

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

## Pigment diversity estimation for hyperspectral images of the Selden map of China

Di Bai, David W. Messinger, David Howell

Di Bai, David W. Messinger, David Howell, "Pigment diversity estimation for hyperspectral images of the Selden map of China," Proc. SPIE 10644, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV, 1064415 (8 May 2018); doi: 10.1117/12.2304041

**SPIE.**

Event: SPIE Defense + Security, 2018, Orlando, Florida, United States

# Pigment Diversity Estimation for Hyperspectral Images of the Selden Map of China

Di Bai<sup>a</sup>, David W. Messinger<sup>a</sup>, and David Howell<sup>b</sup>

<sup>a</sup> Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, 14623, USA, E-mail: db3641@rit.edu, messinger@cis.rit.edu

<sup>b</sup> Bodleian Library, University of Oxford, Broad St, Oxford OX1 3BG, UK, E-mail: david.howell@bodleian.ox.ac.uk

## ABSTRACT

The Selden Map of China, an early 17th century wall map of East Asia was rediscovered in 2008. Recently, the map was imaged using a VNIR hyperspectral imaging system while in the collection at the Bodleian Library, Oxford University. The goal of the collection of the hyperspectral image (HSI) of the Selden Map was to help historians understand the material diversity of its composition and potentially the methods used in the creation of the map. The Selden map has been named one of the "Treasures of the Bodleian" and it poses many questions such as the diversity of pigments used to create the map. In this research, we extract visually common pixels (here, the green pigments) from the Selden Map and estimate the material diversity of the green pixels. Previous pigment analysis on the HSI of the Gough Map<sup>1,2</sup> used an endmember based approach, the Gram Matrix technique,<sup>3,4</sup> to understand the number of distinct materials in a scene and then used spectral angle mapper (SAM) to classify all the pigments. Here, we use the same Gram Matrix technique, but due to the complexity of the Selden Map data, instead of using SAM, we use two spectral unmixing techniques, NNLS (nonnegative linear least squares)<sup>5</sup> and FUMI (functions of multiple instances)<sup>6</sup> to determine the weights of all the endmembers for each data point to study the within-material diversity. Results show that the Selden Map is composed of at least 6 kinds of dominant green pigments with a particular spatial pattern. This research provides a useful tool for historical geographers and cartographic historians to analyze the material diversity of HSI of cultural heritage artifacts.

**Keywords:** Hyperspectral imagery, pigment analysis, dimensionality estimation, spectral unmixing

## 1. INTRODUCTION

The Selden Map of China, dating from c.1619, arrived at the University of Oxford's Bodleian Library in 1659.<sup>7-9</sup> Rediscovered by Robert Batchelor in 2008, this map provided an opportunity to reexamine the history of Chinese cartography and dispute about maritime dimensions of the Ming Empire.<sup>7</sup> The map was drawn with carbon ink and watercolours on Chinese paper, depicting China (Ming dynasty), the East and South China Sea from Japan to Timor, and a network of shipping routes from Indonesia to the western coast of India.<sup>8,9</sup> There are mountains, rivers, ocean waves, forests, plants and flowers drawn using six different colors: red, green, blue, yellow, white and black.<sup>9</sup> The color green dominates this map due to the large areas of the sea in the map.<sup>10</sup> Questions were posed such as pigment analysis for estimation of material diversity in the Selden Map. To answer such questions, hyperspectral images (HSI) of the map were collected at the Bodleian Library, Oxford University in 2015. The collection was done in 12 overlapping "chips" to ensure the full coverage of the entire map with sufficient spatial resolution. Each HSI chip has 334 spectral bands (colors) over the Vis-NIR spectrum ranging from 400nm to 1000nm. The size of each chip varies from 1600 (pixels) x 2300 (pixels) to 1600 (pixels) x 3250 (pixels). Section 3.1 states more collection details. This research is significant because the pigment analysis of the HSI of the Selden Map contributes to the codicological studies of the map (*i.e.*, understanding the tools, techniques, and timeline of its creation and revision).

