

ROGER L. EASTON, JR., AND DAVID KELBE

Multispectral Image Collection and Statistical Processing of Palimpsests at St. Catherine's Monastery*

INTRODUCTION

The text in many parchment manuscripts of historical value often is difficult to read. This is particularly true for manuscripts that have been deliberately erased and overwritten, thus forming 'palimpsests'. In such cases, the contrast between the erased ink and the parchment as seen by the eye usually is reduced to the point where it may be difficult or impossible for human readers even to identify the existence of erased characters. This is due in part to limitations of the human visual system, such as the relatively narrow range of spectral sensitivity of the retinal cells of the eye. The eye is sensitive to wavelengths λ only in the interval from blue light ($\lambda \approx 400$ nm) to red light ($\lambda \approx 700$ nm). Because useful information about the text often exists outside of this range, this is a significant constraint.

Fortunately, erased or damaged writings often may be recovered from images obtained with modern sensors, which can record light over a wider range of wavelengths than that to which the human eye is sensitive. The most common photoelectronic detectors are based on silicon and are sensitive from the ultraviolet part of the spectrum ($\lambda \approx 350$ nm) through the visible region ($400 \text{ nm} \leq \lambda \leq 700 \text{ nm}$) and into the infrared region out to $\lambda \approx 1100$ nm. Sensors based on other materials or compounds, such as indium antimonide (*InSb*) or mercury-cadmium-telluride (*HgCdTe*), are capable of recording light with yet longer wavelengths, but cost and other factors have kept these sensors from being widely applied to cultural heritage imaging.

Other characteristics of the human visual system also detract from the ability to read erased or damaged writing. The human eye-brain visual system is capable of producing useful images over a very wide range of ambient lightness levels (from bright sunlight to faint objects in the night sky), but this capability arises from involuntary physical and photochemical mechanisms that adapt to the ambient lightness level. The physical mechanism is the variation in the diameter of the eye pupil to admit more light from faint scenes, while different photochemical receptors are used at different light levels. The involuntary adaptation of human vision means that perceptions of feature brightness and visibility are not comparable to the perceived responses under other ambient conditions (this is the basis for some familiar 'optical illusions'). Beyond that, the human eye is capable of distinguishing only a limited range of lightness levels at any one level of brightness adaptation, which means that the 'intensity' of light from distinct features must be sufficiently different to be perceptible, even without the added complication of different ambient conditions. One additional observation deserves emphasis: the important point that images from the human eye are retained in human memory, whose familiar imperfections further complicate assessment of similarities or differences in different scenes viewed at different times; remembered appearance is rarely useful for comparisons. In short, it is very difficult for the human visual system to consistently perceive small differences in intensity from different features.

The advantages of modern photoelectronic sensors over the human visual system for the application to

Roger L. Easton, Jr.: Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, USA 14623-5604; easton@cis.rit.edu

David Kelbe: Xerra Earth Observation Institute, 50 Centennial Avenue, Alexandra, 9320 New Zealand; dave.kelbe@gmail.com

* The authors would like to acknowledge the assistance and inspiration provided by Michael Phelps, Damianos Kasotakis, Josephine Dru, and Keith Knox of the Early Manuscripts Electronic Library, Giulia Rossetto of the University of Vienna, Gregory Heyworth of the University of Rochester, and Jana Grusková of the Austrian Academy of Sciences. The students who performed the image processing included Allison Bright and Kevin Sacca of the Rochester Institute of Technology (RIT) and Nicole Polglaze of Grinnell College, who participated as part of the RIT 'Research Experience for Undergraduates' program funded by the National Science Foundation. The late Miss Catherine Carlson of Rochester, New York provided financial assistance that allowed undergraduate students to process the imagery of many of the leaves. This research was funded by a grant from the Arcadia Fund in the UK.

cultural heritage are diverse. First, sensors are capable of measuring smaller differences in light ‘strength’ over a wider range of ambient conditions than the human eye. In addition, the sensors also are not susceptible to variation in the perceived lightness levels due to the involuntary adaptation mechanisms that ‘confuse’ the eye. Put another way, modern sensors can ‘see’ light over a wider range of wavelengths and can repeatedly record more distinct ‘lightness’ levels (often called ‘gray levels’ or ‘digital counts’) than the human visual system, without being affected by psychological factors.

In addition, images measured by modern photoelectronic sensors are recorded as integer numbers in durable memory, and these numerical values may be compared or enhanced to make small differences more visible to human observers. Of even more significance, images recorded from camera systems may be combined digitally (by summing/averaging and/or differencing) or analyzed using statistical metrics, with the goal of enhancing the feature(s) of interest. This may enable the scholar to visualize very small differences in contrast between erased ink and parchment that are not visible to the eye alone.

SPECTRAL IMAGING SYSTEMS

The collection of imagery over wider ranges of wavelength bands than are visible to the eye has become rather common in studies of cultural heritage over the last two decades or so. These ‘spectral imaging systems’ are often classified depending on the number and configuration of the spectral bands and how light from each is dispersed into the spectrum of constituent wavelengths. ‘Multispectral’ imaging (MSI) systems collect a relatively small number of wavelength bands, up to a few dozen. These bands typically span fairly wide spectral ranges and may overlap or be disjoint. These systems restrict the measured light to specific bands either by applying different colours of illumination to the object or by placing ‘coloured’ filters at the camera lens that transmit different wavelength passbands through the lens to the sensor. Most such multispectral systems use two-dimensional sensor arrays and sequentially measure images in each spectral band of the entire object. The spatial resolution of these images is generally rather ‘fine’ (better than 25 pixels per mm, corresponding to 600+ pixels per inch).

