STYLISTIC ANALYSIS OF PAINTINGS USING WAVELETS AND MACHINE LEARNING

Sina Jafarpour, Güngör Polatkan, Eugene Brevdo, Shannon Hughes, Andrei Brasoveanu, Ingrid Daubechies

Princeton University
Departments of Electrical Engineering, Computer Science, and Mathematics
and the Program in Applied and Computational Mathematics
Princeton, NJ 08544

ABSTRACT

Wavelet transforms and machine learning tools can be used to assist art experts in the stylistic analysis of paintings. A dual-tree complex wavelet transform, Hidden Markov Tree modeling and Random Forest classifiers are used here for a stylistic analysis of Vincent van Gogh's paintings with results on two stylometry challenges that concern "dating, resp. extracting distinguishing features".

1. INTRODUCTION

Stylometry, i.e determining a painter's style, is a challenging problem for art historians. Many factors play a role. Technical analyses of the painting, including of pigments present, the materials used and the method of their preparation, the artist's process as documented in the underlayers of the painting (observed through Xray and infrared imaging), etc, provide one type of information. Visual inspection of the painting is of course very important as well, to evaluate and help characterize the visual appearance and style of the work. However, even the sum of all these analyses may prove inconclusive for some works.

A new movement in image processing seeks to use computational tools from image analysis and machine learning to provide an additional source of analysis for such challenging paintings, based on the assumption that an artist's brushwork can be characterized, (at least in part), by signature features (e.g. those arising from the artist's habitual physical movements) and that such distinguishing quantitatively measurable characteristics might be found by machine learning methods and used as an additional piece of evidence in stylometry tasks. Indeed, early attempts in this area have already found considerable success [1, 2, 3].

Recent attempts to characterize paintings of particular style via features discernible by image processing and machine learning algorithms, have often focused on characterizing the statistics of the wavelet coefficients of digital scans of paintings by that artist [1, 4, 5].

This paper uses an approach of this type on a dataset provided by the Van Gogh Museum and the Kroller-Muller Museum in the Netherlands, consisting of high resolution scans of paintings by Vincent van Gogh.

We combine recent image processing and machine learning techniques, in order to tackle two stylometry problems proposed by the two museums: extracting distinguishing features, and a dating challenge. We show how modeling style as a hidden variable, controlling the behavior of the image observables, such as brushstrokes, color patterns, etc, can improve the accuracy of the style analyzer to a significant extent. We use a dual-tree complex wavelet transform [6], that is (almost) shift invariant, to capture quantitatively the effects observable in the image. Next, using Hidden Markov Trees [7], an extension of Hidden Markov Variables, combined with the expectation maximization algorithm [8], we extract the style parameters from the noisy observables. Finally, using standard machine learning techniques, we feed the extracted features to appropriate classifiers, and use the resulting prediction rule for style analysis.

This paper is a sibling of [10], in which similar techniques were used by our team, for authentication purposes instead of stylistic analysis.

2. APPLICATIONS

2.1 Dating Challenge

In the absence of convincing documentation, the dating of a painting is based on where it fits in the chronology of the artist's style, concerning for example, subject matter, materials used, color palette, compositional style, and brushwork. Some undocumented paintings have a mixture of features that seemingly correspond with different periods of their creator's artistic development. Such feature mixes pose difficult dating challenges .

When dating relies on categorizing style and technique issues, computer-based image processing tasks for magnifying the differences in style should prove useful. Furthermore, artificial intelligence and machine learning techniques can provide the right tools for the final decision task.

The dating challenge concerns the dating of paintings by Vincent van Gogh that stem from either his Paris phase (ending early in 1888) or his following late Arles period. The question is to ascertain which features distinguish the two test sets (taking as benchmark the paintings that are unquestionably from the Paris or Arles period), and to use them subsequently to attempt to associate each of the dating candidates with one group or the other.

