Discovering Salient Characteristics of Authors of Art Works

Peter Bajcsy[1a] and Maryam Moslemi[a]
[a]National Center for Supercomputing Applications, 1205 W. Clark St., Urbana, IL, USA

ABSTRACT

We addressed the problem of finding salient characteristics of artists from two-dimensional (2D) images of historical artifacts. Given a set of 2D images of historical artifacts by known authors, we discovered what salient characteristics made an artist different from others, and then enabled statistical learning about individual and collective authorship. The objective of this effort was to learn what would be unique about the style of each artist, and to provide the quantitative results about salient characteristic. We accomplished this by exploring a large search space of low level image descriptors. The motivation behind our framework was to assist humanists in discovering salient characteristics by automated exploration of the key image descriptors. By employing our framework we had not only saved time of art historians but also provided quantitative measures for incorporating their personal judgments and bridging the semantic gap in image understanding. We applied the framework implementation to the face illustrations in Froissart's Chronicles drawn by two anonymous authors. We reported the salient characteristics to be (HSV, histogram, k-nearest neighbor) among the 55 triples considered with 5-fold validations. These low level characteristics were confirmed by the experts to correspond semantically to the face skin colors.

Keywords: salient image features, historical illustrations, statistical learning, face illustrations

1. INTRODUCTION

Understanding paintings and historical illustrations in terms of contributing artistic hands has been known as a very challenging problem. There are several formulations of this problem that bring together researchers from the areas of art history, history of book trade, image processing, pattern recognition, computer vision and machine learning. In our work, the problem is formulated as follows: Given a set of 2D images of historical artifacts with known authors, discover what salient characteristics make an artist different from others, and then to enable statistical learning about individual and collective authorship. The objective of this effort is to learn what is unique about the style of each artist, and to provide the results at a much higher level of confidence than has ever before been feasible, by exploring a large search space in the semantic gap of image understanding.

The previous work in this area addressed the problems of grouping and classification of paintings, where classification label corresponds to the artist who drew the painting. Ideally, one would like to find groups of paintings drawn by a single artist. However, this problem has been frequently relaxed to finding groups of paintings drawn by artists with similar style. In our work, the objective is not to find grouping and classification of paintings but rather to identify salient features that discriminate two individual or two groups of artists. We use grouping and classification techniques applied to paintings and historical illustrations as a mechanism for finding the key discriminating features among artists. Our focus is on identifying the search space dimensions that would capture the characteristics discriminating two artists or two groups of artists by analyzing paintings and illustrations in historical manuscripts. The problem is formulated as a search of salient characteristics of artists from two-dimensional (2D) images of historical artifacts where the search space dimensions are not known a priori, and the dimensionality of the search space requires significant computational resources.

While searching for distinct characteristics (features) of artists is time-consuming, computer-assisted techniques can help humanists to discover salient characteristics and increase the reliability of those findings over a large-volume corpus of digitized images. Computer-assisted techniques can provide an initial bridge from the low-level image units, such as color of pixels, to higher-level semantic concepts such as brush strokes, compositions or scene patterns. The technological questions are related to the design of algorithms that can extract evidence at the low-level image units that

[1] pbajcsy@ncsa.uiuc.edu; phone 1 217 265-5387; fax 1 244 1987; isda.ncsa.uiuc.edu

could be aggregated into higher-level semantic concepts, and support humanists in image interpretation, style understanding and authorship assignment. Additional technological questions are about statistical confidence of hypothesis driven authorship assignment which can be answered by processing volumes of images but could not be answered by visually inspecting images with the current human resources within a reasonable time frame. Thus, our key technological questions are about (a) How to extract knowledge about authorship and (b) How to increase our confidence in the characteristics of authorship.

