https://doi.org/10.2352/issn.2168-3204.2021.1.0.16

Surface Metrology and Data Science/Analytics Applied to Modern Asian Lacquer Surfaces

H. David Sheets, Physics Department, Canisius College, Buffalo N.Y. USA
Patrick Ravines, Patricia H. and Richard E. Garman Art Conservation Department, Buffalo State College, Buffalo , N.Y. USA
Marianne Webb, Webb Conservation Services, British Columbia, Canada

Abstract

This paper presents a new quantitative approach to the study of Asian lacquers using surface metrology, and two data science approaches: feature engineering and convolutional deep neural networks, as used in machine vision or image recognition applications. The types of Asian lacquers and additives have a quantifiable impact on the topography of the resulting surface. To understand the unaged and aged characteristics, 15 different formulas of Asian lacquer were prepared using laccol, thitsiol and urushiol with the most common additives: oils, pigments and resins. These were studied with the surface metrology instrumental technique of confocal microscopy.

Introduction

This paper presents our foray into both surface metrology and data science/analytics to study and better understand Asian lacquer surfaces.

Research at the Getty Conservation Institute on the chemical characterization of Asian lacquers has demonstrated that the type of lacquer mixtures and the additives have an impact on the resulting surface. As part of the GCI's ongoing project to understand the aging characteristics and develop cleaning methods for Asian lacquers, in 2017, 15 different formulas of Asian lacquers for surface texture analyses were prepared. Three types of Asian lacquer, laccol, thitsiol and urushiol, were obtained or purchased from commercially reliable sources. Traditional recipes and preparation protocols were followed to minimize differences and ensure standardization of the final products. The five laccol, four thitsiol and five urushiol panel formulas differ from the next in the series by having had one additional common ingredient added such as oils, pigments and resins. Having different individual panels representing each lacquer formulation allows for the investigation of how each additional ingredient affects the surface texture [1].

The surface metrology instrumental technique of confocal microscopy was used to study 14 Asian lacquer panels before (unaged) and after accelerated aging. For the 14 unaged panels, 12 distinct areas of interest were examined using a 10x (area 1,600 x1,600 μm) objective [1]. For aging, each panel was divided into four sections and each section was artificially aged first with light at 100, 200, 300 and 400 hours (Xenon arc lamp; 0.5W/m2 at 340nm with sodium borosilicate inner filter and soda lime outer filter), and then submitted to relative humidity cycling of 80% per week followed by 20% per week for eight weeks [2].

Three regions of interest have been measured in each of the four aged sections of each of the 14 aged panels. This paper presents data on the unaged and aged laccol, thitsiol and urushiol specimens.

Processing a Surface Scan for Analysis

The measured data is a map of a surface topography from a confocal microscope scan that is 984 pixels by 984 pixels in size with each pixel representing 1.62 microns, with a vertical height measurement in each pixel. Each surface is then represented by 968,256 (984 squared) floating point values, which is similar to a black and white digital image, except that the values are heights rather than greyscale values. To use these images in subsequent statistical analyses, there are two general classes of approaches available, feature engineering, and the use of convolutional deep neural networks which work with raw images.

Feature engineering is the process of extracting useful derived variables from a data set that act as summary variables (features) for further analysis. In ordinary life, people use a range of features to describe objects in the world around them, the process of feature engineering is mathematical in nature but akin to this ordinary human process. The feature engineering approach employed here is an extension of the common division of the surface topography into two components, waviness or contour and roughness.

