapes) to obtain five kinds of decompositions for an image. These five EMDs are performed using double average masks D_Ω where Ω varies among the shapes of square, horizontal line segment, vertical line segment, and two (perpendicular) diagonal line segments. Each mask is centered at the origin. For each EMD using a given shape, we vary the support thereby obtaining IMFs at different scales.

For example, for the double average mask over squares we let Ω_1 have support over a 5×5 square centered at the origin. This yields I_1, the first IMF. For I_2 we choose Ω_2 to have support over the 11×11 square centered at the origin. In general, the length of a side of the support Ω_k is one plus twice that of Ω_{k-1}. The same goes for the other masks, i.e., for a line of a given orientation, the support of the mask at scale k is of length two times the length used for scale $k-1$, plus one. Of course, in general the sizes can be adjusted depending on the resolutions, sizes and types of the images. See Table 1 for the various initial masks.

Figures 2 and 3 show the EMDs of a drawing by Pieter Bruegel the Elder (see Figure 1) using the square and vertical masks. Each figure contains the first four IMFs. The smaller sized masks capture finer details in the strokes and vice versa. Thus, each EMD shares the common trait with a wavelet decomposition that both are multiscale decompositions. By using a square mask the EMD captures stroke details in a uniform way like a shotgun. The line segment masks are designed to capture stroke information *in a given direction*. In general, a line mask is used to highlight strokes in the orthogonal direction. For example, a horizontal line mask will be used to capture vertical strokes.

Such information may be important for stylometric analysis as different artists may exhibit preferences for and subtle characteristics in certain directions. One advantage of this EMD approach is its flexibility in choosing the masks. If for instance we need more directional analysis of strokes we can easily add a line mask of a given direction.

Fig. 1. *Path through a Village*, Pieter Bruegel the Elder, ca. 1552. No 5 from [30].

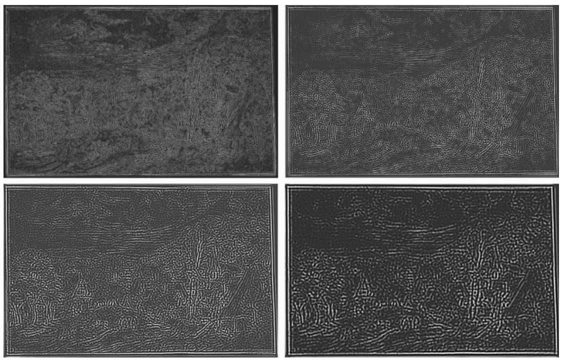


Fig. 2. The first four IMFs of the above Bruegel drawing (Figure 1) using square masks.

The associated IMFs provide the underlying data for the feature vectors associated with the images. From these, classifiers can be designed to address questions pertaining to visual stylometry. In the next section we illustrate this approach with a study of drawings, all of which at one time have been attributed to Pieter Bruegel the Elder.

4 EMD Stylometry Analysis for Bruegel

In this section we apply the EMD framework to a stylometric analysis of a corpus of drawings by the great Flemish artist Pieter Bruegel the Elder (1525–1569) and some well known imitations. Bruegel was famous in his lifetime and thus it is very likely that his style would have been imitated by aspiring artists for training purposes and working artists for business

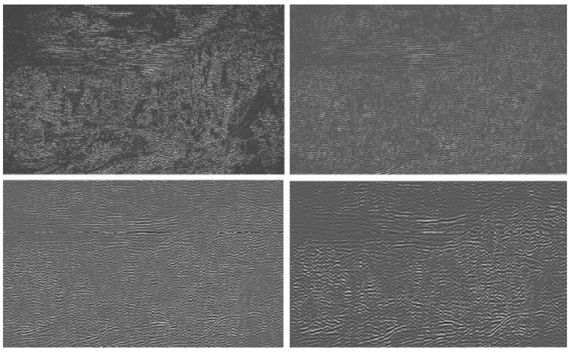


Fig. 3. The first four IMFs of the above Bruegel drawing (Figure 1) using vertical line masks.

MMA Cat. No.	Title	Artist
3	Pastoral Landscape	Bruegel
4	Mountain Landscape with Ridge and Valley	Bruegel
5	Path through a Village	Bruegel
6	Mule Caravan on Hillside	Bruegel
9	Mountain Landscape with Ridge and Travelers	Bruegel
11	Landscape with Saint Jermove	Bruegel
13	Italian Landscape	Bruegel
20	Rest on the Flight into Egypt	Bruegel
7	Mule Caravan on Hillside	-
120	Mountain Landscape with a River, Village, and Castle	-
121	Alpine Landscape	-
125	Solicitudo Rustica	-
127	Rocky Landscape with Castle and a River	Savery

Fig. 4. Authentic (top) and imitations (bottom). The first column corresponds to the New York Metropolitan Museum of Art (MMA) catalog number in [30]. Note that "Savery" refers to the contemporaneous Flemish artist Jacob Savery (1545–1602). A dash indicates unknown authorship.

purposes, both legitimate (i.e., creating drawings and paintings in the style of Bruegel) and illegitimate (e.g., creating works like Bruegel and attempting to pass them off as executed by Bruegel). There are still active debates regarding the authenticity of various works attributed to Bruegel [29].