In distinguishing Van Gogh paintings from these two periods, art historians rely on several general observations regarding shifts in his practice. For instance, small strokes are more prominent in Paris, while brush handling is broader in Arles; colors appear more saturated in Arles due to the filling in of larger areas.

At the initial stage of the challenge, the set of training examples included 33 images each, from the Paris and the Arles periods.

At the final stage, three test paintings were provided. Each test painting exhibits some general features associated with Arles, as well as others associated with Paris. The final goal of this challenge was to come up with a high-confidence

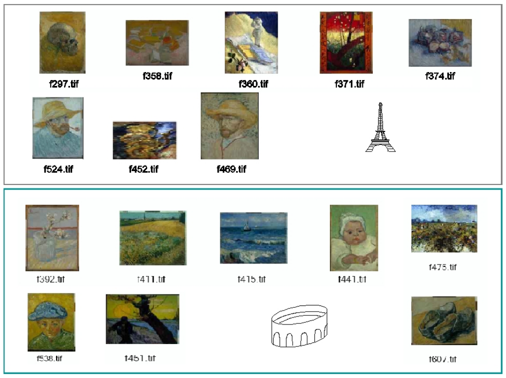

Figure 1: Training set examples for the dating challenge.

classification of test images based on the training images provided. The test paintings for the dating challenge were selected because they represent a real question and have not been easily or consistently dated in the art historical literature. Consequently, digital stylistic analysis provides new and added data that could usefully be taken into account in attempts to date these works more securely.

2.2 Distinguishing Feature Extraction

At the start of Van Gogh's years in Paris he discovered the work of Adolphe Monticelli (1824-1886), in particular his floral still lifes characterized by impasto brushwork. Soon, Van Gogh painted still lifes that reflected Monticelli's brushwork. Art historians have identified several features in the brushwork and color schemes shared by Van Gogh's and Monticelli's floral still lifes, that are absent in the works of their contemporaries. The "distinguishing features extraction" challenge introduced a small set of floral still lifes by Monticelli, Van Gogh, and contemporaries; the task was deducing features discernible by image processing and machine learning algorithms that distinguish Monticelli and Van Gogh floral still lifes, as a group, from the floral still lifes of contemporaries.

In distinguishing styles of painters, art historians rely on several general statements regarding the painting style, including *vigorous brushwork, heavy dark outlines, repetitive, non-overlapping strokes, perspective, brushwork style, color patterns etc*. In Section 4 we show how we can model and magnify some of these features using image processing techniques, and then use the result for final decision making.

The dataset for this challenge, provided by the Kroller-Muller Museum and the Van Gogh Museum, were very high-resolution color scans of the paintings, checkerboarded over part of their surface for security reasons.

3. STYLE ANALYSIS

3.1 Art Historian Style Analysis

Understanding art historians style recognition methods is the first step in providing appropriate automatic style recognition methods. The art history methods should emphasize the key features characterizing the style; moreover, they should be robust against the temporal changes in appearance of the painting due to the deterioration of materials used such as color shift [1]. The art historian then has to combine the noisy observations with his/her previous knowledge in order to come up with a stylometric decision.

Computer-based image processing can assist in this process. Wavelet transforms [6] successfully capture local differences at different scales of images. Appropriate color representation can capture local and global color saturation. Stochastic analysis, often assuming Markov conditions [7], allow extraction of key features of images from the observed wavelet coefficients, despite the noise, and provide robustness. Finally pattern recognition tools [11] provide a variety of different computation classifiers capable of categorizing images based on the extracted features.

3.2 HSL Color Representation

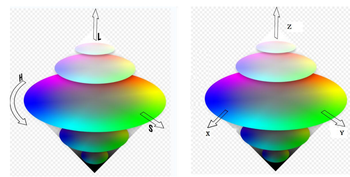

Figure 2: HSL and XYZ color representation domains

In image precessing and computer graphics, HSL is a representation of colors, often provided in an RGB color space, that attempts to describe perceptual color relationships more accurately than RGB, while remaining computationally simple [12]. HSL represents each color by its *hue, saturation*, and *luminance* forming a double-cone or a sphere (with white at the top, black at the bottom, and the fully-saturated colors around the edge of a horizontal cross-section; middle gray is at its center).