In order to extract knowledge about authorship, one would inspect visually many samples by the same artist and then by multiple artists. We try to mimic the visual inspection by presenting a computer system with a set of samples that could be either labeled (artist is known) or unlabeled (artist is unknown) but grouped together. The challenge is to identify similar (common) features within a group of either labeled or unlabeled samples and dissimilar (discriminating) features across multiple groups of samples. The complexity lies in many representations of images (e.g., color spaces), a very large number of possible features derived from image pixels (e.g., color histograms, edge pixels, texture descriptors, and so on), a variety of similarity/dissimilarity metrics that should correspond to visual perception to serve art historians (e.g., symmetry), and the subjective assignments of the dividing values between similar and dissimilar. Our assumption is that there exists a search space encapsulating the complexity of visual inspection to identify key characteristics of each artist alone and within a group of artists, and we would be able to explore this large search space if sufficient computer resources became available.

The contribution of our work is in exploring a computer assisted methodology to help art historians while analyzing manuscript production, comparing different artist's style, or determining whether one set of hands created the work or some of them were painted by imitators. Currently, humanists must look at images of historical artifacts to determine distinct characteristics of certain individual (or groups of) miniaturists and map engravers, scribes, quilters, and so on. Such visual inspection involves identifying objects in 2D images, recognizing specific types of objects, discriminating differences in realization of those objects, classifying realizations into groups with similarities, building cumulative evidence over multiple groups of objects with similarity in realization, and assigning authorship based on temporally evolving expertise in visual inspection. For example, to assign a label of an artistic hand to an illustration in Froissart's Chronicles, one would first identify objects such as boats, castles, crowns, faces, group of knights, horses, landscapes, skies, spears, tents, towns and water. Next, art historians would look for the discriminating differences in all found instances of these objects and group the objects based on similarities. Finally, using meticulous book-keeping of all groups with similarities, one would build a mapping between the groups of classified objects and the potential assignment of authorship. This manual process is very labor-intensive and cannot be scaled up to large volumes of digital artifacts. In addition, the salient characteristics (a collection of discriminating differences) per artist are described at a high semantic level, which makes it difficult to automate the discovery process. Thus, a computer assisted methodology is needed to bridge the semantic gap in image understanding, and to establish the mappings between pixel level image properties and the higher-level abstract image descriptions.

The rest of this paper is organized as follows. The next section reviews some the related works and compares our approach to the previous work. Section 3 consists of general methodology and a specific framework we have built for addressing the problem of finding salient characteristics of artists. In Section 4, we discuss the experimental results and we conclude the paper in section 6 with remarks about computational costs and the size of a search space.

2. RELATED WORK

The past related work can be divided into categories based on the underlying research objectives and the objects of interests. These objectives would aim at (a) classifying paintings based on image analyses [1], (b) determining authenticity of drawings [2, 3], (c) exploring the tools used for painting (e.g., brushes and brush strokes [4, 5]), or (d) discovering methods of painting (e.g., the use of optics during art creation [6, 7]). As illustrated by the variety of objectives, the objects of interests span historical paintings and drawings, illustrations in historical manuscripts, contemporary photographs or other two-dimensional (2D) objects (e.g., images of quilts [8]).

For example, in [6, 7, 9] the authors addressed the problem of understanding the methods used for creating art. In this series of papers, the authors conducted experiments using the existing computer vision techniques to analyze whether the artists of renaissance used optical projections, how artists created shadows and used lighting to create realistic painting, and finally how comparing shapes painted by different artists helps us to determine the copy from the original work.

While our work would benefit from understanding the differences in methods used by multiple artists, we are taking a data-driven approach in comparison to the model-based approaches described above.

The problem of authentication has been addressed in [2, 3] by using wavelet analysis. The authors divide the image into multiple non-overlapping regions and apply low pass and high pass filters at different scales and orientations in order to extract high dimensional feature vectors. Then, the pair-wise similarity of the images is computed using Hausdorff distance [10], and the dimensionality is reduced by multidimensional scaling for visualization purposes. The method was applied to paintings by Perugino and Bruegel. In the first experiment, the authors manually cropped faces from the images and showed that faces painted by the same artist are closer to each other in low dimension than imitations. In the second experiment, the authors compared seven paintings by Bruegel and his imitators. Their choice of features led to closer distances between paintings by Bruegel than between painting by Bruegel's imitators. The results are obtained over a very limited data set of paintings and it is not clear whether the framework could be generalized to other data sets. The work is related to our problem in the quest for features, similarity metrics and lower dimensional spaces that would separate imitators from the actual painters. Nonetheless, the work in [2, 3] is constrained to only wavelets and Hausdorff distance without a mechanism for expanding and exploring a much larger space of discriminating characteristics.