In the typical approach to estimating waviness and roughness, a spatial filter (often a robust Gaussian filter) is applied to the data to smooth it and yield an estimated waviness (the smoothed surface) [3]. The difference between the original surface and the smoothed surface is termed the roughness component, or the residuals of the filter, thus splitting the topography of the surface into the two distinct components. The filter size used in this process is typically taken as ten percent of the size of the surface. The root- mean-square (RMS) of the height value is taken as an overall measure of roughness of the surface. A number of more complex approaches to surface characterizations have been developed, including those based on discrete Fourier transforms [4] or other advanced feature engineering methods [5]. Within this study, the gaussian filter process has simply been extended to more than two components. Rather than apply a single filter of fixed size to produce a single roughness and smoothed surface, a series of many filters of increasing size are used to successful compute a roughness and a contour for each size of the filter. The smallest filter size is taken as the resolution limit of the measurement, meaning a single pixel. The filter size is then steadily increased at some desired rate (often 1 pixel at time), so that the residuals of the filter produce a roughness estimate at the scale of each filter used. The RMS value of the residuals at each filter size thus form a roughness spectrum which describes the magnitude of the roughness at a variety of scales. The entire roughness spectrum of a sample thus forms a set of engineered features called a feature vector [5] that describes the texture of the surface as a series of scale-dependent measurements rather than a single RMS roughness. This iterative filtering process was carried out in R [6] using a robust gaussian filter from the spatstat package. The resulting roughness spectra may then be analyzed with a wide range of multivariate statistical methods.

This digital approach to a roughness spectra approach has several physical analogs. Geologist or civil engineers will pass a collection of sediment through a series of sieves of decreasing size, so that the proportion by weight of each size of sedimentary component may be determined [7]. Different environments will yield different patterns (spectra) of grain sizes. Woodworkers purchase sandpaper by grit size, with low grit sizes indicating coarse or large grained sandpaper and higher grit sizes indicating progressively finer grit sizes. In a roughness spectrum, small filter sizes correspond to fine “grits” and larger filter sizes correspond to coarser “grits”.

The second approach presented here makes use of a deep learning method called a convolutional neural network (CNN) [8], which has proven effective in image analysis in detecting and identifying objects in an image. The surface scans may be thought of as monochrome images, in which the height at each location is false color coded in shades of grey. The CNN used in the current study utilized N x N pixel patches of the entire image, so that than many different training examples were available from each of the 984 x 984 data scans available. Three different patch sizes, 50 x 50, 75 x 75 and 100 x 100 pixels were examined. The CNN was implemented in Python using the Keras API to the TensorFlow library [9]. A simple CNN structure of 3 paired convolution and pooling layers, followed by flattening layer, one hidden layer for the classification and an output layer was used, based on examples from Chollet [8]. This system was then trained to classify input images as coming from one of three different types of lacquer.

Results

The roughness spectra were calculated for all surface scans from all samples, both the aged and non-aged of laccol, thitsiol and urushiol lacquers with no additives, using a total of 19 different filter sizes ranging from 1.62 microns up to 89.1 microns in size. The mean and standard deviation were then calculated for each lacquer type, based on the non-aged surfaces, as shown in Figure 1. There are clear differences in the mean values of the roughness spectra for the three lacquers.

As each individual specimen is represented by the 19 values in the roughness spectra, an ordination method (an unsupervised learning method) called Principal Components Analysis (PCA) [5,10] was used to produce a reduced dimensionality plot. PCA produces a depiction of the largest patterns of variation in the data (the roughness spectra of each sample), by producing a set of engineered features called the Principal Axes. Specimens can be located or placed along these axes based on the roughness spectrum scores of the given specimens. Each PCA axes explains some proportion of the total variance in the data, much as a regression model explains variance. Unlike a regression model, the PCA has no independent predictor variables, it operates by simply summarizing the variance in the data. PCA plots require practice to interpret but provide a simplified image of the large patterns of variation in the data. If different groups of specimens (such as lacquer types) group or cluster on the diagram, it indicates that specimens with similar group variables (lacquer type or age) also have similar patterns in the roughness spectra. Note that the PCA does not use information about the lacquer type in estimating the axes.

The first two PCA axes based on the analysis of the roughness spectra of 111 specimens in our study show a clear cluster or segregation of specimens by lacquer type. The aged (age categories 1-4) urushiol specimens do separate substantially from the unaged specimens (category 0). Aged specimens of the other two lacquers remain close to the unaged specimens. The first PC axis (PC1) explains 75.8% of the variation in the roughness spectra, while the second (PC2) explains 22.4% of the variation, the two combined represent 98.2% of the total variation. The PC axes are linear combinations of the roughness spectra, and by examining a PCA “axis loading”, it is possible to determine which elements of the roughness spectra have high loadings and thus contribute strongly to the PC axis. Figure 3 shows that the 16.2 micron roughness is the major influence on PC1, while PC2 is strongly influenced by roughness at 1.62 microns, and also at 3.2 microns and 16.2 microns again. Figure 1 does indicate that the roughness values at 16.2 microns have high variance within each lacquer type, consistent with the broad bands of each lacquer time visible along the PC1 axis in