The following transformation converts the color representation from the HSL domain to the XYZ domain:

$$\begin{aligned} Z &= L \\ X &= S\cos\left(\frac{2\pi H}{360}\right)\min\{2L, 2(1-L)\} \\ Y &= S\sin\left(\frac{2\pi H}{360}\right)\min\{2L, 2(1-L)\} \end{aligned} \tag{1}$$

In the XYZ domain, the color has not only a Cartesian representation useful in wavelet analysis; it also provides easily readable and valuable information about the saturation and luminance of images, crucial for the style analysis task.

3.3 Dual-Tree Complex Wavelet Transforms

Wavelet transforms [6] separate the details of an image into different scales. For the task of style analysis dealing with very high resolution images, this decomposition permits the extraction of very high resolutional differences, which may not be observable by human eye. The dual-tree complex wavelet transform, applied to the XYZ representation of the images, can detect color patterns well; it is able to separate into details of different orientations, helpful to characterize brushstroke directions; finally, it captures local differences, and is in this respect analogous to the art historians' scrutiny in style analysis.

The dual tree complex wavelet transform [6] comprises two parallel wavelet filter bank trees. Each complex wavelet coefficient can be written as

$$c_i = u_i + jv_i, \quad (2)$$

where the u_i, v_i each constitute DWT coefficients valid in their own right. The magnitudes $|c_i| = \sqrt{u_i^2 + v_i^2}$ are largely shift invariant; they give rise to a more accurate estimate of the image at a given location and scale than the DWT.

The CWT has six subbands of coefficients. Each subband matches the changes in a corresponding direction. Using complex wavelet transforms, we can thus analyze changes in images in six different directions. For each direction, the coefficients form a quad-tree data structure. This means that any coefficient at a coarser level corresponds to exactly four coefficients at the next finer level. Figure 3 shows a two-layer wavelet decomposition of one patch from a self portrait by Van Gogh, along the six orientations (subbands).

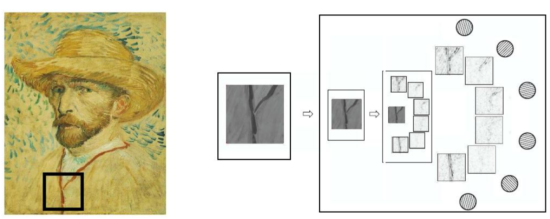

Figure 3: Two-layer wavelet decomposition of one patch from a self portrait by Van Gogh, along the six orientations. Wavelet coefficients with larger magnitudes are depicted in darker gray; smaller magnitude coefficients are lighter

3.4 Hidden Markov Trees

Complex wavelet coefficients capture direction-specific local differences and are appropriate for style analysis. However, for the task of style analysis with high resolution images, the coefficients lie in a very high dimensional space; moreover, the coefficients are still noisy due to the scanning or other processing noise. Consequently, standard, robust dimensionality reduction and feature extraction techniques are required in order to reduce the complexity and noise level of the feature space.

Hidden Markov Trees [7] provides a multiresolution image model that captures the statistical structure of the image. In each subband, wavelet coefficients form a quad-tree with structural local dependencies between wavelet coefficients at different levels. At each scale, hidden variables control the wavelet coefficients. These hidden variables can take two states: ”small”, corresponding to smooth regions, or ”large”, corresponding to edges. The wavelet coefficients are modeled at each scale as samples from a mixture of two Gaussian distributions; one with small variance, controlling the ”smooth” coefficients, and one with large variance, controlling the ”edge” coefficients.

Of course there exist dependencies among the size of the wavelet coefficients at different levels. Hidden Markov Trees capture these dependencies, by two assumptions:

- Hidden: The dependency is between the hidden variables controlling the magnitude of the wavelet coefficients.
- Markov: The dependency is *local: a hidden variable at the finer scale depends only on its parent variable at the coarser scale.*

Hence, the quad-tree structures, in the Hidden Markov Tree model of the subbands, remain independent. The nodes of the tree are the hidden variables controlling the wavelet coefficients, and the observables are the wavelet coefficients themselves.