---

Further author information: (Send correspondence to Di Bai)

Di Bai: E-mail: db3641@rit.edu, Telephone: 1 585 754 3879

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXIV,  
edited by Miguel Velez-Reyes, David W. Messinger, Proc. of SPIE Vol. 10644, 1064415  
© 2018 SPIE · CCC code: 0277-786X/18/\$18 · doi: 10.1117/12.2304041

In cultural heritage codicological studies, faded text enhancement and pigment analysis are always important to historical geographers and cartographic historians. For faded text analysis, popular methods such as multispectral reflectance, fluorescence, and transmission imaging<sup>11–19</sup> were used to enhance faded text for reading. For example, Easton and his colleagues used a Megavision Imaging System to collect and generate multispectral images of manuscripts and then used multiple image processing techniques to enhance the texts. Methods included unsupervised/supervised classification, independent component analysis (ICA), principal component analysis (PCA), renderings in pseudo-color and spectral pseudoinverse calculation to enhance multispectral images of historical manuscripts, such as the Herculaneum Papyri and Archimedes Palimpsest.<sup>11–19</sup> For hyperspectral image analysis, Goltz<sup>20</sup> applied hyperspectral imaging to assess stains on historical documents. He generated optical density slices (which is the minus log function of the image pixel value divided by the white reference) of certain ranges of spectra and created false color images to enhance the faded text.

For pigment analysis, in recent years, hyperspectral Imaging has been used in the cultural heritage artifacts such as material identification and mapping of the work of art.<sup>21–26</sup> Using HSI systems with different classification algorithms allows: (a) restoration and pigment differentiation of the historical paintings,<sup>26–29</sup> (b) study of the methods and materials used in the artifact creation,<sup>30</sup> (c) analysis of an artwork's historic value,<sup>31–33</sup> (d) identification of timeline of a manuscript creation<sup>31</sup> and (e) recovery of erased and overwritten scripts.<sup>32</sup> Different imaging methods like X-ray fluorescence imaging,<sup>21,26</sup> Infrared imaging<sup>34,35</sup> and laser induced breakdown spectroscopy<sup>31</sup> were used to obtain the HSIs. Image processing techniques included PCA,<sup>24,29,30</sup> spectral angle mapper,<sup>28</sup> spectral unmixing<sup>27</sup> and support vector machine.<sup>36</sup> Dooley *et al.*<sup>24</sup> and Delaney *et al.*<sup>25</sup> also used ENVI N-d visualizer to find endmember spectra through clustering and then use spectral angle mapper to classify pigments of hyperspectral images. In 2017, Bai *et al.*<sup>1</sup> presented an approach by using Gram Matrix and end-member analysis to estimate the green pigment diversity over the Gough Map and their results showed that there were at least 6 kinds of dominant green pigments used in the map. That was the first time estimating within material diversity in cultural heritage artifacts. In this paper, we extend Bai's research by using two spectral unmixing techniques to classify the green pigments in the Selden Map. We also propose a new spectral angle spatial pattern method to validate the classification results.

This paper is organized as follows. Section 2 describes 1) the MaxD and Gram Matrix techniques used to estimate material diversity in the HSI of the Selden Map and 2) the NNLS and FUMI spectral unmixing techniques. Section 3 describes application using this technique to estimate the material diversity of green pigments in the Selden Map and presents the results. Section 4 summarizes the results and findings.

## 2. MATERIAL DIVERSITY ESTIMATION & SPECTRAL UNMIXING

This section provides the theoretical basis of understanding the number of distinct materials (green pigments) in the Selden Map and how to use spectral unmixing to output a spatial distribution of the green pigments. Material diversity estimation outputs the representatives (endmembers) of different colors (materials) within a hyperspectral image. Then, spectral unmixing techniques calculate the abundances of those representatives and classify the hyperspectral image.

### 2.1 Material Diversity Estimation

Material diversity estimation of a data set can be related to the dimensionality of the data in the spectral domain. The dimensionality of a data set is the estimation of the least number of free variables required to represent the data without losing information.<sup>37</sup> There have been many approaches used in dimensionality estimation of hyperspectral dataset,<sup>3,37–43</sup> but the most common one is PCA,<sup>39,40</sup> which transforms the original data into a new space through the magnitude of its eigenvalue, but equates the dimensionality of the data with the significance of a dimension in the new space.