Our analysis consists of two parts. In the first, using an EMD-based approach, we revisit the dataset of images analyzed in [13], [23], [19]. In these studies successful approaches to classification were accomplished using sparse coding [13] and quadrature mirror filters [23], [19]. We also achieve successful classification, but by a different means, showing that strong linear classifiers can be constructed from the EMD output. The use of linear classification in stylometry first appeared in an analysis of Jackson Pollock's drip paintings [14]. In the second part of this section we apply the EMD framework to an augmented set that includes new drawings had at one time been attributed to Bruegel and produce evidence for their attribution. The results indicate that the EMD-based framework shows promise as an effective technique for visual stylometry.

4.1 Analysis of the Bruegel and Bruegel-like landscapes

Data preparation and EMD methodology. In the first part of our study we use an EMD approach on a previously analyzed [13], [23], [19] collection of thirteen landscape drawings, each of which had at one time been attributed to Bruegel. It is now accepted that eight of these belong to Bruegel (i.e., are authentic) while the remaining five are believed (or are known) to be imitations [30] (see Figure 4).

More precisely, the initial data is given by digital scans (at 2400 dpi) of 35mm color slides of the eight secure drawings by Bruegel (catalog numbers 003, 004, 005, 006, 009, 011, 013, and 020) and five accepted Bruegel imitations (catalog numbers 007, 120, 121, 125, and 127 – see [30]).[1] The scans are then converted to one byte greyscale images, so that the shades range from 0 (black) through 255 (white). For the purpose of homogeneity we extract a central 2000×2000 pixel square from each image. To obtain enough samples for classification each 2000×2000 pixel image is further partitioned into 25 400×400 pixel sub-images or "samples," giving us in total 200 samples of Bruegel and 125 samples of non-Bruegel.

EMD is applied to every sample and the first three IMFs are obtained for each of the five mask shapes shown in Table 1. The initial size of these masks is 5×5 pixels and successive sizes (localizations) are produced by doubling the previous size and adding one. The initial masks are shown in Table 1. Note that they are simply the double averages of standard moving average filters. Given three IMFs for each shape, we obtain fifteen IMFs per sample.

Feature vector construction. A single feature vector is constructed for each of the 400×400 pixel samples. The feature vector entries encode a collection of summary statistics based on the sample's fifteen IMFs and is formed as follows:

Step 1. To take into account variations within the sample image we divide each of the fifteen IMFs of the sample into an 8×8 grid of 50×50 pixel "patches." We shall label these patches $I^j_{k,l}$ with $1 \leq k, l \leq 8$

1. Slides were provided courtesy of the New York Metropolitan Museum of Art.

$$\frac{1}{81}\begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 4 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 4 & 2 \\ 1 & 2 & 3 & 2 & 1 \end{pmatrix}$$

square

$$\frac{1}{9}\begin{pmatrix} 1 & 2 & 3 & 2 & 1 \end{pmatrix}$$

horizontal line segment

$$\frac{1}{9}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 1 \end{pmatrix}$$

vertical line segment

$$\frac{1}{9}\begin{pmatrix} 1 & & & & \\ & 2 & & & \\ & & 3 & & \\ & & & 2 & \\ & & & & 1 \end{pmatrix}$$

diagonal line segment

$$\frac{1}{9}\begin{pmatrix} & & & & 1 \\ & & & 2 & \\ & & 3 & & \\ & 2 & & & \\ 1 & & & & \end{pmatrix}$$

diagonal line segment

TABLE 1
The initial EMD masks.

indexing the patch location and $1 \leq j \leq 15$ indexing the IMF.

Step 2. For a patch $I_{k,l}^{j}$ we compute the following nine summary statistics of the pixel values:

1) mean
2) standard deviation
3) skewness
4) kurtosis
5) percentage of "outlier pixels" defined as those that are greater than the mean plus the standard deviation
6) mean of the outlier pixels
7) standard deviation of the outlier pixels
8) skewness of the outlier pixels
9) kurtosis of the outlier pixels

These nine statistics together give us a nine-dimensional vector $u_{k,l}^{j}$. Grouping together $u_{k,l}^{j}$ for each $1 \leq j \leq 15$ yields a 135-dimensional vector

$$U_{k,l} = [u_{k,l}^{1}, u_{k,l}^{2}, \ldots, u_{k,l}^{15}]$$

associated with each patch. Note that the statistics we compute are motivated by earlier work. The basic summary statistics (the first four moments) are the same statistics computed in [23] from the wavelet coefficients derived from the samples. The use of the "outlier" statistics is motivated by their successful application in the EMD-based classification of physiological data [26].

From the 64 vectors $U_{k,l}$ where $1 \leq k, l \leq 8$ we construct three 135-dimensional vectors V_1, V_2, V_3 to serve as an initial summary of the sample. The vector V_1 is simply the mean of the vectors $U_{k,l}$ over $1 \leq k, l \leq 8$. The vector V_2 is the mean of the top 50% of the values in each entry. The vector V_3 is the mean of the top quartile (top 25%) of the values in each entry. Finally, the feature vector of the sample is the 405-dimensional vector $W = [V_1, V_2, V_3]$. Since each drawing is divided into 25 samples, each drawing is now associated with a "cloud" of 25 points in the 405-dimensional space.