‘Hyperspectral’ imaging (HSI) systems collect a much larger number M of image bands, where typically $20 \leq M \leq 500$ or so. These spectral bands are usually narrow and often contiguous, though often with coarser spatial resolution. Hyperspectral systems often illuminate the object with white light and disperse the reflected light into its spectrum by a diffraction grating or prism. In the most common systems, the spectra of all pixels along a ‘line’ of the object are measured simultaneously by dispersing the light from each pixel in the line into spectra that are measured on a two-dimensional sensor array. The object is translated past the entrance slit of the optical system to evaluate the spectrum of the next line on the object, and the process continues by scanning the object to build up its spectrum ‘line-by-line’ in sequence. The spectral resolution of hyperspectral systems can be quite fine and may be sufficient to enable identifying and quantifying constituent materials by comparing the observed response at a pixel to a library of recorded spectra. This capability to analyze constituent materials is a significant advantage for HSI systems, though the limited spatial resolution can be problematic. Though the processing methods considered here are applicable to HSI, and though these systems are being utilized more often for studies of cultural heritage, they will not be considered further here.

The last category of spectral imaging system is called ‘ultraspectral’. Such systems use an interferometer to obtain even finer spectral resolution, though the spatial resolution and field of view usually are quite limited. Ultraspectral systems are only now becoming available and their use, particularly in cultural heritage, is very uncommon, so they will not be considered further here.

Based on these generally-accepted definitions, the vast majority of imaging systems used for cultural heritage should be classified as ‘multispectral’, because of the relatively small number of ‘wide’ spectral bands with $\Delta\lambda \geq 40$ nm. As the costs of hyperspectral and ultraspectral systems decrease, we can expect to see more instances of their application to cultural heritage objects. For example, the Conservation Science Laboratory at the Bodleian Library of the University of Oxford has utilized the capability for true HSI for study of cultural heritage artifacts. Significant discoveries have been made about the Gough Map using this system.¹ However, the rest of this discussion is restricted to the use of multispectral imaging to recover erased text from the palimpsests in the collection at St. Catherine’s Monastery.

¹ Bai *et al.*, Hyperspectral analysis of cultural heritage artifacts.

IMAGE MODALITIES

All readers are familiar with images of light reflected from the surface of an object. Useful information is also obtained from images of light that interacts with the object in other ways. The potential value of imaging light transmitted through a palimpsest may be obvious, and has proven to be very useful for cases where ink acids have 'thinned' the flesh side of parchments. Also of great value is imaging of light generated by fluorescence, which is the physical mechanism where light of shorter wavelengths (often ultraviolet) is absorbed by parchment or ink, followed by emission of longer wavelengths. The number of photons generated by fluorescence is largest at wavelengths just longer than the stimulating light, and this number decreases as the wavelength is increased. The spectral character of the fluorescence often conveys important information about erased text, and may be captured by placing transparent coloured filters in the optical path prior to the sensor to reduce the range of imaged wavelengths. In the system at St. Catherine's Monastery, fluorescence is imaged through blue, green, orange, and red filters. These filters further reduce the already small fluorescence signal, often to the point where the light available for imaging is very faint. Such conditions can have significant detrimental consequences for the quality of the recorded images, as will be described shortly.

IMPORTANCE OF IMAGE QUALITY

To ensure the best possible recovery of erased text, it is imperative to record 'high-quality' images for subsequent processing, particularly in cases where the goal is to enhance very subtle differences among parchments and erased inks. The 'quality' of an image depends on the capabilities of both the camera lens and the sensor. The scientific study of 'image quality' is a broad subject, and its effects on processed images is an ongoing research area, so we will only touch on some significant points here.² The 'quality' of an image describes the ability to distinguish two characteristics of the scene from the recording: (1) fine spatial detail and (2) small differences in the 'lightnesses' of features in the scene. The first of these is called image 'sharpness', and is determined by the performance of the lens and by the sizes of pixels in the image sensor. The second attribute of image quality often is described by the metric 'signal-to-noise ratio' (SNR), which quantifies the precision of lightness measurements recorded at each pixel that may be expected from the system. From its very name, it is apparent that SNR is the arithmetic ratio of two different numerical measurements. The numerator in the ratio is the amplitude that ideally would be measured at a pixel based on the physical conditions. The 'noise' term describes the collective random variations in the measured amplitude that could arise from a number of thermal, electronic, or numerical ('statistical') mechanisms.

A preliminary discussion of the nature of light and its impact on imaging is necessary before considering further details of image quality. Note that some researchers restrict use of the term 'light' only to wavelengths visible to the human eye, while insisting that ultraviolet and infrared bands must be called 'electromagnetic radiation'. We adopt the less anthropocentric view that the term 'light' applies equally well to all electromagnetic radiation, whether visible or not.