For dual-tree complex wavelet coefficients, the hidden nodes control the magnitude of the wavelet coefficients. Figure 4 shows the structure of the Hidden Markov Tree for a subband of wavelet coefficients, the hidden variables, and the observable wavelet coefficients. At each level, three parameters control the Hidden Markov Tree :

- α_T : A 2×2 the transition probability matrix $\Pr[\text{child}|\text{parent}]$.
- σ_S: Variance of the narrow Gaussian distribution.
- σ_L: Variance of the wide Gaussian distribution.

Hence, although we are faced with a high number of wavelet coefficients, there are in this model much fewer parameters that control the wavelet coefficients. The HMT parameters can be extracted using the *expectation maximization* algorithm. We then use the set of extracted HMT parameters at all levels as the set of features for our classification task. This is similar to what art historians do, in the sense that each style has some hidden parameters, and art historians try to capture them based on their observations, and on their background knowlege-based model.

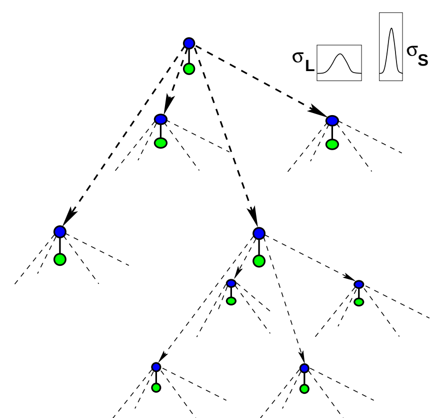

Figure 4: Quad-tree HMT model of one subband of complex wavelet coefficients. Blue nodes represent hidden layer variables (Small or Large), and Black nodes represent the wavelet coefficient. Each layer has is a mixture of two Gaussians with controlling parameters, given by $\alpha_T, \sigma_L, \sigma_S$.

3.5 Style Analysis Algorithm

Since we were provided with very high resolution images, and our approach is based on the local difference analysis, we divided each image into several 256×256 patches. We then treated each patch as an independent training example. The style analysis algorithm is as follows:

Divide images onto 256×256 patches.
for Each patch **do**
Convert the patch to XYZ domain.

Compute complex wavelet coefficients w_x, w_y, w_z.
Compute wavelet norms $w_r = \sqrt{|w_x|^2 + |w_y|^2 + |w_z|^2}$
Extract Hidden Markov Tree Parameters Θ , based on wavelet norms w_r using EM algorithm.

end for

Classify the patches using extracted features, with an appropriate classifier and ten fold cross-validation method.

4. RESULTS

4.1 Dating Challenge

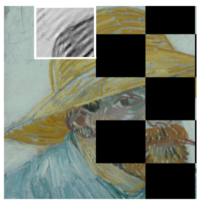

Figure 5: An example of a feature selected as significant for the dating challenge. Usually wavelet coefficients at level 4 and direction 45° are larger, in Paris images. These coefficients are magnified in one patch for one Paris image, to illustrate their meaning.

For the dating challenge, a set of 66 high resolution training paintings were used. The goal of this challenge was the dating of three test images corresponding to paintings that have not been easily or consistently dated in the art historical literature. First the training images were divided into 4727 patches of 256×256 pixels. Using 10-fold cross-validation, several state of the art classifiers, such as Adaboost, SVM, and Random Forest [13] were trained and tested with training examples; as indicated in Table 1, the Random Forest classifier had the lowest generalization error. Figure 6 shows the scatter plot of the prediction made by the above learning methods on the patches of each painting. Black circles indicate the Paris paintings, while green circles indicate the Arles paintings. The horizontal axis corresponds to the number of correctly predicted patches, and the vertical axis corresponds to the number of incorrectly predicted patches. Again the Random Forest classifier performs best.