Dimensionality reduction and proof-of-concept classification. In order to evaluate the features we extracted from the IMFs from each of the Bruegel and non-Bruegel drawings, we take advantage of the ground-truth knowledge we have about the attribution of the drawings considered. Ultimately, with whatever features we extract from these images, we seek to classify them as either authentic or not authentic, and to have confidence in these features we must first evaluate their efficacy on data with known attributions. In this sense, classifying a drawing as "authentic" indicates that the features are statistically similar to those of a secure set – indeed, this is effectively what the connoisseur also means. In order to do this, we perform a leave-one-out cross validation experiment in which we consider all but one Bruegel drawing's samples and attempt to classify the held-out Bruegel using a support vector machine (SVM) trained on the remaining samples (see [8] for a good survey of these tools and techniques). Note that, because we do not wish to assume any stylistic similarity between the non-Bruegel drawings, we consider the classification of the held-out Bruegel with respect to the remaining authentic Bruegel drawings and each non-Bruegel *individually*, since there is no art historical knowledge that indicates that these imitations were executed by the same hand. Indeed, evidence suggests that they were executed by several different artists attempting to imitate the style of Bruegel.

The procedure is as follows:

Step 1 (Hold out an authentic Bruegel drawing). We begin with 8 authentic drawings, and on the ith iteration of the experiment we hold out all 25 samples from drawing i.

Step 2 (SVM via RFE). Using the method of recursive feature elimination (RFE, see e.g., [7]) we train SVMs to distinguish the remaining authentic drawings from each non-Bruegel. Since we have five non-Bruegel drawings, we train five classifiers in each iteration of the experiment. This leaves us with 175 authentic samples (7 drawings × 25 samples each) grouped as Class I and 25 samples from the non-Bruegel grouped as Class II, per classifier. We then construct a classifier using an SVM to separate Classes I and II in the original 405-dimensional space via a linear decision boundary (hyperplane) C_1. We then remove variables that do not strongly contribute to classification by eliminating the half of the variables that have the lowest weight in this preliminary classifier. An SVM is again trained on the remaining half of the variables to produce a linear classifier C_2. This process of eliminating variables is repeated until no more than 10 variables remain, leaving us with a classifier that utilizes at most the 10 variables most important for distinguishing between the two classes.

For robustness we repeat this process 100 times for each non-Bruegel drawing, obtaining each time a classifier with at most 10 variables. From these classifiers, we then determine the set of variables that appeared at least six times among the one hundred replicates. Note that, as before, we consider each non-Bruegel separately, so we retain a set of "optimal" variables chosen to distinguish between each non-Bruegel and the remaining authentic drawings.

Step 3 (Testing on held-out images). Once the optimal set of variables has been chosen to distinguish the authentic Bruegel drawings from each non-Bruegel (with the ith drawing held out), we train an SVM using only the optimal variables and measure the performance of this classifier in correctly identifying the held-out Bruegel as authentic. Table 2 shows the fraction of correctly identified samples for each held-out drawing with respect to the classifier trained on the remaining Bruegel drawings and each non-Bruegel drawing.

		Held-out drawing							
		003	004	005	006	009	011	013	020
Non-Bruegel	007	1.0	1.0	1.0	1.0	0.64	0.76	1.0	0.92
	120	1.0	1.0	1.0	1.0	1.0	1.0	0.96	1.0
	121	1.0	1.0	1.0	0.96	1.0	0.92	1.0	1.0
	125	0.76	1.0	1.0	1.0	0.92	1.0	1.0	1.0
	127	1.0	1.0	1.0	1.0	1.0	1.0	0.92	1.0

TABLE 2
Fraction of correctly identified samples (out of a total of 25) for each held-out authentic Bruegel drawing (columns) with respect to the classifier trained on the remaining Bruegel drawings and each non-Bruegel individually (rows).

In order to say that a held-out Bruegel drawing was correctly classified as authentic, it should be classified as such with respect to each of the non-Bruegel drawings, meaning that a large fraction of its samples are correctly classified. In Table 2, the results of these experiments show that the held-out drawings were correctly identified as authentic with a high rate of success. This result establishes that the EMD features provide enough information to distinguish between authentic and imitation Bruegel drawings.

Analysis of the new corpus of drawings. The classifiers we have built above are based on analyzing the corpus of drawings previously analyzed in [13], [19], [23]. In effect, this is a proof of concept that the EMD-based approach may be of use for visual stylometry.

As a result, it is of interest to use these classifiers to establish attribution for other drawings from the Bruegel corpus, since attributions sometimes change as new information, both about the artist as well as the works, comes to light. Here we construct a classifier in the same manner as before to examine an additional set of drawings (not previously analyzed by "digital" techniques) with varying histories that, as of today, are attributed to Bruegel. These drawings correspond to the New York Metropolitan Museum of Art (MMA) catalog numbers No. 002, 010, 012, 015, 017, 018, and 019 [30].

It is important to emphasize that our interest in examining these works (and indeed all of the works that we analyze) is not to simply provide a determination of authenticity, but rather to also say something about the stylistic characteristics of the works of art. In particular, it is sensible to ask to what extent Bruegel drawings share commonalities with imitations, or what constitutes the stylistic differences between two known authentic drawings? Ultimately, once we have a framework for studying stylistic differences, we can begin to ask questions that are broader than the authentication question. We can analyze the evolution of an artist's style, or understand what role a particular artist played in the evolution of a stylistic movement. We can even begin to examine the stylistic relationships between works of art separated by thousands of years.

In analyzing this new corpus of drawings, we proceed according to the methodology described above. Each of the new drawings is divided into 25 samples of 400×400 pixels in the original 405-dimensional feature space. Once again, we begin by creating a classifier using RFE, except that we consider all 200 of the authentic Bruegel samples as Class 1 and consider each of the non-Bruegel drawings individually as Class 2, building each time a classifier to distinguish between these two classes. As before, we eliminate half of the variables until we have at most 10 remaining. We repeat this procedure one hundred times and select as the set of optimal variables those that were present in at least six of the one hundred classifiers for each non-Bruegel drawing. Once this set of variables has been obtained, we build five classifiers, one for distinguishing between each non-Bruegel drawing and the entire set of authentic drawings.

