A Framework for Computing Artistic Style using Artistically Relevant Features

Catherine A. Buell
Fitchburg State University

William P. Seeley
Bates College

Ricky J. Sethi
Fitchburg State University

***Abstract*—We present two artistically-relevant algorithms to aid in the quantification of artistic style, the Discrete Tonal Measure (DTM) and Discrete Variational Measure (DVM). These quantitative features can provide clues to the artistic elements that enable art scholars to categorize works as belonging to different artistic styles. We also introduce two new datasets for analysis of artistic style: one based on the school of art to which artists belong and one based on the medium used by a specific artist. We show results of initial experiments for classifying paintings on each of these datasets with DTM and DVM using a scientific workflows framework that will allow reuse and extension of many visual stylometry methods, as well as allowing easy reproducibility of analytical results, by publishing datasets and workflows packaged as linked data.**

1. Artistically Relevant Dataset

Visual Stylometry Introduction: The nature of artistic style is the subject of ongoing debate within art history and philosophy of art. Visual Stylometry combines research and methods from art history, computer science, and cognitive science to help quantify the style of an artist [1], [2]. Computational and statistical methods in visual stylometry allow researchers to quantify and compare aspects of artistic style over the course of the career of an individual artist, among artists who share in a common artistic style, and across different schools of art.

Our goal is to develop a quantitative approach that is artistically motivated and has significance for humanities researchers, as well. Since artists and art scholars look at both the distribution of tonal values (the relative lightness of color information within and across different works) as well as the variation of colour and texture in paintings (which is indicative of the style and biomechanics of an artist's brushstrokes) [1], [3], [4], we have developed two measures that can begin to quantify these aspects of artistic style: Discrete Tonal Measure and Discrete Variational Measure.

New Datasets: There are many strategies that machine classifiers might adopt to sort and organize digital images of artworks [5], [6] but these previous approaches and datasets are not targeted to artistically-salient categories or features. It is important that the strategies adopted are consistent with our best understanding of visual processing and we have therefore created two new datasets that are manually curated to concentrate on artistically relevant aspects: one based on the school of art to which artists belong (Impressionism vs Hudson River) and one based on the medium used for a specific artist (medium of tempera vs medium of watercolors). The selection of artworks for each dataset is directed by knowledge of a broad range of normative conventions governing artistic practices. These include conventions governing how a work in a category is normally constructed, what it means for an artist to have chosen to make a work in that category, and what variance in the way these conventions are followed means for experience, understanding, and interpretation of the work. Such considerations guided the careful selection of artworks in each of these datasets. In the end, we included 68 images for the medium of art dataset (32 watercolors/36 egg temperas) and 30 images for the school of art dataset (5 images each from three impressionist painters, Renoir, Sisley, Monet and three Hudson River painters, Bierstadt, Cole, Church), where each image is a high resolution JPG and will be made publicly available. We anticipate continuing to add additional manually curated images in these and other categories as we form new partnerships with museums and digital art repositories.

2. Aristically Relevant Features

Discrete Tonal Measure (DTM): A strong indicator of artistic style is a painter's choice of color palette and how the application of pigment varies across different parts of a painting [7]. One such feature used to gain insight into a painter's artistic style is the use of tone in a painting. **Tone**, in the context of paintings, is how light or dark a color might be. From a computational perspective, we can determine the tone of an image as the distribution of RGB values in various pixel regions. Since artists and art scholars also look at these distributions of tonal values (the relative lightness of color information within and across different works [1], [3], [4]), the **tonal measure** of a painting is a measure of the degree of tonal variance among pixels in a neighborhood. A small area having wide tonal variation will

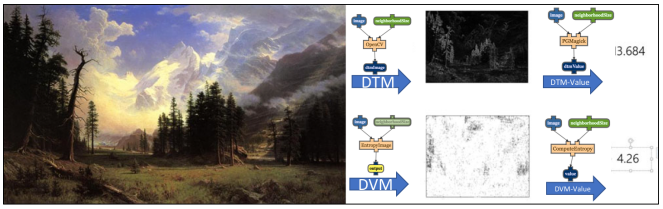

Figure 1. Four workflows in WINGS: 1) DTM producing image; 2) DTM producing value; 3) DVM producing image; 4) DVM producing value.

	c Analysis	b Analysis	MSE
Tempera	88.88%	75%	83.33%
Watercolor	59.375%	59.375%	59.375%

TABLE I. WYETH RESULTS COMPARING TEMPERA VS WATERCOLORS

have a dominant textural appearance, while a neighborhood with little variance will have the appearance of a uniform shade of one tone.

We calculate this variance by first obtaining a grayscale image and scanning it using kernels of varying size. For each kernel, we compute the normalized histogram of occurrences of each pixel value; these normalized frequency values are the probability, P_i, for each pixel, i. Thus, for each pixel in the image, we calculate the P_i, probability of each pixel, and N_g, the number of distinct grey levels in the quantized image. We can then go on to compute the mean for each kernel, k, as $\mu_k = \sum_{i=0}^{N_g-1} iP_i$. This can then be used to calculate the variance, as per, $\sigma_k^2 = \sum_{i=0}^{N_g-1}(i-\mu_k)^2 P_i$, which gives the standard deviation as the square root of the variance and, when averaged over all N_k kernels, yields the *Discrete Tonal Measure (DTM)* as $\frac{\sum_{k=1}^{N_k}\sqrt{\sum_{i=0}^{N_g-1}(i-\mu_k)^2 P_i}}{N_k}$.