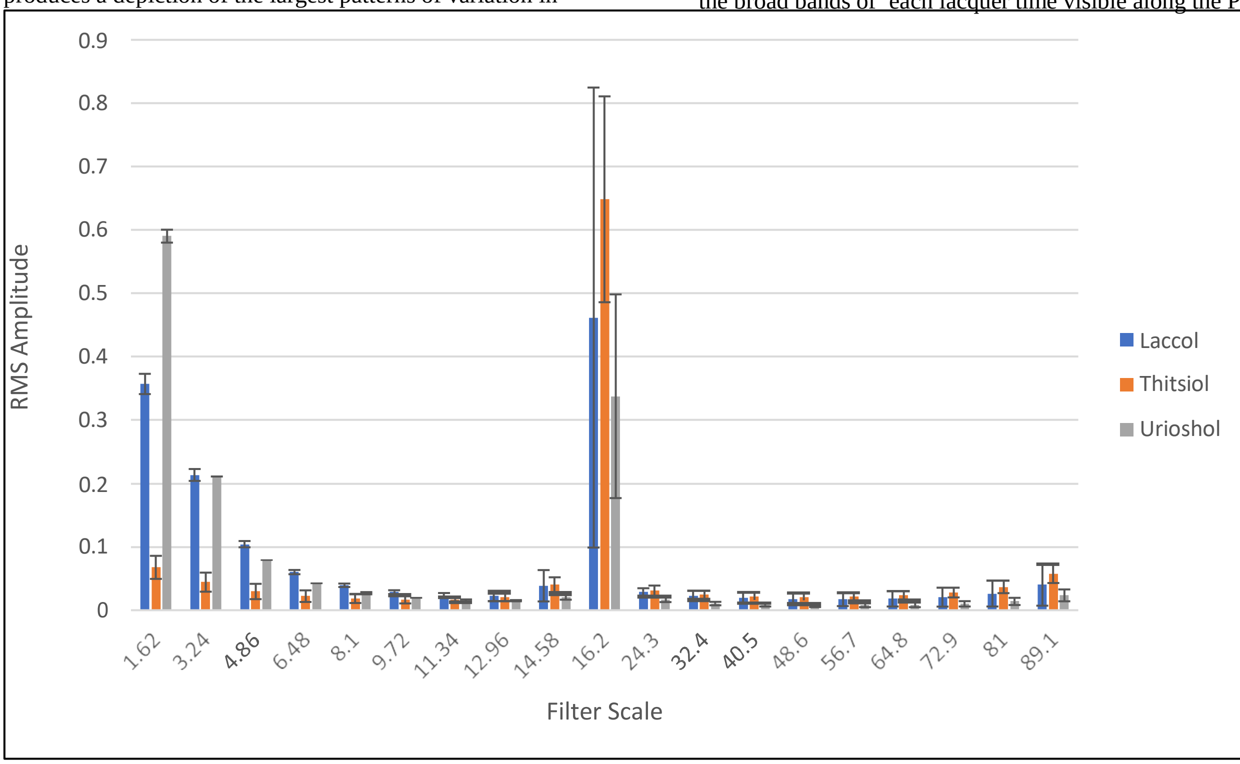

Figure 1: The mean and standard deviation of each roughness spectra component for the three unaged lacquer specimens.

Figure 2. PC2 is more effective in separating the 3 lacquer types in Figure 2, consistent with the difference in the mean values of the roughness spectra at the smallest scales seen in Figure 1.

The roughness spectrum can also be used in a supervised learning method to classify specimens to a lacquer type. In this case, a random forest [11] method using 500 decision trees was trained to classify specimens into lacquer types. The forest was trained on a subset of the data and then evaluated using specimens not used in the training process. In this case, a method called cross validation was used in which the random forest was trained using 90% of the unaged data and then tested using the remaining 10% of the unaged data, this was repeated 100 times to obtain an estimate of the expected classification performance, which was 99% correct for the unaged data. Aged data was not used in the training. The roughness spectra from the aged specimens was then used as the input to the random forest classifier.