We map the new samples into each of the lower-dimensional feature spaces determined by the optimal set of variables for each classifier, and then classify those samples according to each of the classifiers that have been learned. Table 3 shows the fraction of test samples identified as authentic by each classifier trained on a single non-Bruegel drawing's samples. Specifically, this table reports the number of values of the decision function $f_i(v) > 0$ for v, a sample from a drawing from the new corpus, and f_i, the classifier trained to distinguish between the authentic drawings and the ith non-Bruegel drawing.

Because linear support vectors machines do not *guarantee* separability between classes (since such separation may not exist in the feature space considered), it is useful to approach classification from a proba-

		New corpus drawing						
		002	010	012	015	017	018	019
Non-Bruegel	007	0.84	1.0	0.56	0.68	0.56	0.80	1.0
	120	1.0	1.0	0.88	1.0	1.0	1.0	1.0
	121	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	125	0.96	0.44	1.0	1.0	1.0	0.44	0.96
	127	1.0	1.0	1.0	1.0	1.0	1.0	1.0

TABLE 3

Fraction of correctly identified samples (out of a total of 25) for each drawing from the new corpus (columns) with respect to the classifier trained on the authentic Bruegel drawings and each non-Bruegel individually (rows).

		New corpus drawing						
		002	010	012	015	017	018	019
Non-Bruegel	007	0.87	0.95	0.53	0.73	0.60	0.79	0.98
	120	0.99	0.99	0.90	0.99	0.99	0.99	0.99
	121	0.99	0.99	0.99	0.99	0.99	0.97	0.99
	125	0.96	0.48	0.99	0.96	0.95	0.51	0.94
	127	0.99	0.97	0.92	0.99	0.96	0.99	0.99

TABLE 4

Mean posterior probability of belonging to the authentic class for samples from each drawing from the new corpus (columns) with respect to the classifier trained on the authentic Bruegel drawings and each non-Bruegel individually (rows).

bilistic perspective. In particular, we seek a probabilistic method that takes into account the distribution of values of the decision functions f_i for the training data. Using the class conditional distributions $P(f_i(x)\,|\,Class = 1)$ and $P(f_i(y)\,|\,Class = 2)$, $i = 1..5$, for all training datapoints x in Class 1 (authentic drawings) and datapoints y in Class 2 (non-Bruegel drawings), we can, via Bayes' rule, infer an estimate of the posterior distribution over classes [33], [20]:

$$P(Class = 1|f_i(v)) \propto P(f_i(v)|Class = 1)P(Class = 1),$$

where in each case the prior probability of class membership is given simply as the number of training datapoints appearing in that class, and the class-conditional distributions are estimated directly from the training examples. The particular advantage of using this approach to classification is that it takes the distribution of decision values for the training data into account (since in general, there may be misclassified training data) in determining what decision function values constitute high probability of belonging to one class or another. The posterior distribution over class membership is estimated using a sigmoid function (see [33] for details), with parameters A_i, B_i estimated using the training data:

$$P(Class = 1|f_i(v)) = \frac{1}{1 + \exp(A_i f_i(v) + B_i)}.$$

Using this methodology, we estimated the posterior probability of the test images with respect to each of the three classifiers. We let the posterior probability of class membership for a particular unknown be equal to the mean of the distribution over posterior probabilities for each of the 25 samples from that image for each classifier f_i:

$$\frac{1}{25}\sum_j P(Class = 1|f_i(v_j)), j = 1 \ldots 25,$$

for each image v from the new corpus. Table 4 shows the average probability of being authentic for each test image with respect to each of the five classifiers.

This analysis is of course reflective of statistical similarity between the unknowns and the training set – indeed, an unknown image should only be considered definitively authentic if it is highly dissimilar to all of the known non-Bruegels (whose similarity to the authentic drawings is captured by the five classifiers $f_1, \ldots, f_5$). By this reasoning, if we consider the fraction of correctly identified samples for each drawing from the new corpus, we do not have sufficient evidence to attribute the drawings with with catalog numbers 10, 12, and 18 to Bruegel, since for each of these drawings a significant fraction of samples was not correctly identified. The probabilistic results indicate that the same group of drawings could be considered non-Bruegel, since they had a relatively weak probability of belonging to the authentic group of drawings according to one or more classifiers. Typically, however, we do not seek to provide a binary classification (i.e., authentic/inauthentic) of these drawings because of the lack of practicality of such a judgement. Art historians consider a number of factors when determining the authenticity of a work of art, and it is generally more useful to provide a measure of authenticity (such as a probability), rather than a strict classification. Furthermore, a probabilistic approach allows us to measure in some sense the overall stylistic similarity between drawings, which admits a significantly more nuanced interpretation. For example, although a particular drawing might be not be considered authentic by this methodology, high similarity to certain drawings indicates that it possesses similar stylistic tendencies.