In this research, we use an approach developed by Messenger, *et al.* (2011)<sup>3</sup> and applied by Canham, *et al.* (2011)<sup>38</sup> to estimate the diversity of green river and ocean pigments in the HSI of the Selden Map. The Gram Matrix is first used to estimate the dimensionality of hyperspectra data and then the MaxD algorithm is used

to extract associated endmembers. Gram Matrix is an  $k \times k$  matrix, where  $k$  is the number of vectors in the set. The  $(m, n)^{th}$  element of the Gram matrix is the inner product of the  $m$  and  $n$  vectors in the set, such that

$$G_{m,n} = \langle x_m, x_n \rangle. \quad (1)$$

Note that the determinant of  $G$  (known as the Gramian) is the square of the volume of the parallelepiped that encloses the vector set. An important property of the Gram matrix is that if the Gram Matrix is composed of endmembers extracted from the image, when the Gramian approaches zero, the vectors in the set are no longer linearly independent, which indicates the number of distinct materials in the hyperspectral image. After acquiring the number of endmembers, MaxD algorithm is used to extract endmembers, which represent the number of "pure" materials in the HSI image. More details of this approach can be found in the literature.<sup>3,38,41</sup>

## 2.2 Spectral Unmixing

After acquiring the endmembers of the HSI, spectral unmixing is used to identify the abundances of those endmembers in each pixel.<sup>44</sup> NNLS (nonnegative linear least squares)<sup>5</sup> and FUMI (functions of multiple instances)<sup>6</sup> were used in this research. Lawson and Hanson gave the standard algorithm for NNLS in 1995.<sup>5</sup> Mathworks modified the algorithm and renamed it to *lsqnonneg*.<sup>45</sup> The NNLS algorithm solves a nonnegative least squares curve fitting problem by minimizing the objective function:  $\min_x \|C \cdot x - d\|_2^2$ , where  $x \geq 0$ . In this function,  $d$  represents the reflectance vector of each pixel,  $C$  is the endmember matrix and  $x$  is the abundances of those endmembers. More details of this approach can be found in the literature.<sup>46</sup>

The functions of multiple instances (FUMI) technique<sup>6,47</sup> is a generalization of multiple instance learning (MIL). In HSI spectral unmixing, FUMI treats each data point as a function of endmembers. The function parameters define the relationship between each data point and the endmembers. The output will be an abundance matrix of the proportion of the endmembers for each data point. The objective function to minimize is shown below:

$$F = \frac{1-\mu}{2} \sum_{i=1}^N w_{l(x_i)} \left\| (x_i - l(x_i)p_{iT}e_T - \sum_{k=1}^M p_{ik}e_k) \right\|_2^2 + \frac{u}{2} \sum_{k=1}^M \| (e_k - \mu_0) \|_2^2 \\ + \frac{u}{2} \| (e_T - \mu_0) \|_2^2 + \sum_{k=1}^M \frac{\Gamma}{\sum_{n=1}^N p_{nk}^{t-1}} \sum_{i=1}^N p_{ik}$$

where,  $e_T$  is the target signature,  $e_k$  is the nontarget signature,  $p_i$  is the proportion value of the  $k^{th}$  signature in data point,  $\mu_0$  is the global data mean,  $l(x_i)$  is the binary label determining target or nontarget,  $w_{l(x_i)}$  depicts the weight of target and nontarget samples. More details of this approach are in the literature.<sup>47</sup> The algorithm's code is accessible at GitHub.<sup>48</sup>

In this paper, we first use Photoshop to create a green pigment mask of the Selden Map. After that, the green pixels from the Selden Map (including rivers and oceans) are extracted and their dimensionality is estimated by the Gram Matrix technique. Spectral unmixing is then applied to classify all the green pigments into different classes. We also generate spectral angle spatial patterns (the spectral angle between every pixel and its belonging class's average spectra) within each class to validate the classification results. The goal of this research is to estimate the number of distinct green pigments utilized in the creation of the entire Selden Map and to find whether there is a spatial pattern in the distribution of each pigment type across the map. The Gram matrix technique enables us to estimate within-class dimensionality of the data and spectral unmixing technique separates out potentially different green pigments based on their spectra, despite their common visually "green" appearance.

