PURSUING AUTOMATED CLASSIFICATION OF HISTORIC PHOTOGRAPHIC PAPERS FROM RAKING LIGHT IMAGES

C. RICHARD JOHNSON, JR.[1], PAUL MESSIER[2], WILLIAM A. SETHARES[3], ANDREW G. KLEIN[4], CHRISTOPHER BROWN[4], ANH HOANG DO[4], PHILIP KLAUSMEYER[5], PATRICE ABRY[6], STEPHANE JAFFARD[7], HERWIG WENDT[8], STEPHANE ROUX[6], NELLY PUSTELNIK[6], NANNE VAN NOORD[9], LAURENS VAN DER MAATEN[9,10], ERIC POSTMA[9], JAMES CODDINGTON[11], LEE ANN DAFFNER[11], HANAKO MURATA[11], HENRY WILHELM[12], SALLY WOOD[13], AND MARK MESSIER[14]

[1] *Cornell University/Rijksmuseum*
[2] *Paul Messier LLC*
[3] *University of Wisconsin*
[4] *Worcester Polytechnic Institute*
[5] *Worcester Art Museum*
[6] *Ecole Normale Superieure de Lyon, Centre National de la Recherche Scientifique*
[7] *University of Paris*
[8] *Institute de Recherche en Informatique de Toulouse, Centre National de la Recherche Scientifique*
[9] *Tilburg University*
[10] *Delft University of Technology*
[11] *Museum of Modern Art*
[12] *Wilhelm Imaging Research*
[13] *University of Santa Clara*
[14] *Indiana University*

Surface texture is a critical feature in the manufacture, marketing, and use of photographic paper. Raking light reveals texture through a stark rendering of highlights and shadows. Though close-up raking light images effectively document surface features of photographic paper, the sheer number and diversity of textures used for historic papers prohibits efficient visual classification. This work provides evidence that automatic, computer-based classification of texture documented with raking light is feasible by demonstrating an encouraging degree of success sorting a set of 120 images made from samples of historic silver gelatin paper. Using this dataset, four university teams applied different image-processing strategies for automatic feature extraction and degree of similarity quantification. All four approaches successfully detected strong affinities and outliers built into the dataset. The creation and deployment of the algorithms was carried out by the teams without prior knowledge of the distributions of similarities and outliers. These results indicate that automatic classification of silver gelatin photographic paper based on close-up texture images is feasible and should be pursued. To encourage the development of other classification schemes, the 120-sample "training" dataset used in this work is available to other academic researchers at http://www.PaperTextureID.org.

KEYWORDS: *Photographic paper, Texture, Eigentexture, Random-feature texton, Area-scale analysis, Anisotropic wavelet multiscale analysis, Automatic classification, Texture dataset, Computational art history, Digital humanities, Art authentication, Image processing for art investigation*

1. TEXTURE IN PHOTOGRAPHIC PAPER

Texture is a defining attribute of photographic paper. Starting in the early 20th century, manufacturers manipulated texture to differentiate their products and to satisfy the esthetic and functional requirements of photographers. Especially prior to WWII, when

DOI: 10.1179/1945233014Y.0000000024

Table 1. Teams and computational methods

Lead institution	Team members	Computational method	Method classification
University of Wisconsin, Madison (Wisconsin)	William Sethares	Eigentextures	Non-semantic
Tilburg University (Tilburg)	Nanne van Noord, Laurens van der Maaten, Eric Postma	Random-feature texton method	Non-semantic
Ecole Normale Superieure de Lyon (Lyon)	Patrice Abry, Stéphane Jaffard, Herwig Wendt, Stéphane Roux, Nelly Pustelnik	Anisotropic wavelet multiscale analysis	Multiscale
Worcester Polytechnic Institute (WPI)	Andrew Klein, Christopher Brown, Anh Hoang Do, Philip Klausmeyer	Pseudo-area-scale analysis	Multiscale

black-and-white silver gelatin paper was the dominant photographic medium (Messier 2008), dozens of manufacturers worldwide produced a wide array of surfaces. From this period a book of specimen prints by the Belgian company Gevaert lists 25 different surfaces made up of combinations of texture, reflectance, color, and paper thickness (The Gevaert Company of America c.1935). Around the same time, a sample book from the Defender Company of Rochester, New York lists 27 surfaces (Defender Photo Supply Company c.1935), Mimosa 26 (Mimosa c.1930), and Kodak 22 (Eastman Kodak Company c.1935a, c.1935b). Each listed surface was proprietary to the different manufacturers and each was used across their multiple brands of paper with changes, additions, and deletions occurring over a span of many years.

A vital factor in the evaluation of paper surface, texture impacts the visibility of fine detail and other qualitative features, thus providing insight into the intent of the photographer and the envisioned purpose of a particular print. For example, prints made for reproduction or documentary functions tend to be better suited to smooth-surface papers that render detail with sharpness and clarity, whereas more impressionistic or expressive subjects, especially those depicting large unmodulated masses of shadows or highlights, are best suited for papers with rough, broadly open textures (Eastman Kodak Company c.1935a, c.1935b). A result of a careful and deliberate manufacturing process, texture applied to silver gelatin paper is designed to be distinct and distinguishable through processing and post-processing procedures. Likewise manufacturer-applied texture endures despite localized defects such as abrasions and deterioration caused by poor handling, storage environment, and enclosures. Given these attributes, an encyclopedic collection of surface textures could reveal vital clues about a photographic print of unknown origin. Ultimately a method for classifying textures could provide a means to link prints to specific photographers or to other prints of known provenance.

Previous work (Messier et al. 2009; Parker and Messier 2009) established the practicality of the image data collection procedure used in this study and suggested that more sophisticated approaches to automated classification could yield a highly reliable texture matching methodology.