There are three models of light that are applied in different circumstances: (1) light as a 'ray', (2) light as a 'wave', and (3) light as 'packets' of energy (called 'quanta' or 'photons'). The first model, unfortunately called 'geometrical optics', is the simplest and most common view of light. It assumes that energy propagates in straight lines, without spreading, until the light encounters a medium with a different velocity of propagation. At the interface between media, the propagation direction changes as determined by the well-known Snell's law for refraction. This first simple model of light propagation is used for designing lenses and optical systems, but does a poor job of explaining phenomena that are important in imaging, such as image 'sharpness' (also called 'spatial resolution').

The wave model of light (referred to by the equally unfortunate name of 'physical optics') is more mathematically complicated than the ray model, but also more complete because it explains the unavoidable 'spreading' of energy during propagation, which is commonly called 'diffraction'. This spreading of energy means that light energy 'bleeds' into regions that would be 'shadowed' in the ray model. This spreading results in the ultimate limit to spatial resolution of the image. A simplified view of spatial resolution may be visualized

² Peery *et al.*, Image Quality in Cultural Heritage.

using a model of the scene as consisting of two light sources of infinitesimal size—so-called ‘point sources’ of light. These two sources are separated by a dark space of variable width. The desired ideal image would be an exact replica of the two points of infinitesimal size for any possible spatial separation, which would mean that the two sources would always be distinguishable from the recorded image. This fortunate situation is physically impossible due to the wave nature of light, which ensures that light does not propagate as straight ‘rays’ of energy, but rather the energy ‘spreads’ during propagation. One result is that the images of the two point sources formed by a lens are themselves finite in extent instead of the infinitesimal point images in the ray model. The consequence is that the finite-sized image ‘spots’ generated from two closely-spaced point objects will be created at the sensor, and may overlap if the sources are sufficiently close together, to the point where these overlapping spots may not be distinguishable from the recorded image. The smallest separation at which the images of two point objects often is called the ‘spatial resolution’ of the imaging system and is determined by both the performance quality of the lens and by the properties of the sensor. The ultimate limit to spatial resolution often is the size of the diffraction ‘blur spot’ generated from a point source by the lens. This size is proportional to the wavelength of the light and the ‘focal ratio’ or ‘f/number’ of the lens, which is the ratio of the focal length to the aperture diameter. It is also limited by imperfections in the lens design and fabrication (the ‘aberrations’ of the lens).

The third model of light as ‘quanta’ or ‘photons’ of energy is necessary in order to understand the action of sensors, whether of older photochemical emulsions or of modern photoelectronic devices. Einstein showed that the energy conveyed by a photon of light is proportional to the reciprocal of the wavelength, so that photons of short-wavelength light have larger energies than longer wavelengths. The ‘strength’ or ‘intensity’ of light is determined by the number N of photons that reach the sensor from the object, and statistical uncertainties in this number often result in the ultimate limit to the SNR. The energy of a photon is absorbed by the lattice of atoms (usually silicon) in the sensor. The energy is transferred to electrons, which form an electric current that may be measured, which produces an estimate of the strength of the incident light.

To illustrate the impact of the uncertainty in the number of quanta on the measured signal, consider the case where each pixel is exposed to a region of constant ‘lightness’ on the object, which should result in the same numerical measurements at each pixel. However, variations in the measurements introduce error in these numbers, which is quantified as ‘noise’, and which affects the measured values. For example, the uncertainty may be due to unavoidable variations in the number of photons incident on each pixel during the exposure time. For statistical reasons beyond the scope of this paper, the statistical uncertainty in the number of photons N counted at each pixel is the square root $\sqrt{N}$, which means that the SNR is the ratio of the expected number N to the uncertainty $\sqrt{N}$, so the SNR is $\sqrt{N}$. Consider the simple example of a sensor exposed to the field of photons from an object that uniformly reflects light so that each pixel ideally would count 10,000 photons. The statistical uncertainty in the number is $\sqrt{10,000} = 100$ photons, so the resulting SNR is $10,000/100 = 100$. For a faint object field that emits 100 photons in the same exposure time, as might happen in fluorescence imaging, the resulting $\text{SNR} = (100)/(\sqrt{100}) = 10$, smaller by a factor of ten. In words, there is a larger ‘uncertainty’ in the ‘strength’ of the signal as measured by the number of counted photons for weaker signals. Since the number of photons is usually small in spectral fluorescence imaging, so is the resulting SNR, and pixels that should measure identical signals can exhibit wide variations that are due simply to the photon statistics. In cases where the measurement uncertainty approaches the true signal differences, the task of enhancing the differences for assessment by scholars is significantly complicated. Such situations detrimentally impact the results of subsequent image processing intended to enhance differences between object features. The measured number may be increased by some combination of longer exposure times or larger pixels. It is therefore necessary to find the ‘sweet spot’ that relates pixel size to exposure time that produces adequate performance of both spatial and noise aspects. Efforts are underway to define metrics that would better characterize system performance by simultaneously considering both aspects of image quality.