We have created workflow fragments for various image processing algorithms and have also implemented DTM (and DVM, as described below) using the WINGS (http://wings-workflows.org/) semantic workflow system which validates semantic constraints of the visual stylometry algorithms. We can see an example of the scientific workflow for the DTM in Figure 1, along with the results of processing. The final results, as shown in the next Section, show that DTM correctly classified images in the test set by genre (Impressionist vs. Hudson River) in over 80% of the trials and that tonal variance in the foreground/background analysis was more indicative of individual artists than was analysis of the whole painting.

Discrete Variational Measure (DVM): Texture, which can be thought of as the roughness or bumpiness from a perceptual perspective, can be characterized by using the entropy of an image [8]. Entropy measures the uncertainty associated with a random variable and we assume that pixel intensity can be modeled as a random variable. We can then use a histogram of intensities to approximate the probability density function and compute the entropy of an image, which measures the variability of pixel intensity information in the image. We can think of the entropy as $E = \sum_{i=1}^{n} p(x_i)\log_2(\frac{1}{p(x_i)})$, where $p(x_i)$ is is the probability of pixel x with intensity k in the image and this can be re-written as $E = -\sum_{i=1}^{n} p(x_i)\log_2 p(x_i)$. We can then use N_k sliding window kernels to traverse the image and compute the average *Discrete Variational Measure (DVM)* as $-\frac{\sum_{k=1}^{N_k}\sum_{i=1}^{n} p(x_i)\log_2 p(x_i)}{N_k}$.

We show a DVM workflow fragment in Figure 1. High entropy values are associated with tight contours and highly textured regions of a painting. Entropy values can thereby be used to determine both the distribution and frequency of color information within a painting and brushstroke style, or the manner in which an individual artist applied paint to his or her canvas. We use the entropy workflow to create trendlines and we can also combine various workflow fragments (e.g., we can combine a DTM and DVM workflow to run both analyses on an artwork, as well as combining workflows we've created for standard image processing like converting to greyscale, background extraction, etc.).

The trendlines are of the form $y = c\ln x + b$. A larger slope, c, is associated with the DVM values in the tempera paintings than the watercolor paintings. Also, the initial value, b, for watercolor paintings is slightly lower. For classification, we put together a set of 32 watercolor paintings and 36 tempera paintings. In order to have more representative testing values, we took five random samples of five paintings from each set. The watercolor paintings contain larger variation in DVM values and the tempera are much more consistent. Results and Mean Square Error (MSE) are shown in Table 1 for the correctly sorted Wyeth paintings. We also used an SVM classifier using the parameters of slope and intercept with 25 images of each set for training and then classifying the remaining paintings (11 tempera and 7 watercolor) resulting in a classification rate of 83% as shown in Figure 2.

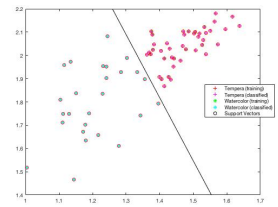

Figure 2. SVM classifier showing classification rate of 83%.

Acknowledgments This research was supported by the National Endowment for the Humanities (NEH) Grant under Award HD-248360-16 and by the Amazon AWS Research Grant program.

References

[1] Daniel J. Graham, James M. Hughes, Helmut Leder, and Daniel N. Rockmore, "Statistics, vision, and the analysis of artistic style," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 115–123, 2012.

[2] Hanchao Qi, Armeen Taeb, and Shannon M. Hughes, "Visual stylometry using background selection and wavelet-HMT-based Fisher information distances for attribution and dating of impressionist paintings," *Signal Processing*, vol. 93, no. 3, pp. 541–553, 2013.

[3] Jana Zujovic, Lisa Gandy, Scott Friedman, Bryan Pardo, and Thrasyvoulos N Pappas, "Classifying paintings by artistic genre: An analysis of features & classifiers," *IEEE International Workshop on Multimedia Signal Processing*, 2009.

[4] Gunnar Goude and Gunilla Derefeldt, "A study of wolfflinś system for characterizing art," *Studies in Art Education*, 1981.

[5] Sergey Karayev, Matthew Trentacoste, Helen Han, Aseem Agarwala, Trevor Darrell, Aaron Hertzmann, and Holger Winnemoeller, "Recognizing Image Style," *BMVC*, pp. 1–20, 2014.

[6] Guilherme Folego, Otavio Gomes, and Anderson Rocha, "From impressionism to expressionism: Automatically identifying van Gogh's paintings," in *ICIP*, 2016.

[7] D.M. Parker and J.B. Deregowski, *Perception and Artistic Style*, Elsevier, 1991.

[8] Andrzej Materka, Michal Strzelecki, et al., "Texture analysis methods–a review," *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, pp. 9–11, 1998.