2. Materials and Methods

2.1 A collaboarative competition

As part of the Museum of Modern Art's (MoMA) Thomas Walther Collection project, close-up raking light images were made from 328 modernist photographs within its collection. Begun in 2008, the overarching research goal of the Walther project was to advance the scholarship about dating and characterization techniques for 20th-century photographic materials and establish a new model for collaborative research, interpretation, and inter-disciplinary dialogue (Daffner 2013). Combined with the prior image classification work cited above, the MoMA texture images stimulated interest in developing an automated scheme to cluster like prints based on surface texture. Initiating a collaborative competition, an appeal was made to university teams to develop methods for sorting texture images. Invited teams had both signal-processing experience and an established interest in developing cultural heritage applications. Four university teams joined the project. Table 1 lists the teams and their computational approaches described in Section 3.

2.2 Texture image preparation

The close-up texture images were acquired with a microscope system assembled using an Infinity 2–3 imager manufactured by the Lumenera Corporation fitted with an Edmund Optics VZM 200i lens, as shown in figure 1. The imager uses an Interline Sony ICX262 3.3 megapixel color progressive scan CCD

FIG. 1. System for raking-light close-up image acquisition.

sensor producing images that incorporate 1536 × 2080, 3.45 μm square pixels. The imaged area on each sample measured 1.00 × 1.35 cm. Raking light close-up images were made with a fixed-point illumination source using a 3-in. LED line light manufactured by Advanced Illumination, placed at a 25° raking angle to the surface of the photographic paper. Each raking light image generated an 18.2 MB, 16-bit, TIFF. Typical images are shown in figure 2, depicting papers manufactured by Ilford and Agfa. The image capture technique is non-contact/non-destructive and therefore easily adapted for use on photographic prints of high intrinsic value. It is also relatively quick and requires minimal specialized handling, meaning large image sets can be produced rapidly (excluding museum registration and logistical issues, the imaging work at MoMA on Thomas Walther collection prints was conducted over a period of three days).

2.3 DATASETS

Sample papers dating from 1908 to 1977 were selected from a large reference collection of photographic paper. Each sample was identified by manufacturer, brand, date, and manufacturer-assigned surface designation. The reference collection and the methods used to identify the samples have been described elsewhere (Messier et al. 2005; Connors-Rowe et al. 2007).

Each of the four teams was given preliminary datasets composed of 50 sample textures with some

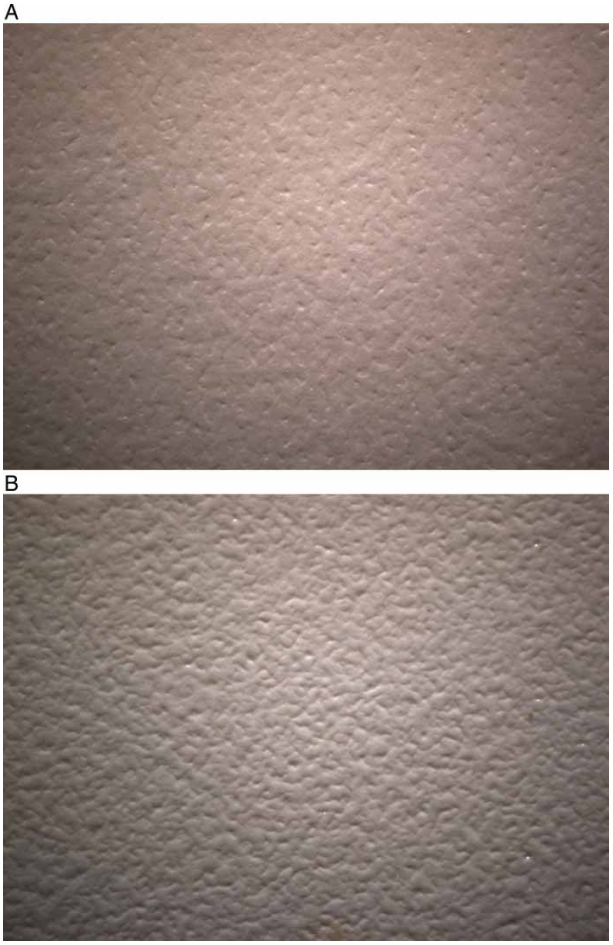

FIG. 2. Examples of raking-light close-up images. Sample 57, Ilford, Plastika, Special Grained, Half Matt, expired in 1940, Messier Reference # 944 (top) and sample 83, Agfa, Brovira, B 119 Crystal, Lustre, expired 1955, Messier Reference # 1791. The surface area of the samples measures 1.00 × 1.35 cm.

known matches. Using this set, the teams developed prototype classification algorithms. Findings from this work established basic parameters that informed the construction of the finished datasets described below: the orientation of the primary paper fiber direction relative to the raking light had no significant impact on results (this finding does not exclude *a priori* that photo paper textures may possess other forms of anisotropy), the depicted surface area and resolution held sufficient texture-related information to enable classification, and the angle of illumination rendered relief without causing large over- and underexposed passages (and subsequent loss of data).

This preliminary work resulted in the design of a 120-sample dataset of raking light images of photographic papers with known metadata including manufacturer, brand, date, gloss, and texture classification, and offering varying degrees of self-similarity. The

Appendix lists all samples used in this study. Using only this metadata, an "expert observer" knowledgeable in the history of photographic papers and able to interpret manufacturer terminology, ranked texture affinities for each possible pairing of samples. For example, the same textures made by the same manufacturer during the same time period would be expected to have a high level of affinity where a pair of textures identified as having a regularly patterned "linen" surface compared to a smooth surface would have a low affinity.

Thus classified, the dataset delivered to the teams for testing was largely composed of nine "affinity groups" of ten paper samples each. Within these groups, there were three similarity subsets: (1) images made from the same sheet of paper, (2) images made from sheets taken from the same manufacturer package of paper, and (3) images from papers made to the same manufacturer specifications over a period of time. A fourth subset, composed of 30 samples, was assembled without concern for texture similarity but instead was selected to span the range of textures associated with historic silver gelatin paper.

Applied to this dataset, the expert observer assessment reflects conventional wisdom that any close-up raking light image taken from different spots on a single sheet of paper should appear nearly identical. Likewise, texture images from different sheets of paper taken from the same manufacturer package also should show strong similarity. Furthermore, shots from papers manufactured to the same specifications but made at different times should show strong similarity, but to a somewhat lesser degree. For the thirty remaining samples, selected to demonstrate diversity, some would appear similar to the group of 90 textures and some would appear to be unique. The challenge posed to the teams was to discover the similarity groupings and mismatches evident to an expert observer.