We now return to the first aspect of image quality, that of ‘sharpness’, which we have stated is determined by the performance of the lens and the characteristics of the image sensor. We have already seen how the ability of the system to resolve closely spaced features is limited by the size of the diffraction spot, but it also is affected by the spacing of the pixels in the sensor; it is not possible to resolve two spots of light that are both incident upon the same single pixel. However, camera vendors often emphasize pixel size, number, and spacing as the primary determinant of image ‘sharpness’, because the sensor spacing is an easier concept for

the user to comprehend. The common strategy is to make sensors with larger numbers of smaller pixels and claim that the resulting camera produces better pictures. For example, one common sensor used in cultural heritage imaging has 50 million pixels, each being a square with linear dimension of 6 micrometres. A new sensor using a different measurement architecture has 100 million pixels with a linear dimension smaller than 3 micrometres. The manufacturer will always claim that a sensor with smaller pixels has better image sharpness than a sensor with larger pixels. There certainly is merit to this approach in many common imaging settings with plenty of light available, as in imaging of spectral reflectance, but the situation for fluorescence imaging of manuscripts is more complicated because of the impact of the SNR of the recorded signal. SNR is critically important in cases where the subsequent image processing is aimed at discriminating between subtly different features (such as 'well-erased' text from the underlying parchment). Since the SNR is proportional to the square root of the number of counted photons, and since a smaller pixel will count fewer photons than a larger pixel under the same exposure, the SNR measurement will be smaller in proportion to the area of the pixel. This difference can result in significant detrimental impacts on the subsequent image processing, particularly when imaging spectral fluorescence with its inherently small number of generated and recorded photons is inherently small.

One other aspect of image 'sharpness' is often not considered with sufficient care in systems that image over the wavelength range wider than that of the sensitivity of the human eye. The system at St. Catherine's Monastery produces images over the range $365 \text{ nm} \leq \lambda \leq 940 \text{ nm}$, whereas camera lenses are normally designed to produce sharply-focused images in visible light only ($400 \text{ nm} \leq \lambda \leq 700 \text{ nm}$). This means that lenses designed to image only at visible wavelengths will not simultaneously focus light outside the visible range upon the sensor; the images at these wavelengths will be 'blurred', thus jeopardizing the success of subsequent image processing. Though it is possible to 'refocus' the lens at the longer wavelengths, the image magnifications will vary with focus setting, which adds complexity to both the collection process and to subsequent image processing.

One additional point of importance is the need to 'calibrate' the collected data to relate the measured numerical values to the physical properties of the object. This is most commonly done by comparing the 'raw' measured gray value of a pixel in each band to the corresponding value of a group of reference pixels with known properties of reflectance, fluorescence, and/or transmittance. Since images are collected beyond the visible range, the references themselves must be well calibrated over the full range of wavelengths as well, which most common 'colour charts' are not. For calibrating reflectance and fluorescence, standard materials made from Spectralon® are commonly used. This is a Teflon derivative with very uniform reflectance over wavelengths from 250 nm in the ultraviolet region to 2500 nm in the near-infrared spectrum—in other words, its spectral reflectance is very 'flat'. Spectralon® references are available from Labsphere, Inc. with different specified reflectance values ranging from 2% ('black') to 98% ('white'). It is recommended that both 'dark' and 'light' standards be included in the field of view; these will allow the images to be calibrated accurately and are essential for subsequent statistical processing. In our work, we use a custom Spectralon® calibration target with reflectances of 2%, 50%, and 98%, which allows accurate calibration of the reflectance.

In 2013, when the images used in the processing example were collected, the image set included 12 reflective bands over the range from the ultraviolet region at 365 nm through the visible to the infrared bands to 940 nm, eight transmissive bands, and eight spectral fluorescence bands (six under ultraviolet illumination and two under blue-light illumination). In addition, two 'raking-incidence' exposures for both $\lambda = 455 \text{ nm}$ (blue) and $\lambda = 870 \text{ nm}$ (infrared) with the illumination at a low 'raking' angle were included. These exposures occasionally reveal low-relief information, but were not used in the processing reported here.

IMAGE PROCESSING AND RENDERING

Once the multispectral images of a palimpsest have been collected, the goal is to combine them to enhance the visibility of the erased text. Based on experience in multispectral imaging and image processing of cultural heritage acquired over more than two decades, we have found that 'pseudocolour' images of selected image bands often result in text that is sufficiently enhanced for legibility. Dr. Keith T. Knox, of the Early Manuscripts Electronic Library, has developed and continually improved methods for creating pseudocolour renderings of multispectral image sets from the time of our initial work on the Archimedes Palimpsest in the first decade of

the new millennium (further details are available in his submission to this volume).³ A further development of this technique was recently applied to multispectral imagery from the *Zacynthius* palimpsest at Cambridge University Library.⁴ Though often very successful, there are cases where these renderings may not sufficiently enhance the contrast of erased text in some areas of damaged leaves or on entire leaves. In such cases, it often is helpful to apply processing algorithms based on the spectral ‘statistics’ of the collected bands within these local areas.