5 Analysis of Rembrandt drawings

With these ideas in mind, we examined whether EMD-based features might prove useful in a very different set of drawings, a collection of pen and ink works historically attributed to Rembrandt van Rijn (1606–1669) and his pupils and imitators. The general problem of distinguishing the works of "Rembrandt and his pupils has a long history (see e.g., [2]). The corpus of drawings considered here [35] includes in

No.	Cat. No.	Title	Artist
03	B95	Jacob lamenting the sight of Josef's blood-stained coat	R
04	B97	Christ carrying the cross	R
07	B164	Adam and Eve	R
09	B519	The return of the prodigal son	
10	B541	Jacob and his sons	R
12	B606	Esau selling his birthright to Jacob	R
15	B876	Josef being sold	R
16	B887	Daniel in the lions den	R
17	B912	Josef interpreting the prisoners dream	R
18	B947	David and Nathan	R
20	B A11	The dismissal of Hagar	R?
01	B66	Temptation of Christ	Fl
02	B80	Josef interpreting the prisoners dream	Fl
05	B108	Christ on the cross	E
08	*	Adoration of the shepherds	Fa
11	B564	Esau selling his birthright to Jacob	B
14	B491	Eliezer and Rebecca at the well	Fa
19	B948	David and Nathan	D
06	B148	Mattathias at Modin	?
13	B652	Christ on the cross	?

Fig. 5. "Rembrandt" set (courtesy of P. Schatborn). Authentic (top) and imitations (bottom). The first column corresponds to the numbering in Schartborn's list and is the numbering used in the Tables in this section. The second column corresponds generally to the Benesch catalog raisonee [1], with catalog number BXXX. Label B A11 indicates an unsure attribution to Rembrandt, at Berlin, Kupferstichkabinett; * indicates not in [1], with unknown attribution, at Amsterdam, Rijksmuseum. "R" denotes R. van Rijn (1606–1669), "Fa" denotes C. Fabritius (1622–1654), "Fl" denotes Govert Flinck (1615–1660), "E" denotes G. van edn Eeckhout (1621–1674), "B" denotes F. Bol (1616–1680), "D" denotes W. Drost (1633–1659), and "?" indicates unknown attribution.

total 20 drawings (see Figure 5). Of these, the first ten are known are now assigned to Rembrandt, No. 11 is likely, but there is still some doubt, two are now given to Carel Fabritius (1622–1654), and five drawings have been identified as the work of various students and imitators of the master. Finally, two drawings do not have secure attribution. It is worth noting that these drawings are quite different from those in the Bruegel corpus – they are drawn more freely, with less precision, and with much more variation in mark width and length and are in a different medium.

The drawings were preprocessed in exactly the same way as in the Bruegel experiments, and features were extracted from these drawings in an identical manner. Unlike in the case of the Bruegel drawings, the Rembrandt and Rembrandt imitation drawings were of various sizes, so that the number of data points per drawing was not consistent. In every other way, the feature vectors extracted from the patches were the same as in the initial experiments.

We initially attempted to replicate the SVM with RFE procedure used in the Bruegel experiments, but this classification scheme yielded poor results on the held-out testing data. In addition, the linear SVM classification boundary was typically unreliable for correctly identifying negative training exemplars (i.e., non-Rembrandt data points).

Having extracted features for all drawings, we embedded the drawings in 22 dimensions via classical multidimensional scaling (MDS). This dimensionality was chosen because it accounts for more then 99% of the total variance in the (embedded) data. Having embedded the points in such a manner, we ran the leave-one-out cross-validation experiments again, learning the parameters of a linear SVM (without using RFE to select features since dimensionality reduction had already been performed). In each iteration of the experiment, we held out all samples from a particular known Rembrandt drawing and trained on the rest, with respect to a single non-Rembrandt drawing. As with the Bruegel drawings, this step was replicated for all non-Rembrandt drawings. Only those (known authentic) drawings whose samples were judged to be authentic sufficiently often were said to be authentic. Tables 5 and 6 show the results for this experiment.

Tables 5 and 6 indicate that two drawings (numbers 09 and 15) were not judged similar enough to other authentic drawings to be considered "authentic" for our purposes. The failure of these drawings to be correctly identified may stem from stylistic variations present in them that are not found in the other authentic drawings, or simply because some of the imitation drawings were particularly successful at capturing certain aspects of the master's style. It is also worth noting that, while the authentic drawing numbered 04 failed to be judged as truly authentic according to the usual SVM decision criterion (i.e., that $f(v) > 0$, for a sample v, given classifier f) with respect to non-Rembrandts 01, when performing classification in a probabilistic manner (shown in Table 6), the posterior probability that drawing 04 was authentic, given the classifier trained against non-Rembrandt number 01, was 0.63, a value large enough to safely consider this drawing authentic. This example highlights the usefulness of a probabilistic approach to classification in this context, since the actual distribution of classifier values may vary significantly from what is expected, because we use SVMs that allow for training points to be misclassified.

Although two genuine Rembrandt drawings were not securely attributed to the master using EMD features, the success of this method in correctly identifying 9 of 11 drawings suggests that EMD does provide a useful means of describing artistic style. With this in mind, we performed the leave-one-out cross-validation experiments again, this time exclud-

		Authentic drawing					
		03	04	07	09	10	12
Non-Rembrandt	01	1.00	0.18	1.00	1.00	1.00	1.00
	02	1.00	1.00	1.00	1.00	1.00	1.00
	05	1.00	1.00	1.00	1.00	1.00	1.00
	08	1.00	1.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	0.38	1.00	1.00
	14	1.00	1.00	1.00	0.00	1.00	1.00
	19	1.00	1.00	1.00	1.00	1.00	1.00

		Authentic drawing				
		15	16	17	18	20
Non-Rembrandt	01	0.00	1.00	1.00	1.00	1.00
	02	0.00	1.00	1.00	1.00	1.00
	05	1.00	1.00	1.00	1.00	1.00
	08	0.00	1.00	1.00	1.00	1.00
	11	1.00	1.00	1.00	1.00	1.00
	14	1.00	1.00	1.00	1.00	1.00
	19	1.00	1.00	1.00	1.00	1.00

TABLE 5
Fraction of authentic Rembrandt samples correctly identified as Rembrandt with respect to the classifier trained on each non-Rembrandt drawing respectively.

ing from our analysis the drawings 09 and 15. Despite excluding these drawings, all other authentic drawings were judged as truly authentic with perfect accuracy with respect to the seven classifiers trained in each iteration of the experiment. More to the point, what this indicates is that the EMD-based features do capture basic multiscale commonalities within the drawings by Rembrandt that generally distinguish his style from that of other artists in the dataset.