3. Technical Approaches

The approaches taken by the four teams can be divided into two categories (Haralick 1979; Gonzalez and Woods 2008) based on the approach to feature definition: (1) non-semantic/Wisconsin and Tilburg and (2) multiscale/Lyon and WPI. The fundamental difference is that non-semantic features are derived directly from a case-by-case analysis of the image data, where the multiscale approaches are based on fixed structural models considered relevant to the encountered data. Accordingly, methods developed by Wisconsin and Tilberg extract and analyze a large number of random patches for each sample, whereas Lyon and WPI analyze the samples sets through the application of predetermined features. These techniques and many of the other terms used in this section are described in the glossary.

3.1 University of Wisconsin, Madison

The Wisconsin method is based on eigentextures. In this approach, a collection of small patches are chosen from each photographic image. These patches are gathered into a large matrix and then simplified to retain only the most relevant eigendirections using a singular value decomposition (SVD) (Moon and Stirling 2000). The preparation stage consists of two steps:

1. For each imaged paper j, randomly pick N $p \times p$ pixel patches $X_{j,i} \in R^{p \times p}$ for $i = 1,2,\ldots,N$ (with $N = 2000$ and $p = 25$ in this case). Lexigraphically reorder the $X_{j,i}$ into column vectors $a_{j,i} \in R^{p^2}$.
2. Create matrices $A_j = [a_{j,1}\ a_{j,2} \ldots a_{j,N}]$ consisting of the N column vectors and calculate the SVDs $A_j = U_j \Sigma V_j$ for all j. Extract the m columns of U_j corresponding to the m largest singular values and call this $\boldsymbol{U_j}$ (with m selected as 15 in this case).

The $\boldsymbol{U_j}$ are the representatives of the classes and may be thought of as vectors pointing in the most-relevant directions. During the classification stage, a number of similarly sized patches are drawn from the tested photographic paper. Each of these patches is compared to the representatives of the classes via a least squares procedure.

3. Select Q (with $Q = 2000$ used here) $p \times p$ pixel patches Q_i from the tested paper and reorder into vectors $q_i \in R^{p^2}$. Calculate the distance from the ith patch to the jth class:

$$d(i,j) = \left|\left| q_i - \boldsymbol{U}_j\left(\boldsymbol{U}_j^T q_i\right)\right|\right|_2 \qquad (1)$$

Every patch is closest to one of the classes, and the number of patches closest to the jth class is recorded.

4. For each patch i, $f_i = \text{argmin}_j\ d(i,j)$ locates the smallest of the $d(i,j)$, indicating that class j is the best fit for patch i. Tally the set of all such f_i, $i = 1, 2, \ldots, Q$.

The commonest entry among the f_i is the most likely class for this image. The second most common entry is the next most likely class for this image, etc.

3.2 Tilburg University

The method developed by Tilberg combines random features and textons, i.e., the random-feature texton method. This method was developed by Liu and Fieguth (2012) and is an adaptation of the texton approach (Varma and Zisserman 2009) using random features. Textons are prototypical exemplar image

patches capturing the "essence" of the texture of an image. Random-features (RFs) are random projections of image patches with $N \times N$ pixels to vectors with D elements ($N = 9$, $D = 20$, $D < N \times N$). More specifically, a RF is defined as a $D \times N^2$ matrix, the elements of which are sampled from the standard multivariate normal distribution $N(0,1)$.

The application of the random-feature texton method on the 120 sample dataset is conducted as follows. A set of X sub-images of $M \times M$ pixels is selected for each gray-value texture image in the 120 sample dataset ($M = 512$). The sub-images are defined to be the central regions of $M \times M$ pixels of which the intensity distributions are normalized to zero mean and unit variance. A sample of 45,000 randomly selected $N \times N$ ($N << M$) patches (represented as vectors of length N^2) of the normalized sub-images are contrast-normalized and subsequently multiplied with RFs, yielding RF vectors of length D.

Subsequently, a texton dictionary is created by applying k-means clustering to all RF vectors of the X sub-images of each texture image of the 120 sample dataset. Each image of the dataset is transformed into a texture histogram by comparing all of its patches (represented as RF vectors) to the entries in the texton dictionary. Finally, the histograms are classified using a k-nearest neighbor algorithm using the χ^2 similarity measure.

3.3 ECOLE NORMALE SUPERIEURE DE LYON

The method developed by the Ecole Normale Superieure de Lyon relies on the use of the Hyperbolic Wavelet Transform (HWT) (Devore et al. 1998; Roux et al. 2013) which is a variation of the 2D-Discrete Wavelet Transform (2D-DWT) (Mallat 2009). The HWT explicitly takes into account the possible anisotropic nature of image textures. Indeed, instead of relying on a single dilation factor a used along both directions of the image (as is the case for the 2D-DWT), HWT relies on the use of two independent factors $a_1 = 2^{j_1}$ and $a_2 = 2^{j_2}$ along directions x_1 and x_2, respectively. The hyperbolic wavelet coefficients of imaged paper i, denoted as $T_i((a_1, a_2), (k_1, k_2))$ are theoretically defined as:

$$T_i((a_1, a_2), (k_1, k_2)) =< i(x_1, x_2), \frac{1}{\sqrt{a_1 a_2}} \psi\left(\frac{x_1 - k_1}{a_1}, \frac{x_2 - k_2}{a_2}\right) > \quad (2)$$

From these HWT coefficients, structure functions, consisting of space averages at given scales a_1, a_2, are defined as:

$$S_i((a_1, a_2), q) = \frac{1}{n_a} \sum_{\underline{k}} |T_i((a_1, a_2), (k_1, k_2))|^q, \quad (3)$$

where n_a stands for the number of $T_i((a_1, a_2), (k_1, k_2))$ actually computed and not degraded by image border effects.