The spectral statistics are numerical descriptions of the ‘lightnesses’ in all bands for each pixel in the image field. We can visualize the concept by considering a system that collects images under M different wavelength bands in reflectance, where $M \geq 12$ for reflectance imaging with modern multispectral imaging systems. The numerical gray values of each pixel in a M -band multispectral image may be represented as a vector with M components, which places each pixel at a specific location in an M -dimensional space. Because our experience is constrained to three spatial dimensions, we cannot visualize pixel vectors from spectral images with $M > 3$, but such vectors may be represented, characterized, and manipulated in digital computers. We expect that the statistics of the vectors for an ensemble of pixels belonging to a specific ‘feature’ of the image (such as ‘erased undertext’) should exhibit spectral similarities and therefore the corresponding vectors for pixels belonging to a specific class should form a local ‘cluster’ in the M -dimensional space. Pixels from the other object classes should have different colours and lightness levels, and therefore should cluster around other distinct locations in the space of vectors. Of course, in bands where the SNR is low, the uncertainties in the measured signals will affect the locations of the endpoints of the vectors so that the corresponding clusters ‘spread out’ and/or the centroid of the clusters may be displaced. As a result, the clusters corresponding to different features on the leaf may overlap (e.g., erased text and nearby parchment), which makes the task of distinguishing among the image features more difficult. This is a further illustration that the SNR of the measured signal must be ‘sufficiently large’ for the image features to be distinguished correctly.

Different methods are available for analyzing the statistics of local regions and may be applied to ‘enhance’ or ‘tighten’ this clustering and therefore improve the ability to segment the pixels that belong to each of the different feature classes. These methods will be considered individually.

The relative utility of results obtained from the two general processing methods of pseudocolour renderings and statistical processing has resulted in a two-tiered workflow. All leaves are batch processed to form pseudocolour images, which are examined and transcribed by the scholars. In the course of the transcription, the scholars identify local areas (or even full leaves) that remain difficult to read, which are then prioritized for subsequent statistical processing using the methods to be described in this paper.

MULTISPECTRAL IMAGE PROCESSING METHODS APPLIED TO THE ST. CATHERINE’S PALIMPSESTS

The primary algorithms based on the spectral statistics that were used to process multispectral imagery from the collection of palimpsests at St. Catherine’s Monastery are: Principal Component Analysis (PCA), Minimum Noise Fraction (MNF), Independent Component Analysis (ICA), and Spectral Angle Mapping (SAM). Because the different methods use different assumptions about the spectral characteristics of the image data, it should not be surprising that the results obtained from the various methods will differ. For this reason, it often is useful to form pseudocolour renderings of outputs from the different methods, which may be manipulated subsequently by the imaging scientist or scholar to enhance the visibility of different features of interest. As one example, dynamic variation of the ‘hue angle’ of the pseudocolour image often reveals information to the eye that was previously imperceptible.

The calibrated image bands are collected to form an image ‘cube’ of data utilized by *ENVI*. Each cube is ‘calibrated’ to the numerical reflectance from the Spectralon® target in the scene for the reflectance and fluorescence bands; spectral imagery in transmittance are calibrated based on unobscured areas. This calibration evaluates the ratio of the measured 16-bit integer value at each pixel to the average signal reflected from the 98% broadband Spectralon® reflectance in that spectral band. The result is a 32-bit floating-point number for

³ Christens-Barry *et al.*, *Imaging and Image Processing Techniques*.

⁴ Houghton *et al.*, *The Palimpsest Catena of Codex Zacynthius*.

each band at each pixel, which means that the calibrated image cube requires twice the storage space of the original image data.

Our experience has shown that the processed results are often improved by an additional preprocessing step, to remove gross variations in the background of the spectral image. This step evaluates the ratio between each band and a replica that has been locally averaged, or 'blurred,' over a linear scale larger than the features of interest (e.g., larger than the stroke width of the pen used to write the overtext). This 'blur-and-divide' step attenuates the visibility of any large-scale 'patchiness' in the parchment background that is often due to staining or shading. The statistics used in the subsequent processing will then not be dominated by such gross spatial variations, which rarely have significance to the readings. The images after the 'blur-and-divide' preprocessing appear to have a 'flat' background, with the original large-scale variations much reduced.

For ease of comparison in this discussion, the four different processing methods were applied to one leaf: f. 103r from Sin. ar. 514. The images were collected in 2013, with a monochrome sensor producing images of size 6132 pixels by 8176 pixels at a spatial resolution of approximately 687 pixels per inch (27 pixels per millimetre). Figure 1 shows the visual appearance of the full image, including the Spectralon® reflectance references and a colour calibration chart from X-Rite, Incorporated. These were located at the bottom and left side of the image, respectively. Since the colour chart is calibrated only for visible light, it is used only to create the image of the true visible appearance – the reference values are not useful for bands in the ultraviolet and infrared regions of the spectrum. We should note that restricting the illustrated examples to this one leaf will not illustrate all methods in their best light, since the specific statistical properties of this leaf can be expected to respond to some methods better than others.

In the time frame of the project, the image set included twelve reflective bands from ultraviolet through infrared, four bands with raking illumination, seven bands of fluorescence under blue and ultraviolet illumination, and eight bands of transmission from blue through infrared. In nearly all cases, statistical processing is applied to subsets of these 31 bands, such as just to the reflectance bands. In all cases shown, the processing was applied only to the 7 fluorescence bands after 'blur-and-divide' preprocessing, where each spectral band was locally averaged over a square area with linear dimension of 101 pixels.