		Authentic drawing					
		03	04	07	09	10	12
Non-Rembrandt	01	1.00	0.63	1.00	1.00	1.00	1.00
	02	0.99	1.00	1.00	1.00	1.00	0.99
	05	0.94	0.98	1.00	0.95	0.95	1.00
	08	1.00	1.00	1.00	1.00	1.00	0.90
	11	0.95	0.98	1.00	0.69	0.96	0.99
	14	0.83	1.00	1.00	0.32	1.00	0.99
	19	1.00	0.99	1.00	0.84	0.99	1.00

		Authentic drawing				
		15	16	17	18	20
Non-Rembrandt	01	0.49	1.00	0.89	0.99	1.00
	02	0.67	0.99	0.98	1.00	1.00
	05	0.99	0.99	1.00	1.00	1.00
	08	0.22	0.99	0.95	0.99	1.00
	11	1.00	1.00	1.00	0.99	1.00
	14	0.97	1.00	1.00	1.00	0.97
	19	1.00	1.00	0.98	1.00	0.98

TABLE 6
Mean posterior probability of belonging to the authentic class for samples from each authentic Rembrandt drawing with respect to the classifier trained on each non-Rembrandt individually.

Analysis of questionable drawings. Having confirmed that the set of nine authentic drawings is accurately classified using EMD features, we build a linear SVM classifier that discriminates between the entire set of authentic drawings and each non-Rembrandt individually. We then use this to judge whether the two questionable drawings (numbers 06 and 13) are authentic or inauthentic. Current art historical opinion is that these drawings are *not* genuine Rembrandt drawings (Peter Schatborn, personal communication). Our analysis seeks to provide evidence – one way or the other – of these drawings' authenticity or lack thereof. Tables 7 and 8 show the fraction of correctly identified patches and the mean posterior probability of patches being authentic, respectively.

		Questionable drawing	
		06	13
Non-Rembrandt	01	1.0	1.0
	02	1.0	1.0
	05	1.0	1.0
	08	1.0	1.0
	11	0	1.0
	14	0.71	1.0
	19	0	1.0

TABLE 7
Fraction of samples from questionable drawings identified as authentic with respect to each classifier trained on all secure authentic samples and each non-Rembrandt individually. Drawing 13 is judged to be authentic by this method, while drawing 06 is not.

		Questionable drawing	
		06	13
Non-Rembrandt	01	0.99	0.99
	02	0.99	0.99
	05	0.99	0.98
	08	0.99	0.99
	11	0.38	0.99
	14	0.67	0.99
	19	0.21	0.99

TABLE 8
Mean posterior probability of belonging to the authentic class for samples from each questionable drawing with respect to the classifier trained on each non-Rembrandt individually.

These results indicate that drawing 06 is not an authentic Rembrandt drawing, confirming art historical analysis. Indeed, this drawing shares stylistic qualities with authentic drawing 09 (in particular, the use of ink wash shading), which was excluded from this part of the analysis because it failed to be correctly identified as authentic. Details of these drawings are shown in Figure 6. This may partially account for this drawing's exclusion as authentic; however, even when drawing 09 was included in the training set, drawing 06 was still rejected as authentic (results not shown).

Interestingly, drawing 13 was in fact judged to be an authentic Rembrandt drawing, according to our method. The quantitative similarity of this drawing to all of the secure authentic drawings suggests that closer examination of the drawing's style, its materi-

Fig. 6. Two details from authentic Rembrandt (drawing 09, top), *The Return of the Prodigal Son*, and questionable work (drawing 06, bottom), *Mattathias at Modin*, highlighting the similarity of the drawings in their use of ink wash.

als, and its provenance may ultimately reveal it to in fact be a genuine Rembrandt.

6 Conclusion

We present here a novel adaptation of empirical mode decomposition (EMD) for image analysis, based on iterative convolution, for the analysis of artistic style (visual stylometry) in the drawings of Pieter Bruegel the Elder and Rembrandt van Rijn. In particular, we show that the statistics of the derived intrinsic mode functions (IMFs) can be used to construct effective linear classifiers for a test set of drawings by each artist. We have shown that the marginal statistics derived from the IMFs as well as marginals derived from the outliers provide useful representations of the stylistic regularities present in the works of these artists. This brings a new approach to the problem of visual stylometry. Using classifiers based on these features, we supplied evidence toward attributions of works whose attributions have shifted over time. We believe that mathematical analyses of style such as the one presented here will become increasingly important in assisting technical art historians and conservators resolve outstanding questions of authenticity and attribution by providing objective analytical techniques that can be used to supplement current methods.

The flexibility of this approach and this first largely successful application suggest that the EMD approach is a useful technique for visual stylometry, and may also be useful for image classification in general. In future work, we will study in more detail the aspects of style captured by EMD, as well as the application of this method to different types of visual artistic media, such as paintings.