To measure proximity between two images i and j, a cepstral distance between their structure functions $S_i((a_1, a_2), q)$ and $S_j((a_1, a_2), q)$ is computed. It consists of a classical L^p norm computed on log-transformed normalized structure functions:

$$d(i, j) = \left(\sum\nolimits_a |S_i(a, q) - S_i(a, q)|^p\right)^{\frac{1}{p}} \quad (4)$$

with

$$\tilde{S}_i(a, q) = \ln \frac{S_i(a, q)}{\sum_{a'} S_i(a', q)} \quad (5)$$

3.4 WORCESTER POLYTECHNIC INSTITUTE

The method developed by Worcester Polytechnic Institute used an area-scale analysis technique, which has been applied to various problems in surface metrology (Brown et al. 1993). Much as the measured length of a coastline depends on the scale of observation and therefore the resolvability of small features, the measured area of a surface is also a function of the scale of observation. The area-scale approach uses fractal analysis to decompose a surface into a patchwork of triangles of a given size. As the size of the triangles is increased, smaller surface features become less resolvable and the "relative area" of the surface decreases. The topological similarity of two surfaces is computed by comparing relative areas at various scales. The technique has traditionally been employed on topographic datasets containing height information over a surface. Though lacking a direct measure, area-scale analysis can be applied to the images using light intensity as a proxy for height.

The proposed approach proceeds in three steps: (1) preprocessing, (2) feature extraction, and (3) classification. The preprocessing step extracts a square $N \times N$ region from the center of the image (where N was chosen to be 1024), and normalizes the intensity of the resulting extracted image. The $N \times N$ grid of equally spaced points (representing pixel locations) is decomposed into a patchwork of

$$2\left(\frac{N-1}{s}\right)^2 \quad (6)$$

isosceles right triangles where s is a scale parameter representing the length of two legs of each triangle. The pixel values at each of the triangle vertices are then taken as the "pseudo-height" of each of the vertices. The area of each triangle in 3-D space is then computed and the areas of all triangular regions are summed, resulting in the total relative area A_s at the chosen scale s. To conduct feature extraction, the

relative area for an image is computed over a range of scale values; in this study, 8 scale values were used ranging from 1 to 34 pixels, which correspond to lengths of 6.51 μm to 0.221 mm, respectively. Finally, to classify and compare the similarity of two images i and j, a χ^2 distance measure $d(i,j)$ is computed via

$$d(i,j) = \sum_{s \in S} \frac{(A_s^{(i)} - A_s^{(j)})^2}{A_s^{(i)} + A_s^{(j)}} \qquad (7)$$

where $A_s^{(i)}$ is the relative area of image i at scale s and S is the set of chosen scale values. Small values of $d(i,j)$ indicate high similarity between images i and j, while large values indicate low similarity.

4. Results as Affinity Maps

Based on the results from the expert observer and each team's automatic classifiers, the degree of similarity (affinity) was tabulated for each possible pairing of images in the 120 sample dataset. These scores were then converted to a gray-scale with the darkest intensities indicating the greatest affinity and the lightest the least affinity.

To visualize these values a table containing 120 rows and 120 columns was created, one row and column for each sample in the dataset. Each of the resulting 14,400 cells in the table was shaded according to the similarity of compared samples with black describing an exact match, white a total mismatch and gray-scale values in between describing a range of better or worse similarities. For example, the top diagram in figure 3, shows predicted similarities within the sample group suggested by the metadata including manufacturer, texture, brand, and date. As expected, the nine dark blocks starting in the upper left and continuing down along the diagonal show a high degree of affinity (dark gray and black) as these blocks depict the nine groups of similar textures. Lesser degrees of similarity are scattered throughout the figure with the 30 samples selected to show diversity (poorer levels of similarity) falling in the lower right quadrant and along the right side and bottom edge.

Gray-scale affinity maps produced to display the results from each of the four teams are also shown in figure 3. The principal similarity among the five affinity maps in figure 3 are the nine dark squares along the upper left to lower right diagonal. Given the construction of the dataset, these blocks should be dark due to the high affinity of the samples in these groups. The light stripes in the right and bottom quarters of the affinity maps, due to some relatively matchless textures among samples 91–120, are also shared by all five affinity maps. While small local differences among the five maps indicate that work remains to find an ideal automated scheme, striking fundamental similarities between the metadata-based affinity map and the four produced by automated schemes indicate