In the recent work of J. Shen [1], global and local features were used to classify western paintings. Color, texture, shape and color layout were the global features while small image patches filtered using the Gabor filter [11] became the local features. Next, a neural network is employed to learn the labels and by combining the classification score for global and local features. The highest classification score of a pair (image of a painting, label of a painting) serves for establishing the relationship. Given 25 artists and an average of 30 images per artist (1080 paintings altogether), the authors achieved the classification accuracy of 69.6%. The work of Shen [1] could be viewed as similar to our work based on the use of multiple features for painting classification. However, in [1] the search space was constrained to one type of data representation (CIE) and to one type of machine learning method (neural network). In addition, our objective of finding salient characteristics is different from the classification objective. Finally, our approach to achieving scalability (and statistical significance) is different.

Based on our knowledge, there has not been an effort to explore the search space of all possible characteristics that could discriminate artists and understand the computational requirements of such a search.

3. METHODOLOGY

3.1 General Methodology

The problem of discovering salient characteristics that discriminate artists of 2D art is decomposed into several phases. First, we design a search space of salient characteristics and its multi-scale dimensions. The word 'multi-scale' refers to the fact that each dimension of the search space can be sampled at multiple frequencies, and can include or exclude various variables in a given dimension. Second, we form a data workflow for the evaluations of salient characteristics and the overall optimization framework. The data workflow can later be executed using workflow engines such as the NCSA Cyberintegrator [12] on local or remote high performance computing (HPC) resources. Third, we select initial sample points in a high-dimensional search space and prototype their implementations. The sample point is associated with software that explores salient characteristics. Fourth, we design an evaluation and search framework that reports a pair-wise score between two images and a group-wise score over a set of images for a given sample point. Fifth, the search results over a given collection of images are reported for interpretation by art historians with the follow-up annotations. In this phase, the interactions between computer scientists and art historians would narrow the gap between the low-level semantic descriptions of salient characteristics reported by the computer and the high-level semantic descriptions used by humanists. The mapping between low-level and high-level descriptions is captured by annotations enabled in the NCSA Cyberintegrator workflow engine. Sixth, the annotations and interpretations lead to exploring new sample points of the search space and repeating the phases as illustrated in Figure 1 (the arrows from the right most box). The initial implementations of sample points can be substituted by any other implementations and the sampling of each dimension can be changed by adding new sample points using a plug-and-play interface of workflow engines. These computations are indicated in an extra box in Figure 1. In the remainder of this paper we focus primarily on the first three phases.

Design of a search space: In this phase, we analyzed the data-driven processing techniques applied to digital images of 2D artifacts and selected three dimensions used for characterizing artists: (1) types of image representations, (2) categories of feature descriptors, and (3) classes of machine learning methods and similarity metrics for assignments of authorship. Based on our observations, these dimensions are orthogonal (although we have not proved it mathematically) as the choice of a point in each dimension is independent of the other two dimensions and vice versa.

Data workflow: The three main dimensions of the search space also correspond to a sequence of choices in a data-driven processing flow: Load→Transform to image representation→Extract image features→Evaluate pair-wise score using machine learning methods→Compare group-wide scores. A user could view images after loading and after transformation to a selected image representation, as well as the results of feature extraction and then the scores corresponding to low-level semantic descriptors of discriminating salient characteristics at the sample points.

Initialization: We have no a priori knowledge about what could constitute salient characteristics. Our initial selection is based on the past work of researchers in the areas of content-based image retrieval [13, 14].