## 3. APPLICATION TO THE SELDEN MAP

In this paper, our goal is to estimate the within material diversity of the green pigments in the Selden Map through analysis of the hyperspectral imagery. In this section, we introduce the application of the above-described analysis techniques to the green pigments.



### 3.1 Collection Parameters

The RGB image of the Selden Map is shown in Fig. 1(a) and the map with landmarks' names is in Fig. 1(b). The actual size of the Selden Map is 100cm x 160cm. To ensure full coverage of the map with sufficient spatial resolution, the Selden Map was imaged in 12 spatially overlapping chips and labeled as chip 1 to chip 12 as shown in Fig. 1(c). Note that chip 2 was lost during its collection.

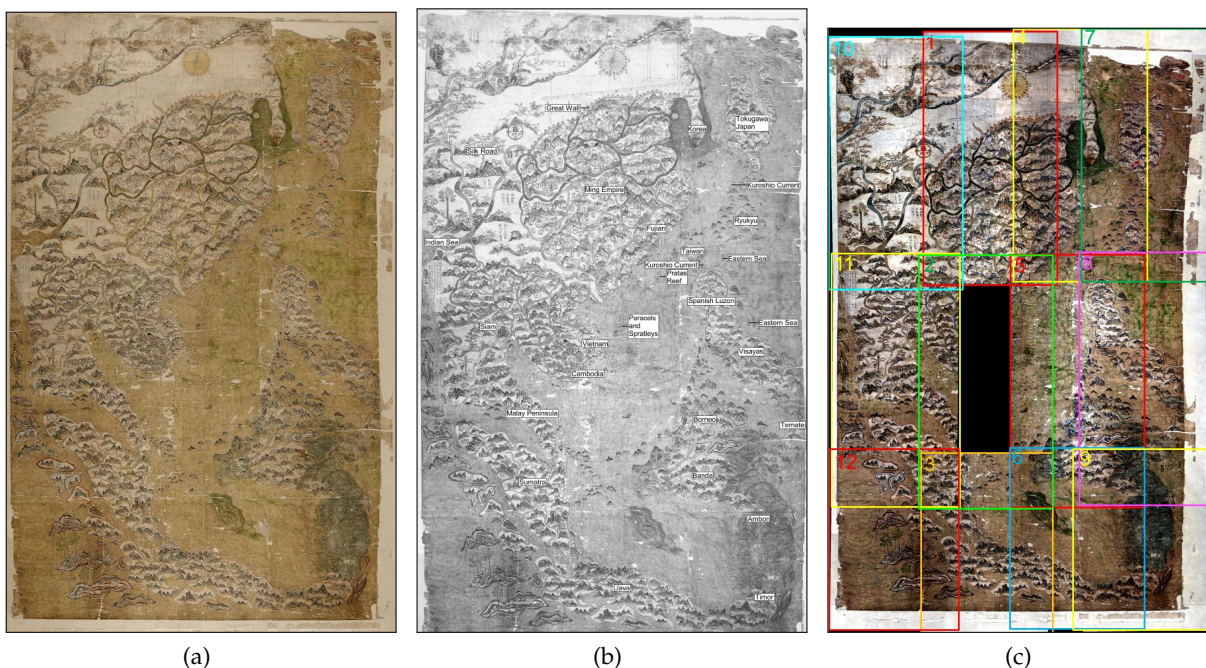


Figure 1. High resolution images of the Selden Map. (a) The RGB image of the map showing particular topographical features, including mountains, rivers, oceans, islands, straits, provincial boundaries, cities and coastal routes in different colors and patterns. The green (yellowish green) pigment dominates this map. (b) Labelled Selden Map with major landmarks. (c) RGB renderings of the HSI chips representing the collection of the entire Selden Map. Note that HSI data of Chip 2 was lost during the collection process.

Each HSI chip has 334 bands with the size varying from 1600 (pixels) x 2300 (pixels) to 1600 (pixels) x 3250 (pixels). The entire Selden Map is about 4784 (pixels) x 7657 (pixels) x 334 (bands). Given the actual size of the Selden Map: 100cm x 160cm, the spatial resolution is about  $\Delta x \simeq \Delta y \simeq 48$  pixels/cm. In addition, each chip has 334 bands over the visible and near infrared (Vis-NIR) range from 398.7 nm to 1000.3 nm, so that  $\Delta\lambda = 1.8$  nm.