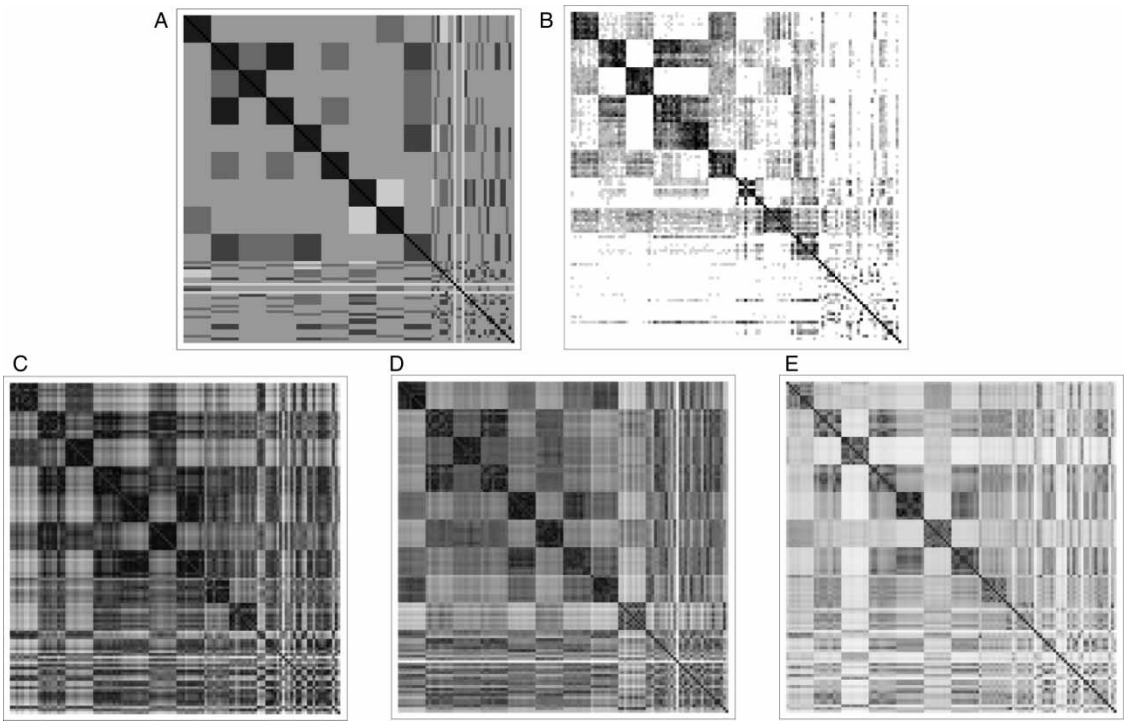

Fig. 3. Affinities (dark: strong, light: weak) for each possible pairing of surface texture images. Expert domain knowledge (A), Wisconsin (B), Tilburg (C), Lyon (D), WPI (E).

raking light close-up images have sufficient texture information to support the automated classification of historic photographic papers.

5. Observations

As shown in figure 3 there is a relatively high level of agreement between the affinity pairings prepared by the classification algorithms and those derived from metadata and subject-matter expertise. As discussed in the previous section, the principal correspondence among the five affinity maps is the nine dark squares along the diagonal running from upper left to lower right. Given the construction of the dataset, the samples in these blocks are very similar and these texture affinities were recognized both by a subjective metadata sort and by the four automated solutions. In addition, both the expert observer and automated solutions were sensitive to the increased levels of diversity within samples 81–90 (ninth dark block on the diagonal) that track a single proprietary surface over a fairly extended period of 22 years. Besides the nine similarity groupings added to the dataset by design, both the expert observer and the automated solutions discovered another strong affinity between subsets 11–20 and 31–40 (shown along the cross-diagonal axis adjoining the third dark square on the diagonal). As shown in the Appendix, these samples have the same manufacturer, brand, surface designation, and date but are taken from different paper packages.

These findings are reinforced by figure 4, which shows a normalization of the distances between each texture pairing within the tested groups. The shape of the curves is remarkably consistent with the automated solutions and the expert observer detecting very similar degrees of affinity across the groups. The chart confirms that there is no measurable difference between texture images made from the same sheet of paper as compared to images made from different sheets from the same manufacturer package. Further, depending on the technique, textures within the same manufacturing standard produced over time show fair-to-good levels of similarity. These results, though not a surprise given high levels of manufacturing regularity, are important for the possible development of future systems that rely on indices of known "exemplar" textures to identify unknowns.

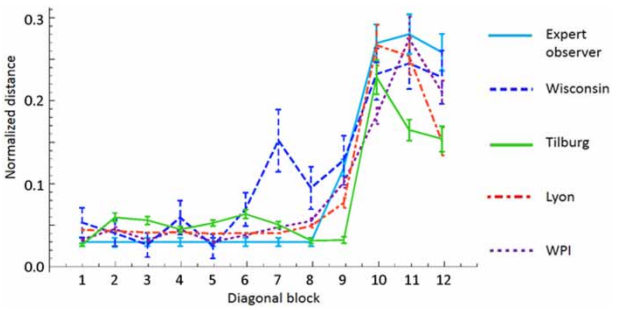

Fig. 4. Normalized and averaged image pair distances with standard deviations for the dataset of texture images, diagonal blocks: 1–3 same sheet, 4–6 same package, 7–9 same manufacturing standard, 10–12 diverse samples.

6. Conclusions and Next Steps

This project opens a path toward a machine vision system that provides meaningful results for the study of photographic prints. To have meaning, an automated classification system cannot produce results simply based on an internal, self-referential "sameness/difference" parameter, but instead must render results that are relevant to trained practitioners such as conservators and curators. For example, the images made from 10 spots on the same sheet of paper, though totally different images, need to be recognized as the "same." Likewise the two other similarity groups made from different sheets from the same manufacturer package and from papers manufactured to the same standard must be recognized as related. A useful system needs to reliably cluster these groups together though be discriminating enough to set these groups apart from others that might have similar textures but, for example, are made by different manufacturers. Using different techniques, each of the four teams met this standard. The fundamental outcome of this experiment is the intuitive expert observer conception of a classification system based on sameness/difference can be replicated through imaging and signal-processing techniques.

Specific to each technique, more work is required to discover potential strengths and weaknesses. Though all techniques produce results similar to those predicted by the expert observer, subtle differences among the techniques are possible and have yet to be determined. For example, one technique may show advantages over another when it comes to evaluating different types of surfaces (e.g., patterned versus random feature distribution). Most likely fully realized systems for classifying textures would incorporate the means to toggle between each of these techniques as well as to aggregate results across multiple methods.

Such systems could engender new modes of scholarship based on the discovery of material-based affinities. Work at MoMA is underway to determine how these techniques might meaningfully be applied to prints in its Thomas Walther collection. Moving forward,

reference libraries of surface textures, containing papers grouped by photographer or paper manufacturer can be assembled and used as a basis of comparison. This work is already underway through the assembly of large paper reference collections categorized by manufacturer, brand, surface finish, and date as well as for individual artists including Man Ray (1890–1976) and Lewis Hine (1874–1940). With standardized imaging techniques and a networked infrastructure, conservators could query such texture libraries to detect similar papers held by other collections, potentially characterizing and identifying works in their collection as well as revealing relationships within an artist's body of work and between artists. These methodologies are being applied to other media, including ongoing work with ink-jet papers (Messier et al. 2013) and the platinum papers of F. Holland Day (1864–1933).

A website, www.PaperTextureID.org, has been created to distribute the dataset of silver gelatin textures used as the basis for this study. An ink-jet paper dataset is also posted at this site. The availability of these image sets should encourage other teams to develop their own automated classification and sorting schemes.

Acknowledgments

The authors wish to thank Jill Sterret and Theresa Andrews of the San Francisco Museum of Modern Art for providing a meeting venue during the summer of 2012, Andrew Messier, Lincoln Laboratories, Massachusetts Institute of Technology, for help designing the imaging system, and Ian Holland, Lumenera Technical Assistance Center, for assistance with imaging specifications.