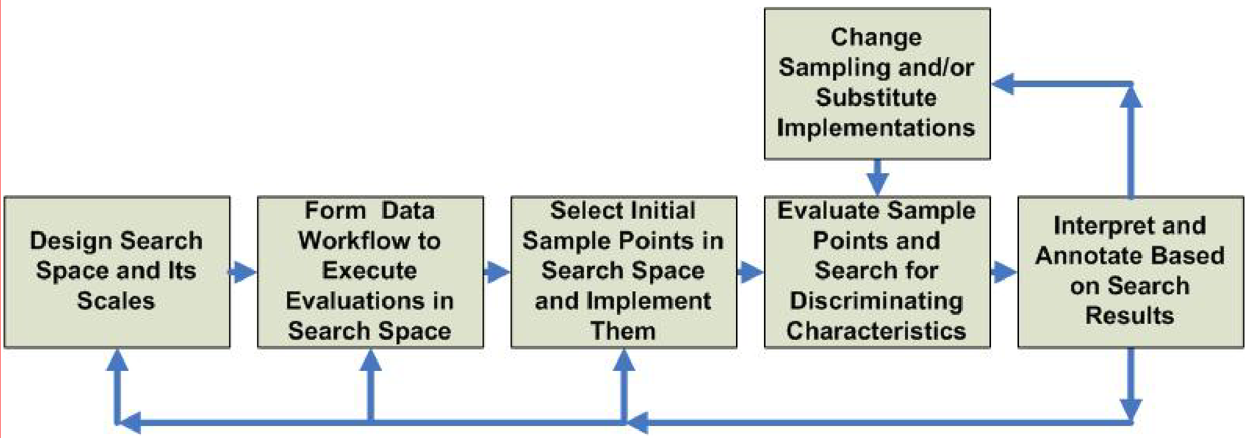

Figure 1. An overall overview of phases to address the problem of discovering salient characteristics.

3.2 Specific Framework

Following the general methodology, we focused on investigating a specific framework in which one could discover discriminating characteristics of artists in a large dimensional space consisting of all possible combinations of image representations, feature descriptors, supervised machine learning methods and their parameters. These characteristics per artist are selected based on the accuracy reported by supervised machine learning methods that compare predicted authorship assignment using the data-driven models with the provided authorship labels. The result of such extensive searches would lead to an n-tuple that provides the highest discrimination with reference to two artists. Our specific framework is illustrated in Figure 2.

For instance, let us assume that the n-tuple found consists of (a) hue color channel in Hue-Saturation-Value (HSV) image representation, (b) frequency of occurrence of each hue value - hue histogram, and (c) similarity of hue histograms measured by chi-squared error and aggregated into groups using the k-nearest neighbors based machine learning method with k=3. Then, a humanist could interpret the discriminating characteristics of two artists to be a hue component of the paint used, a statistical distribution of the hue in the image at the resolution corresponding to the image pixels, and a statistical similarity of at least four (k+1) of these hue distributions across multiple paintings. Thus, visual inspections by a humanist would be assisted by a computer-driven recommendation to focus on a hue component of color images and the similarity of hue distributions in images (or the similarity of hue value frequencies across images). This would reduce the search time of a humanist and could change the role of visual inspection from searching to verification and validation. Furthermore, the images would be delivered for visual inspection in the appropriate representation (e.g., hue channel and its hue histogram) rather than leaving a humanist to recover the hue representation from another color space representation by color transforming images inside of his/her brain (as in the case of visually inspecting hard copies).

Similarly, these pair-wise (artist-to-artist) analyses would lead to a matrix of discriminating characteristics that could be summarized and presented to a humanist. The summaries serve as a decision support into research about what salient characteristics of an artist dominate within a group of artists, a school of artists or a community of artists. Furthermore, they are useful in forensic studies when unseen images are presented to determine authorship.