The final decision was then made using the trained Random Forest classifier on the patches of the test image, as shown in Figure 7. The non-checker-boarded part of each image was divided into as many 256×256 squares as possible, and each square was allocated to the Arles or Paris class (or reduced to a tie) according to the classification results of a Random Forest classifier, trained on the 66 paintings.

As a final note, the obtained results are aligned with the art historians' conclusions. Many art historians agree that the "crab" painting (Figure 7, right) was painted during the Arles period; however, some suggest the Paris period based on comparison with another painting. The "Willows" painting (Figure 7, middle) is suggested to be painted in March 1888, during the Arles period. The "potatoes" painting (Figure 7, left) was traditionally dated in the Paris period, but was recently suggested to be one of the firsts from the Arles Period, i.e. from February 1888.

One of the features that highly separates Paris patches from Arles patches is the variance of the large Gaussian at the fourth level (middle scale) and subband 45°, meaning that wavelet coefficients in Paris are usually smaller, which agrees with the art historian observation that the number of strokes is greater in Paris than in Arles. Figure 5 highlights this type of wavelets.

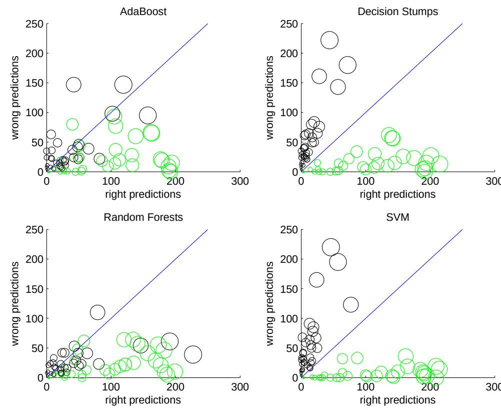

Figure 6: Scatter plot of the prediction made by different learning methods on the total patches of each painting.

Table 1: Comparison of the generalization performance of different learning algorithms using cross-validation in the Dating Challenge. SVM: Support Vector Machine, AB: AdaBoost, DS: Decision Stump, RF: Random Forest)

SVM	AB	DS	RF
66.4%	68.5%	62.12%	73.7%

4.2 Extracting Distinguishing Features

In order to obtain the distinguishing features, several learning classifiers were tried on the HMT parameters, extracted via the EM algorithm. Since the challenge focused only on styles for painting flowers, irrelevant patches were removed from the training examples. In order to make the final decision, several state of the art learning classifiers were examined. The accuracy of the learning algorithm was measured in terms of the error-rate in 10-fold cross-validation as a proxy for the generalization error. Table 1 shows the results.

Table 2: Comparison of Learning Algorithms in terms of their generalization error, (SVM: Support Vector Machine, AB: AdaBoost, BN: Bayes Net, DS: Decision Stump, RF: Random Forest).

SVM	AB	BN	DS	RF
25%	24.9%	29.9%	25.2%	23%

As seen in Table 2, the Random Forest classifier has again superior accuracy compared to the other classifiers. The decision of the Random Forest classifier is based on a weighted average of several features, among which we can identify several superior distinguishing parameters. They are:

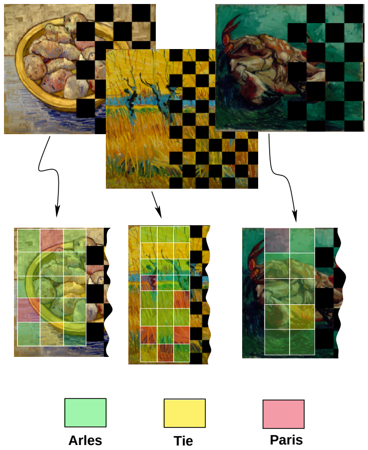

Figure 7: Classification results for the 3 test images in the " dating challenge".