Glossary

2D-Discrete Wavelet Transform (2D-DWT). The 2D-DWT is a classical image processing tool that provides a multiscale, exact, and invertible representation of an image by assessing the collection of image versions band-pass filtered at different scales.

Cepstral distance. The word cepstrum was invented in 1963 and derives from a reordering of the first letters in spectrum. A cepstrum is the result of analyzing signals or images with (a normalized version of) the logarithm of the absolute value of a classical spectrum, so as to better permit comparisons in terms of change rate, magnitude, phase, power, and other features. The cepstral distance is the net result of this comparison.

Eigentextures. For a given image, eigentextures are the columns of an imposed matrix that form the basis for a set of all depicted textures. The eigentextures provide a rank ordering of the most significant of these sub-spaces, thus providing a reduced dimensional approximation, or distillation, of the texture pattern.

***k*-means clustering.** *k*-means clustering is an adapted signal-processing technique used to analyze vectors and other complex variables for the purposes of modeling large datasets and to derive affinity groupings. *k*-means clustering is often used for data mining and machine learning applications.

Least Squares. Least squares is a standard method for data fitting and regression analysis. Data fitting refers to the determination of a curve or mathematical function that best describes (or "fits") a set of data points according to a particular measure, i.e. the sum of squared errors to each data point. Similarly, regression analysis is used to estimate the relationships between variables.

Multiscale. Multiscale methods refer to a broad range of signal processing and classification techniques that analyze data jointly or simultaneously at different scales or resolutions. Often, multiscale techniques rely on fitting and comparing data against exemplar structural building blocks, also at different scales.

Non-semantic. Non-semantic methods also refer to a broad range of signal processing and classification techniques. As opposed to multiscale techniques, non-semantic methods are based on values measured directly from the sample under consideration versus the application of prototypical patterns and data structures.

Hyperbolic wavelet transform (HWT): The hyperbolic wavelet transform extends the 2D-DWT by allowing the use of different dilation factors (changes of scale) on the horizontal and vertical axes of a data matrix. Image analysis using HWT allows both multiscale and anisotropic feature extraction.

Singular value decomposition. Singular value decomposition is a mathematical method for a data-driven derivation of substructures, patterns, correlations, and variations from complex, multidimensional data matrices.

Textons. Textons are the most basic texture elements that, when repeated, fully define an image depicting a textured surface. These fundamental micro-structures are often described as the "atoms" of texture perception.

Appendix: Paper Samples Used In The Silver Gelatin Dataset

For the following table ID is the sequential numbering system suggested by the teams following image processing. Date refers to the paper expiration dates applied to manufacturer packages or estimates made based on packaging. M ID is the Messier Reference Collection catalog number. Other descriptors, such as brand and paper characteristics, are taken directly from the manufacturer packaging.

ID	Manufacturer	Brand	Texture	Reflectance	Date	M ID
10 samples from the same sheet (×3 sheets)						
1–10	Kodak	Vitava Athena	C – Smooth	Matte	1943	10
11–20	Kodak	Kodabromide	E – Fine grain	Lustre	1967	2952
21–30	Leonar	Rano Kraftig	N/A	Chamois	1910	4869
10 samples from the same package (×3 packages)						
31–40	Kodak	Kodabromide	E – Fine grained	Luster	1967	2216
41–50	Ilford	Contact	1P	Glossy	1955	2750
51–60	Ilford	Plastika	Special grained	Half Matt	1940	944
10 samples from the same manufacturer surface finish (×3 manufacturers)						
61	Kodak	Velox	F – Smooth	Glossy	1938	97
62	Kodak	Kodabrom	F – Smooth	Glossy	1939	1019
63	Kodak	Kodabrom	F – Smooth	Glossy	1939	1020
64	Kodak	Azo	F – Smooth	Glossy	1931	1503
65	Kodak	Azo	F – Smooth	Glossy	1935	1530
66	Kodak	Azo	F – Smooth	Glossy	1937	1531
67	Kodak	Azo	F – Smooth	Glossy	1937	1532
68	Kodak	Azo	F – Smooth	Glossy	1930	2370
69	Kodak	Vitava Athena	F – Smooth	Glossy	1928	2447
70	Kodak	[no brand]	F	Glossy	1930	2924
71	Dupont-Defender	Apex	A	Semi Matte	1947	112
72	Dupont-Defender	Velour Black	A	Semi Matte	1951	1427
73	Dupont-Defender	Velour Black	A	Semi Matte	1951	1434
74	Dupont-Defender	Velour Black	A	Semi Matte	1951	1435
75	Dupont-Defender	Velour Black	A	Semi Matte	1951	1436
76	Dupont-Defender	Velour Black	A	Semi Matte	1951	1440
77	Dupont-Defender	Varigam	A	Semi-Matt	1953	2302
78	Dupont	Varigam	A	Semi Matte	1958	2921
79	Dupont-Defender	Velour Black	A	Semi-Matt	1953	4842
80	Dupont-Defender	Velour Black	A	Semi-Matt	1953	4843
81	Agfa-Gevaert	Brovira	B 119	Lustre	1965	167
82	Agfa-Gevaert	Brovira	B 119	Luster	1976	1540
83	Agfa	Brovira	B 119 crystal	Lustre	1955	1791
84	Agfa	Brovira	B 119 crystal	Luster	1955	1838
85	Agfa-Gevaert	Brovira	B 1	Glossy	1965	2079
86	Agfa	Brovira	B 119 crystal	N/A	1960	2365
87	Agfa-Gevaert	Brovira	B 119	Lustre	1974	2438
88	Agfa-Gevaert	Brovira	B 119 – Filigran	Glossy	1964	2547
89	Agfa-Gevaert	Brovira	B 119 – Fine grained	Lustre	1964	2634
90	Agfa-Gevaert	Lupex	N/A	Glossy	1964	2640