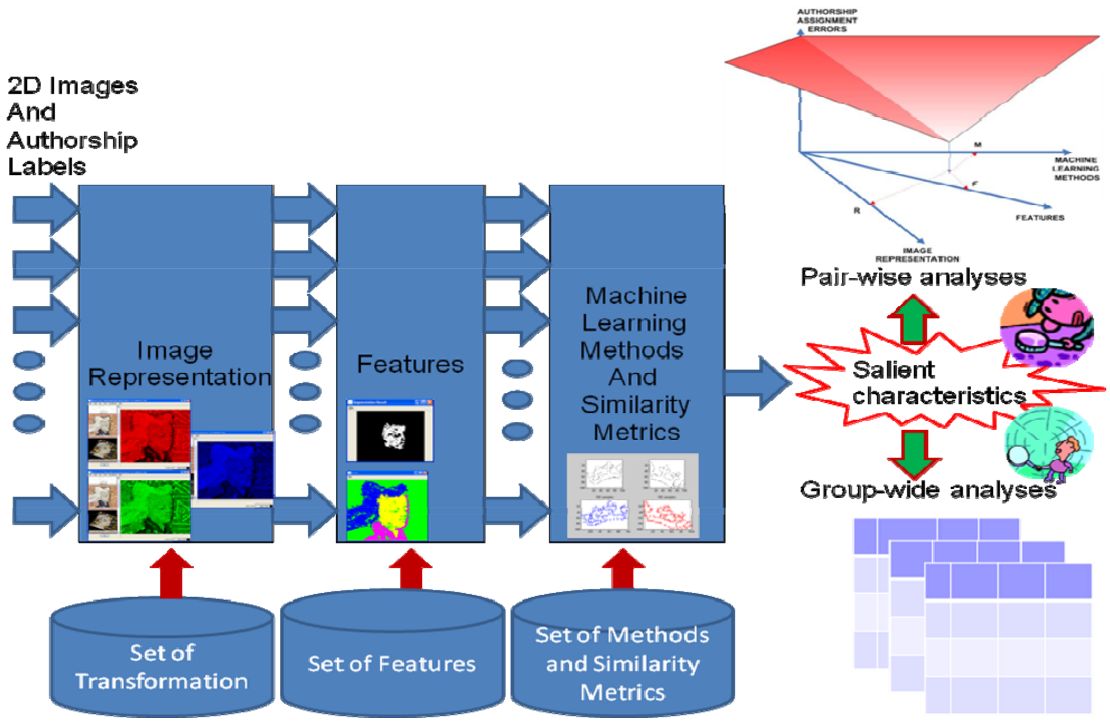

Figure 2. A specific framework for discovering salient characteristics of artists from 2D images of historical artifacts.

<u>*Sampling Image Representation Dimension*</u>
Image Representation: the image representations refer to various ways in which digital images could represent the information about physical objects. The representations include Color spaces (e.g., RGB, HSV, YUV, CIE) [15], Frequency transforms (e.g., Fourier, Hough or Discrete Cosine Transform), Special transforms (e.g., Gabor filters, co-occurrence matrices), Decomposition transforms (principal components, wavelets) and Edge transformations (Sobel, Canny, Robertson, etc. [11]). While there have been studies of what representations are close to human perception following Gestalt psychology principles [16] [17], it has not been established how the image representations map towards discriminating artists and to higher-level semantic descriptions. We explore the search space of the above image representations and choose the one with highest accuracy.

<u>*Sampling Feature Dimension*</u>
Feature descriptors: Once an image representation has been selected, there is a need to describe properties of image pixels (called features) that capture local and global, deterministic and statistical, spatial and spectral image characteristics. The extraction of features can focus on color, shape, or texture properties of 2D images. We explore the search space of the most common features including 1D vector of values, color histogram, shape context [18], or texture descriptors (e.g., extracted from co-occurrence matrices) [19].

<u>*Sampling Method and Similarity Dimension*</u>
Methods and similarity metrics: After extracting the features we need to find the best method which can include both metric and classification scheme to finish the classification task. There are two categories of methods. One class is supervised methods and the other one is unsupervised techniques. In the supervised techniques we need to manually label part of the data and train the classifier using them and test on the rest of the data to measure the accuracy. Support vector machines and nearest neighbor (k nearest neighbor) are examples of supervised techniques. In unsupervised methods we do not need to label any of the data and we can start clustering it based on the similarities of features. An example of unsupervised technique is k-means clustering.