The image processing algorithms used here are available in the *ENVI* software package from L3Harris, Inc. *ENVI* was designed and is usually used for the processing of imagery for environmental remote sensing, but it has proven to be also very useful in studies of historical artefacts. Fortunately, both authors resided at the time at the Chester F. Carlson Center for Imaging Science of the Rochester Institute of Technology, which is an *ENVI Center of Excellence*, and were therefore able to access the software easily.

PRINCIPAL COMPONENT ANALYSIS (PCA)

This method for statistical analysis was invented by Pearson in 1901.⁵ PCA combines the images in a set of M spectral bands to create an 'equivalent' set of M bands that exhibits a specific correlation property. The two sets are 'equivalent' because the transformations in both directions are available from the calculation. In other words, the original set of M bands may be reconstructed from the M output principal components. The specific property of the processed bands is that M PCA bands are 'uncorrelated'—any statistical similarities among the images in the original set of bands will not exist in the set resulting from the processing. In addition, the M output bands are sequenced in order of statistical 'variance'; the band with the largest variance (loosely corresponding to the widest range in numerical values or widest contrast) appears first in the sequence of M output bands, while each succeeding processed band exhibits both no statistical correlation to and less variance than any previously-evaluated output band in the sequence. In this way, the last bands in the calculated sequence exhibit the smallest variance and therefore the smallest variation in image contrast.

The first step in the PCA calculation is to select the region where the spectral statistics are evaluated; these statistics determine the subsequent processing. This region might be the entire scene, but it is much more common to select a small region dominated by the feature of interest (e.g., erased text). Since all pixels of the same object feature tend to have the same 'colour', we expect the endpoints of vectors for pixels belonging to each class of object will 'cluster' in the M -dimensional spectrum space, so that pixels from distinct object

⁵ Pearson, *On Lines and Planes*.

features will appear at noticeably different locations in the M -dimensional space. The PCA calculation assumes that each ‘cluster’ may be described by a Gaussian statistical distribution (the well-known ‘bell curve’), which means in turn that only second-order statistics (variances) are significant, and that there are no asymmetries in these clusters, such as ‘skew’ or ‘kurtosis’.

The first step is to determine the axis in the M -dimensional data space that spans the widest variance of data for the selected region. The original data for the entire image are then projected onto this axis in the M -dimensional space, with one end of the data projection specified as ‘black’ and the other as ‘white’. The end result is a monochromatic image of the first principal component. Next, the axis orthogonal to the first such that the data subset spans the widest remaining variance is then determined, the M -dimensional data are projected, and the endpoints mapped to ‘black’ or ‘white’ to evaluate the second principal component. The process repeats for each of the M dimensions, resulting in a new representation of the original M bands of data sequenced in order of image ‘variance’. The calculation is fairly intensive and often requires considerable time to run.

Note that the calculation of the orthogonal axes depends upon the scaling of the image data. To obtain reproducible and comparable results, the original spectral image data must be accurately calibrated based on measured reflectance standards, lest the results be affected by nonintrinsic factors, such as different absolute exposures in different wavelength bands.

The fact that the axes in the calculated coordinate system are orthogonal means that the output image bands are ‘uncorrelated’. The *ideal* outcome of PCA would be a set of output bands such that each displays a single feature of the scene, which means that the desired feature (usually, the erased text) would appear in a single PC band and that all other features of the scene would have identical gray values that differ from the level of the feature of interest. However, this happy situation rarely occurs—the desired feature generally appears in multiple principal components. In these cases, it is often useful to render ‘pseudocolour’ images by inserting PC bands into the red, green, and blue channels of a visible colour image; this image may then be enhanced, either statically by the imaging scientist or dynamically in display software by the scholar to enhance the visibility of the desired text.

Often the highest-order PCA bands with the small variances are dominated by random variations in the image data (which is called ‘noise’), while the ‘true’ image structure is most visible in the lowest-order bands. Since the noise is now segregated from bands with ‘nonrandom’ or genuine structure, the higher-order bands with noise may be discarded. The inverse PCA calculation can then reconstruct the original image bands without the random variations. This means that PCA is useful for ‘denoising’ spectral image data. It also is useful for determining the effective number of image bands in the original data.

An example comparing the visual appearance of the leaf with a pseudocolour result from PCA evaluated for the eight fluorescence bands on the leaf Sin. ar. 514, f. 103r is shown in Figures 2(a) and 2(b), respectively.

MINIMUM NOISE FRACTION (MNF)

Similar to PCA, the minimum noise fraction transform also produces the set of M image bands into a new set of M bands that are sequenced according to a noise-content description of ‘image quality’. The end result is similar to that from PCA, but MNF often does a better and more reliable job of identifying and removing noise from the data than is possible with PCA alone.⁶ The goal of MNF is to identify component bands that contribute to the SNR of the ensemble of bands in the most significant manner. The covariance of the noise is estimated from the image data directly, and often produces a result that is better than that from PCA alone.

A pseudocolour result from MNF of the seven fluorescence bands of Sin. ar. 514, f. 103r is shown in Figure 2(c). The colour rendering was manipulated by subsequent processing by Adobe Photoshop® to maximize the visibility of the erased text.

⁶ Green *et al.*, A transformation for ordering multispectral data.