1. $\Pr[large|small]$: probability of transition from small coefficients to large coefficients at scale 6, and subband $-45°$, as illustrated in Figure 8, right. This feature corresponds to very sharp steep brushstrokes in the work of Van Gogh and Monticelli.
2. Variance of "Large Gaussian", at level 4, and subband 15°, as highlighted in Figure 8, left. This feature corresponds to horizontal smoothness in the work of the contemporaries.

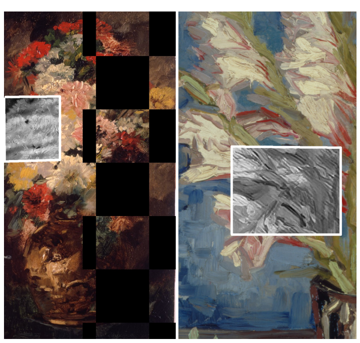

Figure 8: Distinguishing feature for Van Gogh and Monticelli images (right), and the contemporary images (left), In the Van Gogh and Monticelli images, wavelet coefficients at level 6 and direction $-45°$ are magnified to highlight the feature. In contemporary images, wavelet coefficients at level 4 and direction 15° are magnified to highlight the feature.

Note: in this challenge, we did not consider the very finest scales in the images. We found they were more indicative for contrast levels than for difference in style or in brushstroking techniques, so that they could possibly be a confounding factor if they were included. (This would be similar to the confounding blur factor in [10].)

5. CONCLUSION

We used a dual-tree complex wavelet transform, and Hidden Markov Tree modeling for feature selection, and then the Random Forest classifier for classification in "dating" and "extracting distinguishing features" of Vincent van Gogh's paintings to assist art historians in stylistic analysis. In future work, we intend to use additional feature extraction and classification approaches including entropy considerations. We will also devise new methods to reduce the likelihood of overfitting to the training examples, in which will separate the training and test paintings, instead of cross-validating over a mixed bag of patches, provided by all paintings.

REFERENCES

[1] C. R. Johnson, E. Hendriks, I. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang. Image processing for artist identification: Computerized analysis of Vincent van Gogh's brushstrokes. *IEEE Signal Processing Magazine*, 2008.

[2] S. Lyu, D. Rockmore, and H. Farid. A digital technique for art authentication. *Proc. National Academy of Sciences*, 101(49):17006–17010, December 2004.

[3] E. Hendriks and S. M. Hughes. Van Gogh's brushstrokes: Marks of authenticity? In *Proceedings of "Art, Conservation, and Authenticities: Material, Concept, Context," University of Glasgow, Scotland*. Archetype Publications, December 2008.

[4] N. H. Berezhnoy, E. O. Postma, and H. J. van den Herik. Computer analysis of van Gogh's complementary colours. *Pattern Recognition Letters*, 28:703–709, 2007.

[5] M. Barni, A. Pelagotti, and A. Piva. Image processing for the analysis and conservation of paintings: Opportunities and challenges. *IEEE Signal Processing Magazine*, 22:141–144, September 2005.

[6] I. Selesnick, R. Baraniuk, and N. G. Kingsbury. The Dual-Tree Complex Wavelet Transform. *IEEE Signal Processing Magazine*, 2005.

[7] H. Choi, J. Romberg, R. Baraniuk, and N. G. Kingsbury. Hidden Markov tree modeling of complex wavelet transforms. *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2000.

[8] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):138, 1977.

[9] C. Johnson, E. Hendriks, and L. Tilborgh. Digital Painting Analysis. `http://www.digitalpaintinganalysis.org`, 2008.

[10] G. Polatkan, S. Jafarpour, A. Brasoveanu, S. Hughes, and I. Daubechies. Detection of Forgery in Paintings Using Supervised Learning. *Submitted*, 2009.

[11] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. 2nd Edition, Morgan Kaufmann, San Francisco, 2005.

[12] M. Agoston. *Computer Graphics and Geometric Modeling: Implementation and Algorithms*. Springer, 2005.

[13] H. Tim Kam. Random Decision Forest. *Proc of the 3rd Int'l Cof. on Document Analysis and Recognition*, 1995.