Continued

CONTINUED

ID	Manufacturer	Brand	Texture	Reflectance	Date	M ID
30 samples showing diversity						
91	Defender	Argo	N/A	Matte	1912	2775
92	Ansco	Cyko	N/A	N/A	1918	971
93	Darko	Darko Developing Paper	N/A	Matte	1923	3205
94	Kodak	Velvet Velox	N/A	Semi gloss	1921	25
95	Agfa ansco	Convira	B	Glossy	1938	2306
96	Kodak	Velox	F – Smooth	Glossy	1944	2277
97	Kodak	Kodabromide	G – Fine grained	Lustre	1953	1851
98	Unicolor	B & W	N/A	N/A	1973	2234
99	Ansco	Cyko	Linen	Buff	1914	321
100	Darko	[no brand]	N/A	Velvet	1924	3208
101	Kodak	Carbon Velox	N/A	Matte	1908	98
102	Agfa ansco	Cykora	N/A – Silk	N/A	1948	203
103	Kodak	Ektamatic SC	F – Smooth	Glossy	1977	2626
104	Kodak	Azo	A – Smooth	Luster	1916	235
105	Defender	Argo	N/A	Normal gloss	1916	1444
106	Kodak (Canadian)	Azo	F	Glossy	1926	1981
107	Ansco	Cyko	N/A	Buff	1925	994
108	Defender	Veltura	N/A	Matte	1932	38
109	Ilford	Clorona	Porcelain	N/A	1938	2761
110	Kodak	Velox	F – Smooth	Glossy	1946	1040
111	Delaware Photographic Co.	Enlarging Paper	N/A	Semi matte	1940	9
112	Kodak	Kodabromide	E – Fine grained	Buff luster	1950	1709
113	Kodak	Panalure	F – Smooth	Glossy	1969	2623
114	Agfa-Gevaert	Brovira	B 111	Glossy	1975	335
115	Agfa ansco	Brovira	B 119	Lustre	1974	2439
116	Agfa ansco	Convira	N/A	Glossy	1950	857
117	Kodak	Kodabromide	F – Smooth	Glossy	1959	864
118	Defender	Veltura	N/A	Matte	1932	40
119	Agfa-Gevaert	Brovira	B 119	Lustre	1974	2438
120	Agfa ansco	Cykora	Kashmir	N/A	1948	204

REFERENCES

Brown, C., P. Charles, W. Johnsen, and S. Chesteras. 1993. Fractal analysis of topographic data by the patchwork method. *Wear* 161: 61–67.

Connors-Rowe, S., P. Whitmore, and H. Morris. 2007. Optical brighteners in black-and-white photographic paper: appearance and degradation. *Journal of the American Institute for Conservation* 46: 199–213.

Daffner, L. 2013. Abstract: The proof is in the print – characterization and collaboration in the Thomas Walther Collection Project at the Museum of Modern Art. *Topics in Photograph Preservation* 15: 25.

Defender Photo Supply Company. Circa 1935. *Velour black specimen prints*. Rochester, NY: Defender Photo Supply Company.

DeVore, R.A., S.V. Konyagin, and V.N. Temlyakov. 1998. Hyperbolic wavelet approximation. *Constructive Approximation* 14: 1–26.

Eastman Kodak Company. Circa 1935a. *Surface characteristics of Kodak photographic papers*. Rochester, NY: The Eastman Kodak Company.

Eastman Kodak Company. Circa 1935b. *Commercial and illustrative photographic papers*. Rochester, NY: The Eastman Kodak Company.

Gonzalez, R.C., and R.E. Woods. 2008. *Digital image processing*. 3rd ed. Upper Saddle River, NJ: Prentice Hall.

Haralick, R.M. 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE* 67: 786–804.

Liu, L., and P.W. Fieguth. 2012. Texture classification from random features. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34: 574–86.

Mallat, S. 2009. *A wavelet tour of signal processing: the sparse way*. 3rd ed. Burlington, MA: Academic Press.

Messier, P. 2008. Les Emulsion Industrielles. In *Le Vocabulaire Technique de la Photographie*, ed. A. Cartier-Bresson. Paris: Les Editions Marval. 454–456.

Messier, P., V. Baas, D. Tafilowski, and L. Varga. 2005. Optical brightening agents in photographic paper. *Journal of the American Institute for Conservation* 44: 1–12.

Messier, P., M. Messier, and C. Parker. 2009. Query and retrieval systems for a texture library of photographic papers (abstract). *Proceedings of the International Conference on Surface Metrology*, Worcester, MA: Worcester Polytechnic University. I: 10.

Messier, P., R. Johnson, H. Wilhelm, W. Sethares, A. Klein, P. Abry, S. Jaffard, H. Wendt, S. Roux, N. Pustelnik, N. van Noord, L. van der Maaten, and E. Postma. 2013. Automated surface texture classification of inkjet and photographic media. *Technical Program and Proceedings: NIP29: The 29th International Conference on Digital Printing Technologies*, Springfield, VA: IS&T: The Society for Imaging Science and Technology. 85–91.

Mimosa. Circa 1930. *Mimosa papiere*. Dresden: Mimosa AG.

Moon, T.K., and W.C. Stirling. 2000. *Mathematical methods and algorithms for signal processing*. Upper Saddle River, NJ: Prentice Hall.

Parker, C., and P. Messier. 2009. Automating art print authentication using metric learning. *Proceedings of the Twenty-First Innovative Applications of Artificial Intelligence Conference*. Palo Alto, CA: Association for the Advancement of Artificial Intelligence. 122–127.

Roux, S.G., M. Clausel, B. Vedel, S. Jaffard, and P. Abry. 2013. Self-similar anisotropic texture analysis: The hyperbolic wavelet transform contribution. *IEEE Transactions on Image Processing* 22: 4353–4363.

The Gevaert Company, Inc. Circa 1935. *The book of Gevaert paper samples*. New York, NY: The Gevaert Company of America.

Varma, M., and A. Zisserman. 2009. A statistical approach to material classification using image patch exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31: 2032–2047.