INDEPENDENT COMPONENT ANALYSIS (ICA)

Just as for PCA, the goal of ICA is separation of M spectral bands into individual signals from the constituent components, but ICA produces 'statistically independent' rather than 'statistically uncorrelated' signals.⁷ Though conceptually similar to PCA, ICA applies less stringent assumptions to the original data and uses higher-order statistical correlations to determine the transformation. This means that ICA is somewhat more capable of segmenting feature(s) of interest into different output bands.

A common illustration of the concept of ICA is the solution to the so-called 'cocktail party problem', where each of an ensemble of observers hears a mixture of sounds from a roomful of independent conversations, and the processing task is to isolate the individual conversations from signals recorded by different microphones scattered about the room. Each conversation is analogous to an individual feature on a palimpsest leaf (e.g., 'overtex', 'undertext', 'mould', 'parchment', etc.), and the measured signal from each individual microphone corresponds to a different spectral image band. The ICA calculation looks for differences in the measured signals that satisfy a condition for 'statistical independence'. If the different conversations are truly independent (as they would usually be in a cocktail party), ICA usually is able to segment the signals into their constituent components.

By analogy with the variances of the different PCA bands, ICA bands may be sequenced based on a calculated metric called 'spatial coherence'. This is evaluated by mathematically comparing, or 'correlating', translated replicas of each ICA image band to the undisplaced band. If the correlation between a replica displaced by a large physical distance across the scene remains high, then the image is assumed to have spatial structure corresponding to useful information and thus a large 'spatial coherence'. If the correlation between replicas for large displacements is small, then the image is assumed to consist primarily of 'random noise' that is not usefully interpreted, so the 'spatial coherence' is small. The set of ICA output bands may be sequenced in descending order of spatial coherence to form a set of ICA outputs that are analogous to the PCA sequence.

Again, as in PCA, the desired feature of erased text rarely appears in a single ICA band, so dynamically changing renderings of pseudocolour images may reveal or enhance subtle text information. However, in the example of ICA processing Sin. ar. 514, f. 103r from the seven fluorescence bands, ICA band #2 in Figure 2(d) rather successfully shows the erased text. The pseudocolour result from three of the seven ICA output bands in Figure 2(e) is only marginally different from the single band.

SPECTRAL ANGLE MAPPING

Recall the visualization of each pixel in a scene imaged in M bands as forming M -dimensional vectors. If we know that a pixel or small group belong to a feature of interest (such as erased undertext), then we can construct a M -dimensional 'reference' vector for this feature. Even in this M -dimensional space, the angle between any pair of vectors is one-dimensional and quite easily evaluated from the scalar product of the two vectors. In spectral angle mapping (SAM), the angle between the vector corresponding to each pixel in the scene and a 'reference' vector for a pixel known to belong to the feature of interest (e.g., erased text) is evaluated simply and very quickly from a simple trigonometric formula. The resulting gray-scale output image is rendered by converting the angle for each pixel to a proportional gray value—small angles are mapped to 'dark' output pixels and large angles to 'bright' pixels. The resulting gray-scale image depicts the spectral similarity of each pixel to that of the desired feature. It is important to note that the 'lightness' of the pixel has no effect on the output image—only the spectral angle determined from the M -dimensional vector is relevant. This last characteristic of SAM can be very useful for features such as erased text, which may be considerably 'lighter' than areas of dense ink but may still have a measurable spectrum.

An example of a SAM result from the eight fluorescence bands for Sin. ar. 514, f. 103r is shown in Figure 2(f). The result is very similar to that for ICA band 2 in Figure 2(d), but the process is much quicker to run.

⁷ Hyvarinen, Fast and robust fixed-point algorithms.

CONCLUSION

Image processing methods based on multiband statistics have proven to be useful for subtle characters in many of the image sets from palimpsests in the collection at St. Catherine's Monastery. The methods are still being refined and further improvements may be expected.

REFERENCES

- Bai, Di, David W. Messinger, and David Howell, Hyperspectral analysis of cultural heritage artifacts: pigment material diversity in the Gough Map of Britain, *Optical Engineering* 56(8) 081805 (2017).
- Christens-Barry, William, Roger L. Easton, Jr., and Keith T. Knox, Imaging and Image Processing Techniques, in: Reviel Netz, William Noel, Natalie Tchernetska, and Nigel Wilson (eds.), *The Archimedes Palimpsest*, vol. 1 (Cambridge, 2011) 175–197.
- Green, Andrew A., Mark Berman, Paul Switzer, and Michael D. Craig, A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Transactions on Geoscience and Remote Sensing* 26(1) (1988) 65–74.
- Houghton, Hugh A. G., Panagiotis Manafis, and Amy C. Myshrall, *The Palimpsest Catena of Codex Zacynthius: Text and Translation* (Piscataway, NJ, 2020).
- Hyvarinen, Aapo, Fast and robust fixed-point algorithms for independent component analysis, *IEEE Trans. Neural Networks* 10(3) (1999) 626–634.
- Pearson, Karl, On Lines and Planes of Closest Fit to Systems of Points in Space, *Philosophical Magazine* 2(11) (1901) 559–572.
- Peery, Tyler, Roger L. Easton, Jr., Rolando Raqueño, Michael Gartley, and David W. Messinger, Image Quality in Cultural Heritage, *Manuscript Cultures* 14 (2020) 99–104.