7 Acknowledgements

Thanks to N. Foti for many useful conversations and to P. Schatborn for suggesting the Rembrandt experiments and providing the images. JMH and DNR gratefully acknowledge the partial support of the Samuel H. Kress Foundation.

References

[1] O. Benesch, *The Drawings of Rembrandt: Complete Edition, 6 Volumes*, (enlarged and edited by Eva Benesch), London, Phaidon Press, 1973.

[2] H. Bevers, L. Hendrix, W. W. Robinson, and P. Schatborn, *Drawings by Rembrandt and His Pupils: Telling the Difference*, J. Paul Getty Museu, Los Angeles, CA, 2010.

[3] Q. Chen, N. Huang, S. Riemenschneider and Y. Xu, A B-spline approach for empirical mode decomposition, *Advances in Computational Mathematics* **24** (2006), 171-195.

[4] J. Coddington, J. Elton, D. N. Rockmore, and Y. Wang, Multifractal analysis and authentication of Jackson Pollock's paintings, in *Computer image analysis in the study of art*, D. G. Stork and J. Coddington, eds., vol. 6810, pp. 68100F1–12, IS&T/SPIE, (Bellingham, WA), 2008.

[5] J. C. Echeverria, J. A. Crowe, M. S. Woolfson and B. R. Hayes-Gill, Application of empirical mode decomposition to heart rate variability analysis, *Medical and Biological Engineering and Computing* **39** (2001), 471–479.

[6] C. Ginzburg, Morelli, Freud and Sherlock Holmes: Clues and scientific method, *History Workshop Journal* (1980) 9(1): 5-36 doi:10.1093/hwj/9.1.5

[7] I. Guyon, J. Weston, S. Barnhill and V. Vapnik, Gene selection for cancer classification using support vector machines, *Machine learning* **46** (2002), 389–422.

[8] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*, Springer (2009).

[9] D. I. Holmes and J. Kardos, Who was the author? an introduction to stylometry, *Chance* **16**(2) (2003), 5–8.

[10] N. Huang *et al*, The empirical mode decomposition and the Hilbert spectrum for nonlinear nonstationary time series analysis, *Proceedings of Royal Society of London A* **454** (1998), 903–995.

[11] N. Huang, Z. Shen and S. Long, A new view of nonlinear water waves: the Hilbert spectrum, *Annu. Rev. Fluid Mech.* **31** (1999), 417–457.

[12] J. M. Hughes, D. J. Graham, D. N. and Rockmore, Stylometrics of artwork: Uses and limitations. *Proc. SPIE: Computer Vision and Image Analysis of Art* 7531, (2010), 75310C.

[13] J. M. Hughes, D. Graham, and D. N. Rockmore, Quantification of artistic style through sparse coding analysis in the drawings of Pieter Bruegel the Elder, *PNAS*, vol. 107, no. 4 (2010), pp. 1279–1283.

[14] M. Irfan and D. G. Stork, Multiple visual features for the computer authentication of Jackson Pollock's drip paintings: Beyond box-counting and fractals, *Electronic Imaging: Image processing: Machine vision applications II*, Kurt S. Niel and David Fofi (eds.), vol. 7251, (2009) pp. 72510Q1–11, .

[15] A. J. Izenman, *Modern Multivariate Statistical Techniques: Regression, Classification, and Manifold Learning*, Springer, New York, 2008.

[16] C. R. Johnson, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, J. Z. Wang, Image processing for artist identification, *Signal Processing Magazine, IEEE* **25** No. 4, July (2008), 37–48.

[17] K. Jones-Smith, H. Mathur, and L. Krauss, Drip paintings and fractal analysis *Phys. Rev. E* **79**, 046111 (2009), DOI:10.1103/PhysRevE.79.046111

[18] K. Jones-Smith and H. Mathur, Fractal analysis: Revisiting Pollock's drip paintings, *Nature* **444** (2006), E9 – E10.

[19] G. Leibon and D. N. Rockmore, Steps toward digital authentication, in M. Emmer, ed. *Matematica e cultura 2007*, Springer Italia, Milano, 2007, pp. 107 –118.

[20] H.-T. Lin, C.-J. Lin, and R. Weng, A note on Platt's probabilistic outputs for support vector machines, *Machine Learning* 68 (2007).

[21] L. Lin, Y. Wang, and H. Zhou, Iterative filtering as an alternative algorithm for empirical mode decomposition, *Advances in Adaptive Data Analysis* Vol. 1, No. 4 (2009) 543–560.

[22] B. Liu, S. Riemenschneider and Y. Xu, Gearbox fault diagnosis using empirical mode decomposition and Hilbert spectrum, *Mechanical Systems and Signal Processing* **20** (2006), 718–734.

[23] S. Lyu, D. N. Rockmore, and H. Farid, A digital technique for art authentication, *Proc. Nat. Acad. Sciences*, **101** No. 49, (2004), pp. 17006–17010.

[24] B. Manaris, P. Roos, P. Machado, D. Krehbiel, L. Pellicoro, and J. Romero, A corpus-based hybrid approach to music analysis and composition, *Proceedings of 22nd Conference on Artificial Intelligence (AAAI-07)*, Vancouver, BC, Jul. 2007, pp. 839–845.

[25] B. Manaris, J. Romero, P. Machado, D. Krehbiel, T. Hirzel, W. Pharr, and R.B. Davis, Zipf's law, music classification and aesthetics, *Computer Music Journal* **29** No. 1, (Spring 2005), MIT Press, pp. 55–69.