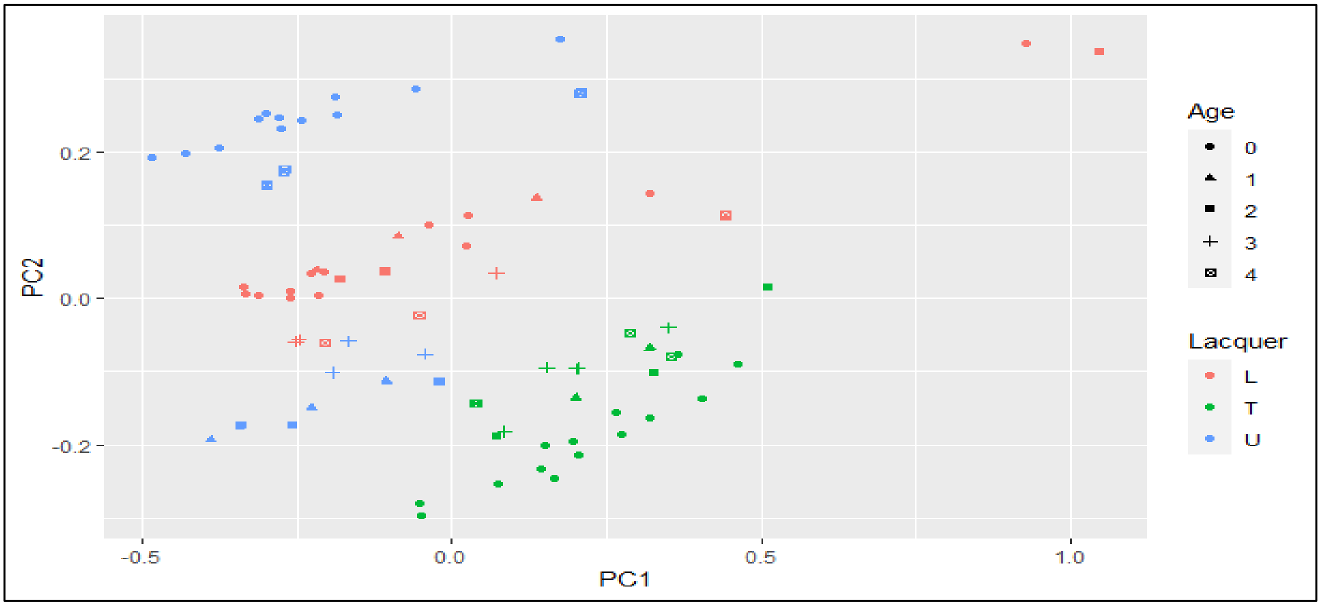

Figure 2: PCA Scores of all unaged (Age=0) and aged specimens. The age and lacquer type are not inputs to this unsupervised learning method, these categories are only used in plotting. The three unaged lacquers form three vertically separated "bands" while many of the aged urushiol specimens lie between the band of the laccol and thitsiol specimens.