### 3.2 Methodology

The diversity of the green pigments in the Selden Map is estimated by using the Gram Matrix technique as described above. Two spectral unmixing techniques: FUMI and NNLS are used to classify the Selden Map. The methodology is outlined here:

1. Use Photoshop to create a green pigments mask of the Selden Map.
2. Extract green pixels from each HSI chip of the Selden Map.
3. Hyperspectral data preprocessing.
  - (a) Sphere the data of each chip by unit magnitude normalizing each pixel to eliminate brightness differences in the reflectance.

- (b) Bin the data of each HSI chip of the Selden Map (spectrally) from 334 bands to 111 bands, so as to increase the SNR.
4. Use the MaxD and Gram Matrix techniques to estimate the dimensionality for each chip and compute endmembers.
  - (a) extract  $k$  endmembers ( $k \approx 10$ )
  - (b) iterate  $i$  from 3 to  $k$  endmembers
    - i. represent the  $j$  endmembers in the Gram matrix
    - ii. compute the convex hull volume for  $j$  endmembers
    - iii. save that volume as a function of  $j$
  - (c) the estimated dimensionality of the HSI chip is the number of endmembers where the convex hull volume function approaches zero.
5. Compute global endmembers of the green pigments across entire Selden Map.
6. Classify green pixels per chip.
  - (a) For each chip, use NNLS (or FUMI) spectral unmixing method to compute the abundances of each global endmember.
  - (b) Use K-means classification method to classify the abundances matrix into certain classes.
  - (c) Label each pixel based on the classification results of the K-means.
  - (d) Map the classification results back to the image for visual interpretation.
7. Compute average spectral within each class and compute spectral angle spatial patterns of each class to validate the classification accuracy.

Details are provided in following subsections on the steps outlined above.

### 3.2.1 Create mask and extract green pixels

To ensure a full and accurate coverage of all the green pigments, we create a mask of each chip using Photoshop. An example of extracting the green pigments from chip 4 of the Selden Map is shown in Fig. 2. Fig. 2(a) is the original RGB image of chip 4 and Fig. 2(b) is the mask of all the green pigments. Then we simply add a threshold to convert (b) to a binary image shown in Fig. 2(c), so that the black pixels in (c) represent all the green pixels in the original chip, automatically extracted. Finally, the data of green pixels in each chip are exported for endmember analysis and meanwhile, the coordinates of the green pixels are also saved for future mapping back into the image space.

### 3.2.2 Hyperspectral data preprocessing

The next step is to preprocess the hyperspectral data, which includes sphering the data and binning the data. We first sphere the data of the green pigments within each chip of the Selden Map. In each collection, a white reference image for each chip was also obtained and then, by normalizing out this white reference, the estimated reflectance of each pixel in the Selden Map was acquired. However, to ensure all of the chips have the same brightness levels, we sphere the data (*i.e.*, unit normalize each pixel vector) to make each pixel have a vector magnitude of 1. This results in a normalized reflectance for each pixel.

Then, we spectrally bin the data of each HSI chip of the Selden Map to increase the SNR. Previous research<sup>1</sup> showed that the pixels' spectral signatures are generally slowly varying, especially in the spectral range of 400nm – 500nm and 800nm – 1000nm. In order to boost the signal to noise ratio (SNR), we spectrally bin each chip by three. Compared to the original HSI chip with 334 bands ( $\Delta\lambda = 1.8$  nm), now we have 111 bands ( $\Delta\lambda = 5.4$  nm).