In some of these machine learning techniques, we also need to use a metric for comparing the features in order to group them. Similar to the representation and descriptor selection there are many different metrics that could be used for classification. For example in k-means we can compute the distance between cluster centers and points using L^1 norm, L^2 norm, absolute distance, chi-square distance or matching cost between the features. Although some of these metrics are more meaningful for particular classes of features, in general, any metric can be applied to comparing two sets of identical features. Table 1 shows examples of representations, features (descriptors) and methods.

Table 1. Examples of representations, features (descriptor), and methods.

Representation	Feature	Method
Red-Green-Blue	Vector	K-nearest neighbor
Hue-Saturation-Value	Histogram	Support Vector Machine (SVM)
Principal Component Analysis	Shape context	Neural Networks
Discrete Cosine Transform		K-means
Edge Map		

4. EXPERIMENTAL RESULTS

We have conducted experiments with 48 images of illustration from the 15th century Froissart's Chronicles (Besançon, Bibliothèque d'Etude et de Conservation MS 864 & MS 865). We extracted and labeled manually faces from the illustrations based on the inputs from art historians. There were two artists (two labels) and 48 faces per label. Figure 3 shows two examples of faces drawn by different artists and their representations (left) and histogram features (right).

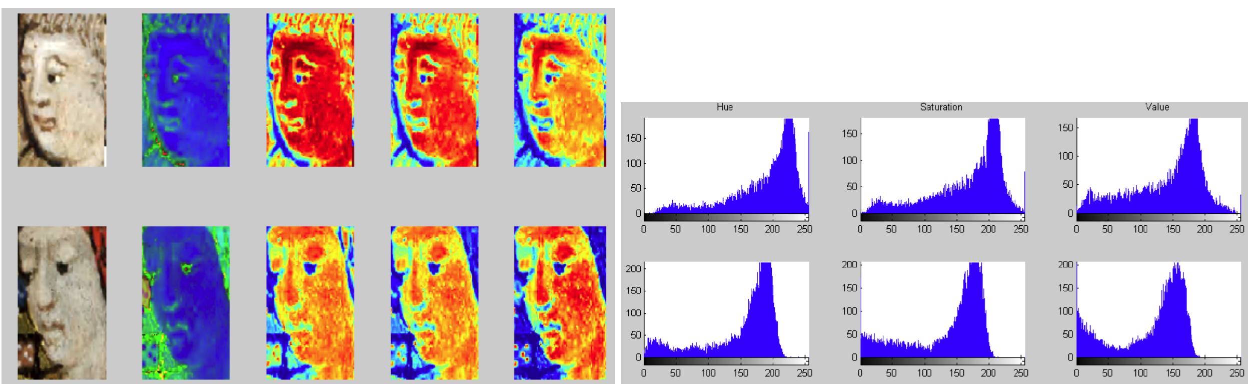

Figure 3. Left – two faces drawn by two different artists. The images to the right of the left most red-green-blue representation correspond to hue-saturation-value representation, pseudo-colored Hue channel, pseudo-colored Saturation channel, and pseudo-colored Value channel. Right – histogram distributions of the hue-saturation-value channels for the two faces.

In our experiments, we used five types of image representations (RGB, HSV, PCA, DCT, and edge map), three types of feature descriptors (1-D vector, shape context, and histogram) and three types of machine learning methods (k-nearest neighbor (KNN), support vector machines (SVM) and K-means clustering). The parameters of various methods were set as follows: the number of histogram bins = 100 (we conducted the sensitivity study to arrive to this number), the number of clusters in K-means clustering = 10 (we measured the purity of clusters to find the accuracy), the k parameter in k-nearest neighbor = 30 (we performed sensitivity analyses to arrive to this number) and 5-fold cross validation for training to report the final score. The number of initial sample points explored is illustrated in Figure 4 as a connected graph.