Author Biographies

C. Richard Johnson, Jr. received a PhD in Electrical Engineering from Stanford University, along with the first PhD minor in Art History granted by Stanford, in 1977. He is currently the Geoffrey S. M. Hedrick Senior Professor of Engineering and a Stephen H. Weiss Presidential Fellow at Cornell University. Since 2005, his primary research interest has been computational art history. Professor Johnson founded the Thread Count Automation Project (TCAP) in collaboration with the Van Gogh Museum in 2007, initiated the Historic Photographic Paper Classification (HPPC) challenge in cooperation with the Museum of Modern Art in 2010, and launched the Chain Line Pattern (CLiP) Matching Project with the Morgan Library & Museum in 2012, with the Rijksmuseum and the Metropolitan Museum of Art joining the project in 2013. He was an Adjunct Research Fellow of the Van Gogh Museum from 2007 through 2011. In 2013, Professor Johnson was appointed a Scientific Researcher of the Rijksmuseum and Computational Art History Advisor to the RKD – Netherlands Institute for Art History. Address: School of Electrical and Computer Engineering, 390 Rhodes Hall, Cornell University, Ithaca, NY 14853. Email: johnson@ece.cornell.edu

Paul Messier is an independent conservator of photographs working in Boston, Massachusetts. Founded in 1994, his studio provides conservation services for private and institutional clients throughout the world. The heart of this practice is unique knowledge and ongoing research into photographic papers. The Messier Reference Collection of Photographic Papers plays a vital role in this work. He received a Masters of Arts and certificate of advanced study in the conservation of works on paper from the art conservation program at the State University College at Buffalo (SUNY). Messier is the corresponding author. Address: 103 Brooks Street, Boston, MA 02135. Email: pm@paulmessier.com

A full list of author biographies for this paper can be found at http://www.maneyonline.com/doi/suppl/10.1179/1945233014Y.0000000024.S1

Résumé – La texture de la surface du papier photographique est une caractéristique essentielle de sa fabrication, de sa commercialisation et de son utilisation. La lumière rasante révèle la texture du papier par un rendu saisissant des lumières et des ombres. Bien que les gros plans en lumière rasante documentent efficacement les caractéristiques de la surface des papiers photographiques, le nombre et la diversité des textures des papiers historiques empêchent une classification visuelle efficace. Le présent travail démontre que la classification automatique par ordinateur de textures documentées avec une lumière rasante est possible, en démontrant un degré encourageant de succès dans le tri d'un ensemble de 120 images réalisées à partir d'échantillons de papiers historiques au gélatino-bromure d'argent. À partir de ce groupe d'images, quatre équipes universitaires ont appliqué différentes stratégies de traitement d'image afin d'extraire automatiquement des caractéristiques et de quantifier le degré de similitude des papiers. Les quatre approches ont chacune réussi à détecter de fortes affinités, ainsi que des images avec caractéristique

aberrante intégrées dans le groupe. La création et le déploiement des algorithmes ont été réalisés par les équipes sans connaissance préalable de la répartition des similitudes et des images aberrantes. Ces résultats indiquent que la classification automatique des papiers photographiques au gélatino-bromure d'argent basée sur des gros plans de texture est possible et doit être poursuivie. Pour encourager le développement d'autres systèmes de classification, les données des 120 échantillons utilisées dans ce travail sont mises à la disposition d'autres chercheurs scientifiques à l'adresse suivante: www.PaperTextureID.org.

Resumen – La textura superficial es un rasgo crítico en la manufactura, comercialización y uso de papel fotográfico. La luz rasante revela la textura a través de una representación nítida de áreas luminosas y sombras. A pesar de que las imágenes con luz rasante documentan efectivamente las características superficiales del papel fotográfico, la cantidad misma y la diversidad de texturas usadas en papeles históricos hace imposible una clasificación visual eficiente. Este trabajo proporciona evidencias de que la clasificación automática de la textura documentada con luz rasante, realizada por computadora es factible al demostrar un motivante grado de éxito al clasificar un grupo de 120 imágenes hechas de muestras de papel histórico de gelatina de plata. Usando esta base de datos, cuatro equipos universitarios aplicaron diferentes estrategias de procesamiento de imágenes para la extracción automática de rasgos y cuantificación del grado de similitud. Las cuatro estrategias detectaron con éxito afinidades fuertes así como los casos atípicos incluidos en la base de datos. Los equipos crearon y usaron algoritmos sin conocimiento previo de las similitudes y los casos lejanos al promedio. Estos resultados indican que la clasificación automática de papel fotográfico de gelatina de plata basada en imágenes que muestran acercamientos de la textura es viable y debería realizarse. Para incentivar el desarrollo de otros esquemas de clasificación el grupo de 120 muestras de "entrenamiento" usado en este trabajo está disponible para otros investigadores en www.PaperTextureID.org.

Resumo – Textura de superfície é um elemento crucial na fabricação, comercialização e uso de papéis fotográficos. Luz rasante revela a textura através da renderização acentuada de altas luzes e sombras. Ainda que imagens feitas com luz rasante em close-up efetivamente documentem as características da superfície dos papéis fotográficos, a enorme quantidade e diversidade de texturas utilizadas nos papéis históricos impedem a classificação visual eficiente. Esse trabalho fornece evidências de que a classificação de texturas documentadas com luz rasante em base de computador, automatizada, é viável pela demonstração bem sucedida e promissora da ordenação de um conjunto de 120 imagens feitas de amostras de papel histórico de gelatina e prata. Com esse conjunto de dados, equipes de quatro universidades aplicaram estratégias diferentes de processamento de imagem para extração automatizada de características e quantificação de graus de semelhança. Todas as quatro abordagens detectaram com sucesso relevantes afinidades e variações atípicas (desvios-padrão) incorporadas ao conjunto de dados. A criação e implantação de algoritmos foram realizadas por equipes sem conhecimento prévio das distribuições de similaridades e de desvios-padrão. Os resultados indicaram que a classificação automatizada de papéis fotográficos de gelatina e prata fundamentada nas imagens em close-up das texturas é possível e deve ser alcançada. Para estimular o desenvolvimento de outros esquemas de classificação, o conjunto "teste" de dados com 120 amostras utilizado nesse trabalho está disponível para outras pesquisas acadêmicas no www.PaperTextureID.org.