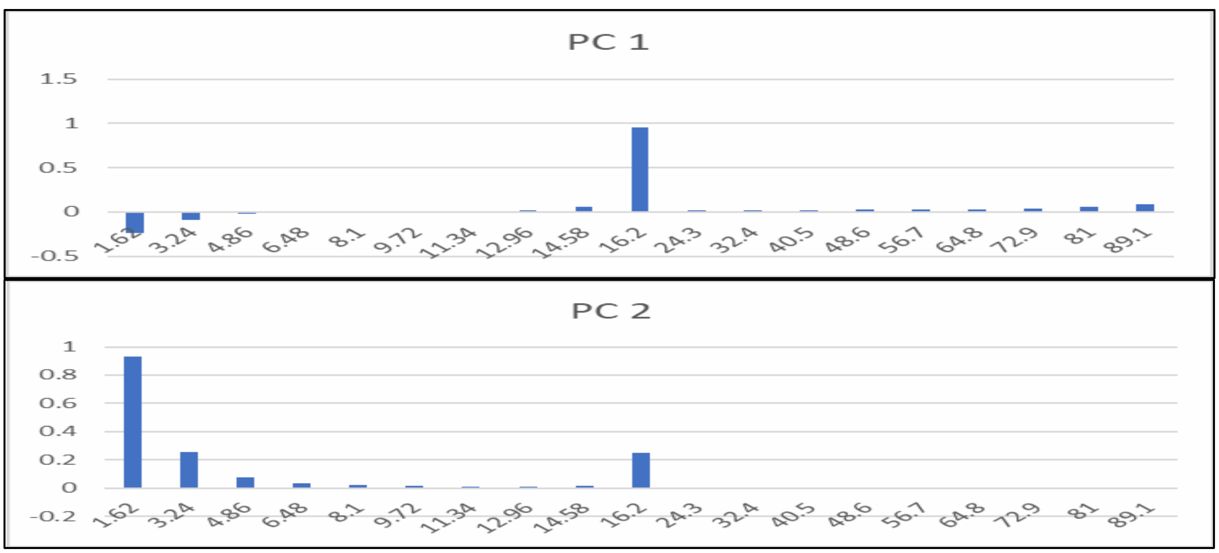

Figure 3: Plots of the axis loadings for the two PCA axes, used to determine which roughness features are represented along each axis.

The overall rate of correct assignments of the aged specimens was 62.7%, with the classification rates for each lacquer shown in Table 1 below. Not surprisingly, the aged urushiol specimens were typically misclassified, as might be expected by the group of urushiol examples close to the laccol and thitsiol specimens in the PCA plot (Figure 2). The specimens of aged laccol and thitsiol had correct assignment rates of 75% and 84.2% respectively, while aged urushiol had a 33.3% correct rate, which is essentially random. However, the aged urushiol was always misclassified as laccol, never as thitsiol.

	Lacquer Type		
Predicted Type	Laccol	Thitsiol	Urushiol
Laccol	18	3	16
Thitsiol	2	16	0
Urushiol	4	0	8
Percent Correct	75	84.2	33.3

Table 1: *Assignments of aged lacquers using a Random Forest based on roughness spectra. Correct assignments are along the diagonal, off-diagonal values are errors.*

The CNN based classifier rapidly reached over 99% accuracy in cross-validation testing using randomly selected 100 by 100 pixels (162 micron x 162 micron) slices of the surface scan. The overall rate of correct assignments of the aged specimens was 67.1%. Again, it proved difficult to correctly assign aged urushiol specimens, as seen in Table 2. Slice sizes of 50 x 50 pixels and 75 x 75 pixels were also tested (results not shown), the correct assignment rate increased with the size of the slice.

	Lacquer Type		
Predicted Type	Laccol	Thitsiol	Urushiol
Laccol	21	2	10
Thitsiol	0	20	8
Urushiol	3	0	6
Percent Correct	87.5	90.9	25.0

Table 2: *Assignments of aged lacquers using a Convolution Neural Network based on 100 x 100 pixel slices of the data.*

Conclusions

The goal of this study was to determine if confocal microscope scans of the surface topography of three different lacquer coatings could be used to determine the type of lacquer used and to detect indications of aging in the surface. Our sample size consisted of a total of 72 scanned regions; 36 scans from three unaged lacquer panels (12 per panel), and 36 scans of the three aged lacquer panels with four aged regions. While this represents a substantial data collection effort, the sample size is limited relative to the typical data sets available in other applications of machine learning.

The feature engineering-based approach led to the development of a roughness spectra that produces scale dependent characterizations of the roughness that can be used in subsequent analyses. A simple plot of the mean and standard deviations of the roughness spectra for the three lacquers in unaged form show clear differences in the roughness spectra (Figure 1), which are most noticeable at the 1.62 and 3.24 micron scales. A PCA analysis of the data shows clear segregation of the three unaged lacquers, with aged laccol and thitsiol specimens remaining close to the unaged specimens, while urushiol specimens changed in more complex ways. It should be noted that one person prepared the thitsiol and laccol panels, while another person prepared the urushiol, so the source of the variation cannot entirely be determined.

Clearly, though, there were detectable changes in the roughness spectra associated with specimen aging. Effective classification of aged specimens based on the roughness spectra will require a detailed understanding of the specimen changes associated with aging, and larger sample sizes.