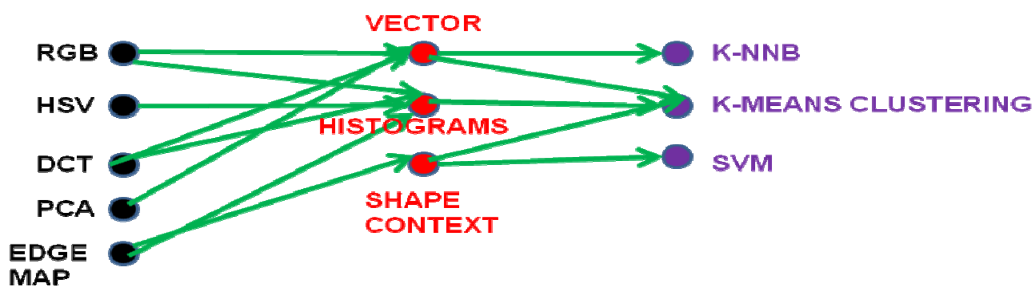

Figure 4. Sample points in a search space represented as a connected graph.

We evaluated all initial sample points to arrive to the triple (representation, feature, method) with the highest score (accuracy of machine learning model after training). The pair-wise (artist-to-artist) analyses have led to a table of triples of representations, descriptors and methods that was summarized and presented to humanists. Based on the experiments conducted with the face illustrations in Froissart's Chronicles drawn by two anonymous authors, we concluded that the triple (HSV, histogram, k-nearest neighbor) led to the highest classification accuracy of 81% among all possible triples. This automated extraction of salient characteristics has been interpreted into higher semantic characteristics (skin color, occurrence of colors used in faces) and visually confirmed by experts by viewing HSV representation of the images.

5. CONCLUSIONS

This paper presented the problem of discovering salient characteristics of artists contributing to 2D art work. We outlined a general methodology and prototyped a specific framework that was applied to a small set of face illustrations from the 15th century Froissart's Chronicles. The computer generated results were consistent with the conclusions of art historians. The key contributions of the work are in the design of a methodology and search space dimensions to discover discriminating characteristics of artists based on data-driven analyses of digital images.

The outlined methodology is computationally intensive as the dimensionality of the search space defined by (image representations, features and machine learning methods) is huge. The computational requirements can be estimated by multiplying the size and number of image pairs to be evaluated times the number and computational complexity of image representations times the number and computational complexity of features times the number and computational complexity of machine learning methods including the number of cross validations. For example, computing the optimal triplet that discriminates between two artists, each represented by 10 images, over 2 color spaces (HSV, H, S, V, HS, HV, SV, RGB, R,G, B, RG, RB, GB), 1 feature type (histogram of each band with the number of bins varying from 100 to 255), one machine learning method (k-nearest neighbors with the k parameter taking values from 1 to 5) and 5-fold cross validation requires (10+1)*5 of image pairs to be evaluated times 14 image representations times 155 features times 5 machine learning variations times 5 cross validation evaluations. This number is equal to 2,983,750 computations with many floating point operations during each accuracy evaluation of machine learning models. In the future, we plan on finishing all components of the general methodology and running the search on high performance computing resources. We anticipate the number of computations to be close to 300 million.

6. ACKNOWLEDGEMENT

The authors would like to thank Dr. Anne D. Hedeman and her doctoral student Ms. Natalie A. Henson at UIUC for providing valuable insights on the process of artistic hands to illustrations, and for selecting illustrations from the 15th century Froissart's Chronicles. The digital copies of illustrations from the 15th century Froissart's Chronicles were kindly made available to our research by Dr. Peter Ainsworth and Dr. Mike Meredith from the University of Sheffield, UK, using Virtual Vellum.

REFERENCES

1 Shen, J.: 'Stochastic modeling western paintings for effective classification.', Pattern Recognition, 2009, 42, (2), pp. 293–301

2 Rockmore, D., Lyu, S., and Farid, H.: 'Wavelet analysis for authentication.', in Editor (Ed.)^(Eds.): 'Book Wavelet analysis for authentication.' (2005, edn.), pp.