Fig. 1: Visual appearance of f. 103r of Sin. ar. 514, which was used as the test object for the processing. The small section near the bottom outlined in red is the region where the processing methods are compared. The Spectralon® reference standards with 2%, 50%, and 98% broadband reflectivity are at the bottom of the image field, and the colour chart that is used only for calibrating the visible colour image is at the lower left of the image field

(© St. Catherine's Monastery, Sinai, Egypt)

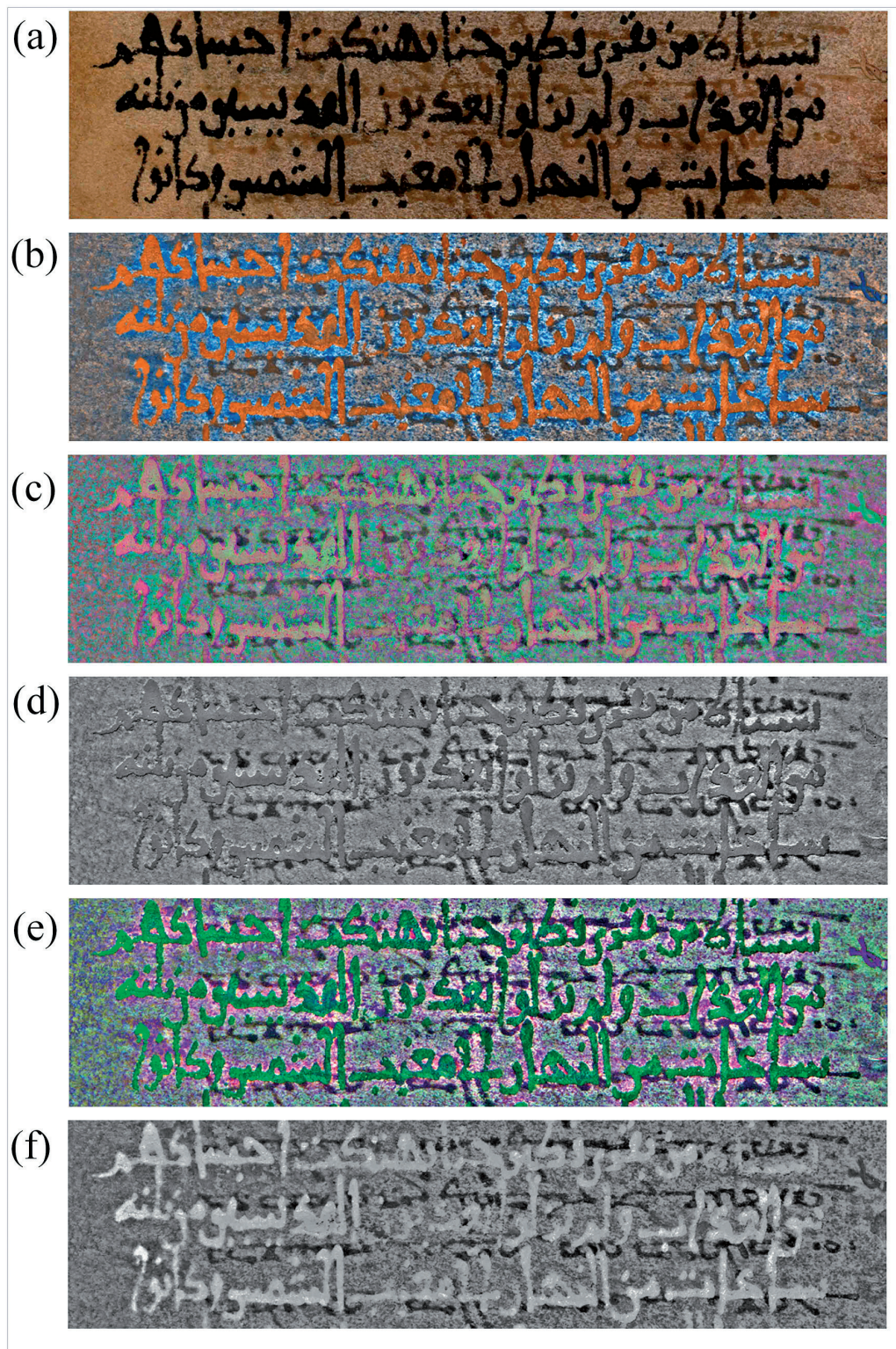


Fig. 2: Comparison of processing results of the small section of f. 103r of Sin. ar. 514. All processing is applied to the seven spectral fluorescence bands: (a) visual appearance, showing that some of the erased text is marginally legible; (b) pseudocolour PCA result, showing some interference from the coloured overtext; (c) pseudocolour MNF result, which shows the undertext in similar contrast to the PCA example, but the overtext has much less contrast; (d) ICA band 2, showing good contrast of the erased undertext; (e) pseudocolour ICA image, showing the colour contrast between the erased undertext and the overtext; (f) result of spectral angle mapping, where the gray value is 'dark' if the angle between the vectors for the image pixel and the reference undertext pixel is small.

Note the significant contrast between the dark undertext and overtext and the 'crisp' edges of the undertext

(© St. Catherine's Monastery, Sinai, Egypt)