[26] D. Mao, Y. Wang, and Q. Wu, A new approach for physiological time series, preprint, (2009).

[27] A. de Morgan,"Letter to Rev. Heald 18/08/1851," in *Memoirs of Augustus de Morgan* by Sophia Elizabeth de Morgan with Selections from his Letters, (S. Elizabeth and D. Morgan, eds.), London: Longman's Green and Co., 1851/1882.

[28] J. R. Mureika, C. C. Dyer and G. C. Cupchik, Multifractal structure in nonrepresentational art, *Physical Review E* **72**, 046101 (2005).

[29] N. M. Orenstein, Followers and fakers of Pieter Bruegel, *International Foundation for Art Research Journal*, **6**(3), (2003), pp. 12–17.

[30] N. M. Orenstein, *Pieter Bruegel the Elder*, Yale University Press, New Haven and London, 2001.

[31] A. Pawlowski, Wincenty Lutoslawski - a forgotten father of stylometry, *Glottometrics,* 8 (2004), 83–89.

[32] D. Pines and L. Salvino, Health monitoring of one-dimensional structures using empirical mode decomposition and the Hilbert-Huang transform, *Proceedings of SPIE* 4701(2002), 127-143.

[33] J. C. Platt, Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods, in *Advances in Large Margin Classifiers*, MIT Press (1999).

[34] C. S. Sapp, Comparative analysis of multiple musical performances, in *ISMIR Proceedings* (2007), pp. 497–500.

[35] Provided by P. Schatborn (2008).

[36] R. P. Taylor, R. Guzman, T. P. Martin, G. D. R. Hall, A. P. Micolich, D. Jonas, B. C. Scannell, M. S. Fairbanks, and C. A. Marlow, Authenticating Pollock paintings using fractal geometry, *Nature* 399 (2006), p. 422.

[37] R. P. Taylor, A. P. Micolich, and D. Jonas Fractal analysis: Revisiting Pollock's drip paintings (Reply), *Nature* **444**, (30 November 2006) E10-E11.

[38] R. P. Taylor, A. P. Micholich and D. Jonas, Fractal analysis of Pollock's drip paintings, *Nature* **399**, 422 (1999).

[39] R. Wollheim, Giovanni Morelli and the Origins of Scientific Connoisseurship," In *On Art and Mind: Essays and Lectures*. London: Allen Lane, 1973.

[40] Y. Wang and Z. Zhou, On the convergence of EMD in l^∞, preprint.

[41] Z. Yang, L. Yang, D. Qi and C. Y. Suen, An EMD-based recognition method for Chinese fonts and styles, *Pattern Recognition Letter,* Vol. 27 (2006), 1692 – 1701.

[42] Z. Yang, L. Yang, and D. Qi, Detection of spindles in sleep EEGs using a novel algorithm based on the Hilbert-Huang transform, in *Wavelet Analysis and Applications*, T. Qian, M. I. Vai and Y. Xu eds. *Applied and Numerical Harmonic Analysis*, Birkhauser 2006, pp. 543–559.

[43] Z. Yang, L. Yang, and Y. Y. Tang, Illumination-rotation-invariant feature extraction for texture classification based on Hilbert-Huang transform, preprint.

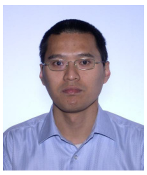
[44] T. Zheng, L. Xie, and L. Yang, Integrated extraction on handwritten numeral strings in form document, *Pattern Recognition and Artificial Intelligence*, In press.

James M. Hughes is a doctoral student in Computer Science at Dartmouth College. His research focuses mainly on the quantification and visualization of artistic & literary style, applied statistics, and network analysis. He received the degree of Bachelor of Arts summa cum laude with honors in Computer Science and honors in German Language & Literature from Kalamazoo College in 2007.

Dong Mao graduated from Peking University, China, got a PhD of Applied Mathematics at Syracuse University, NY. Current is a senior engineer at Brion Technology Co.

Daniel N. Rockmore aniel N. Rockmore-aniel N. RockmoreD received the B.A. degree in mathematics from Princeton University, Princeton, NJ, in 1984, and the M.S. and Ph.D. degrees in mathematics from Harvard University, Cambridge, MA, in 1986 and 1989, respectively. He is currently a Professor of computer science and John G. Kemeny Parents Professor of Mathematics at Dartmouth College, Hanover, NH, where is also Chairman of the Department of Mathematics. He is also on the External Faculty of the Santa Fe Institute. His research interests are in computational harmonic analysis and complex systems. Dr. Rockmore was a 1995 recipient of an NSF Presidential Faculty Fellow- ship, 2005C2007 Sigma Distinguished Lecturer, and the 2008 SIAM I. E. Block Community Lecturer.

Yang Wang received his PhD in mathematics in 1990 from Harvard University under the supervision of David Mumford. From 1989 to 2007 he was a faculty at Georgia Institute of Technology. He is currently a professor in mathematics and head of the mathematics department at Michigan State University. He had also served as a program director at the National Science Foundation in the Division of Mathematical Sciences from 2006 to 2007. His research interests include analysis, wavelets, empirical mode decomposition and their applications to various areas in digital signal processing.



Qiang Wu received his PhD in mathematics in 2005 from City University of Hong Kong. He is currently an Assistant Professor in the department of mathematical Sciences at Middle Tennessee State University. His research interests include statistical learning, data mining, and their applications.