3 Rockmore, D., Lyu, S., and Farid, H.: 'A digital technique for authentication in the arts', International Foundation for Art Research (IFAR) Journal, 2006, 8, (2), pp. 21–29

4 Melzer, T., Kammerer, P., and Zolda, E.: 'Stroke detection of brush strokes in portrait miniatures using a semiparametric and a model based approach.', in Editor (Ed.)^(Eds.): 'Book Stroke detection of brush strokes in portrait miniatures using a semiparametric and a model based approach.' (IEEE Computer Society., 1998, edn.), pp. 474

5 Sablatnig, R., Kammerer, P., and Zolda, E.: 'Hierarchical classification of paintings using face- and brush stroke models.', in Editor (Ed.)^(Eds.): 'Book Hierarchical classification of paintings using face- and brush stroke models.' (1998, edn.), pp. 172–174

6 Stork, D.G.: 'Computer vision, image analysis, and master art: Part I', IEEE Multimedia, 2006, 13, (3), pp. 16–20

7 Stork, D.G., and Johnson, K.: 'Computer vision, image analysis and master art, Part II: Finding the illuminant in realist paintings', IEEE Multimedia, 2006, 13, (4), pp. 12-17

8 Richardson, J., Fegan, M., Kornbluh, M., Rehberger, D., and MacDowell, M.: 'Bits & Bolts to Bits & Bytes: The Quilt Index On-line Repository and Distributed Archival Management System', in Editor (Ed.)^(Eds.): 'Book Bits & Bolts to Bits & Bytes: The Quilt Index On-line Repository and Distributed Archival Management System' (Archives & Museum Informatics, 2004, edn.), pp.

9 Stork, D.G., and Duarte, M.: 'Computer vision, image analysis and master art, Part III: Quantifying shape in realist art', IEEE Multimedia, 2007, 14, (1), pp. 14–18

10 Huttenlocher, D.P., Klanderman, G.A., and Rucklidge, W.J.: 'Comparing images using the hausdorff distance', IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15, (9), pp. 850–863

11 Russ, J.C.: 'Image Processing Handbook' (CRC Press, Inc.,, 2002, 4th edn. 2002)

12 Kooper, R., Marini, L., Minsker, B., Myers, J., and Bajcsy, P.: 'CyberIntegrator: A Highly Interactive Scientific Process Management Environment to Support Earth Observatories'. Proc. Geoinformatics, San Diego, CA, May 16-18, 2007 2007 pp. Pages

13 Veltkamp, R.C., and Latecki, L.J.: 'Properties and Performances of Shape Similarity Measures.', in Editor (Ed.)^(Eds.): 'Book Properties and Performances of Shape Similarity Measures.' (Springer, 2006, edn.), pp. 47-56

14 Veltkamp, R.C., and Tanase, M.: 'Content-Based Image Retrieval Systems: A Survey', in Editor (Ed.)^(Eds.): 'Book Content-Based Image Retrieval Systems: A Survey' (Department of Computing Science, Utrecht University, 2002, revised version from 2000 edn.), pp.

15 Schwarz, M.W., Cowan, W.B., and Beatty, J.C.: 'An experimental comparison of rgb, yiq, lab, hsv, and opponent color models.', ACM Trans. Graph, 1987, 6, (2), pp. 123–158

16 Jadacki, J.: 'On Forms of Objects', in Albertazzi, L. (Ed.): 'Shapes of Forms: From Gestalt Psychology and Phenomenology to Ontology and Mathematics' (Kluwer Academic Publisher, 1999), pp. 341-361

17 Albertazzi, L.: 'Shapes of Forms: From Gestalt psychology and phenomenology to ontology and mathematics.' (Dordrecht: Kluwer Academic., 1999. 1999)

18 Belongie, S., Malik, J., and Puzicha, J.: 'Shape matching and object recognition using shape contexts.', IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24, (24), pp. 509–522

19 Duda, R.O., Hart, P.E., and Stork, D.G.: 'Pattern Classification ' (Wiley-Interscience, 2000, 2nd edn. 2000)