A random forest classifier trained on unaged specimens performed well in classifying unaged specimens, but the performance deteriorated when aged specimens were classified with the same random forest classifier. The CNN classifier which made use of the raw data matrix without any feature engineering and showed marginally better performance that the random forest classifier operating on the roughness spectra. Given more data, it will be possible to train both classifiers on both aged and unaged data, thus incorporating information about the aging process into the classifier to improve performance.

This pilot study shows that advanced analytic approaches can be used effectively to extract information from surface texture scans. Both feature engineering and convolutional network based approaches can successfully classify surfaces to the correct lacquer types, as well as identifying and quantifying the impact of surface aging.

References:

1. Ravines, P, D. Sheets and M. Webb. Surface metrology concepts and practice applied to modern Asian lacquer surfaces. Future Talks 019 Surfaces proceedings, papers from conference held at the Die Neue Sammlung, Munchen, Germany, November 2019. In press, expected November 2021, (2021).
2. Han J. et al. in preparation and personal communication, (2021).
3. Raja, J., Muralkishnan B. and Fu, S. "Recent advances in separation of roughness, waviness and form". Prec. Eng. 26 (2002), pg. 222-235
4. Massarao, A., Dipierro, G. Cannella, E. and Galiano, A. M. "Comparative analysis among discrete Fourier transform, K-means and artificial neural networks image processing techniques oriented on quality control of assembled tires". Information 2020 (2020), 11(5) p 257.
5. Vorburger, T.V. Song, J. and Petraco, N. "Topographic measurements and applications in ballistics and tool mark identifications". Surf. Topog Metrol Prop. 4 013002 (2016)
6. R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org
7. Poppe, L.J., Eliason, A.H., Fredericks, J.J., Rendigs, R.R., Blackwood, D., Polloni, C.F., 2000. Chapter 1. Grain-size analysis of marine sediments e methodology and data processing. In: Poppe, L.J., Hastings, M.E., Eliason, A.H., Fredericks, J.J., Rendigs, R.R., Blackwood, D.S. (Eds.), U.S.G.S. East-coast Sediment Analysis: Procedures, Database, and Georeferenced Displays. U.S. Geological Survey, Woods Hole, MA. U.S. Geological Survey Open File Report 00e358.

8. Chollet, F. Deep Learning in Python. (Manning Publications, Shelter Island, NY. 2018)

9. Chollet, F. & others, Keras. Available at: https://github.com/fchollet/keras. (2015)

10. Pearson, K. "On lines and planes of closest fit to systems of points in space", Phil. Mag. 2(11) (1901).

11. Ho, T.K. "The random subspace method for constructing decision forests". IEEE Trans. On Patt.Analy. and Mach. Int. 20(8) 832-844. (1998)

Author Biography

Patrick Ravines has been director and professor in the Patricia H. and Richard E. Garman Art Conservation Department, Buffalo State College, since 2010. From 2007 to 2010 he was senior project conservator and research fellow and from 2005-2007 he was an Andrew W. Mellon Fellow in the Advanced Residency Program in Photograph Conservation at George Eastman Museum, Rochester, NY. He was chief of the Conservation Office, Baha'i World Centre, Haifa, Israel, from 1986 to 2005.

H. David Sheets has a wide range of research interests, all based on the study of dynamical processes, from a mathematical and statistic perspective. This has included the study of the growth and diversification of biological organisms, such as long-term change within lineages, as well as large scale patterns of biodiversity changes, particularly mass extinctions. Some of this work has also led to an involvement with forensics, where Dr. Sheets has been a strong critic of forensic bitemark analysis, and has served on a national level scientific panel charged with establishing best practices in forensic handwriting analysis.

Marianne Webb is an independent conservator and researcher in Vancouver, Canada. Previously she was the Decorative Arts Conservator at the Royal Ontario Museum, Toronto, where she developed her interest in Asian and western lacquers. Currently she is collaborating with the Getty Conservation Institute on their research into the characterisation of Asian lacquer and developing cleaning techniques for these complicated surfaces. Marianne received an honour's degree in Fine Art from the University of Toronto and a diploma in Art Conservation Techniques from Sir Sanford Fleming College.