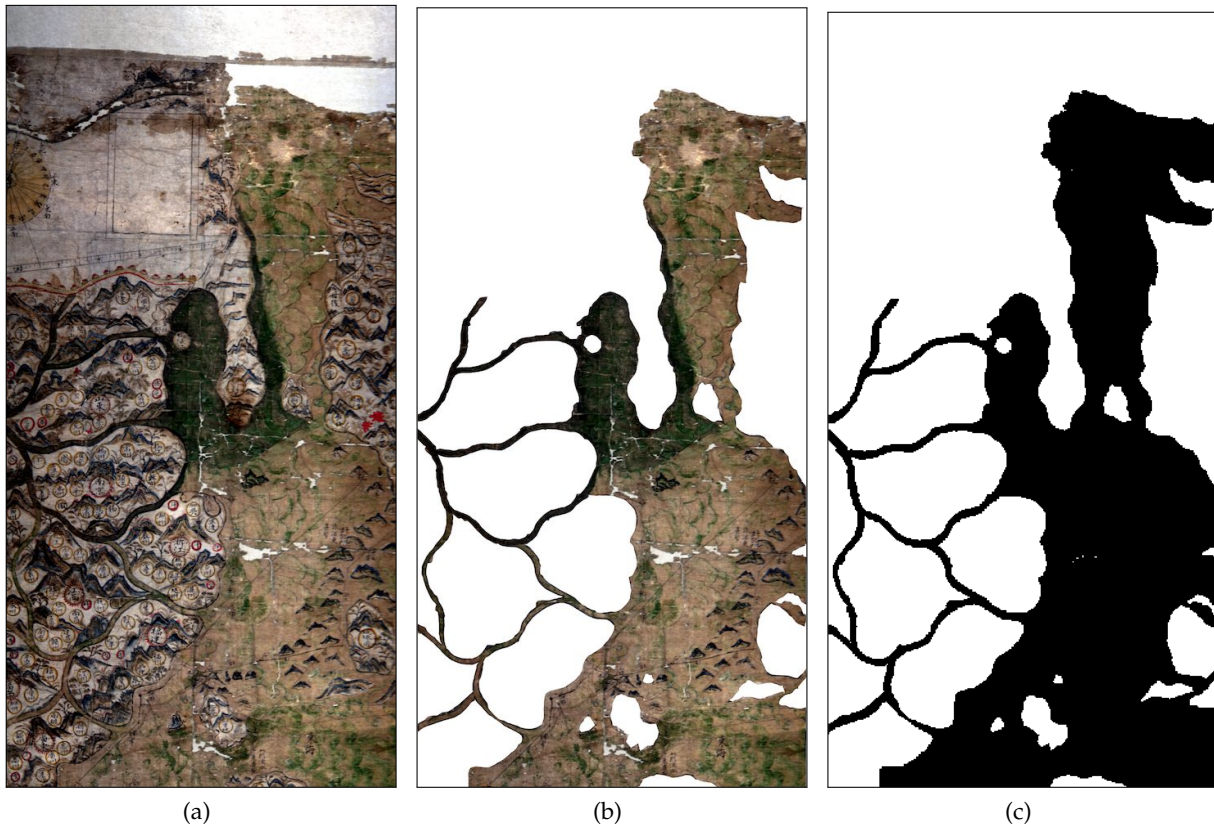


Figure 2. Example process of extracting green pixels in one chip of the Selden Map. (a) Original chip 4 from the Selden Map; (b) Green region mask specified in Photoshop; (c) Binary image of (b) with black pixels representing green pixels of the original chip 4.

### 3.2.3 Extract endmembers and estimate material diversity

The third step uses the MaxD and Gram Matrix techniques to extract endmembers and estimate the number of distinct green pigments used in each HSI chip. Specifically, because the green pigments in chip 10 are very few and the data of chip 2 was lost during its collection, we only analyze the green pigments in the remaining 10 chips. In Fig. 3(a), as described in the Messinger, *et al* (2011),<sup>3</sup> here we plot the estimated convex hull volume vs the number of endmembers (dimensionality) of an example (chip 1) of the Selden Map and when the dimensionality at which the volume function approaches zero, the endmembers indicate the number of distinct materials in the scene. The reflectance of the six endmembers of chip 1 are plotted in Fig. 3(b). We apply the same method to the remaining chips and find that all of them show that there are six distinct green pigments within each chip. The similar results are expected since common materials and methodology were likely used across the entire Selden Map. Given the dimensionality of the green pixels of each chip is six, we output six endmembers for each chip. Those endmembers are actual spectra of green pixels in each corresponding chip.

### 3.2.4 Compute global endmembers

In order to present a distribution pattern of the green pigments across the entire Selden Map, we need to compute global endmembers. First, we use the MaxD and Gram Matrix approach to estimate the number of distinct green pigments from the 60 primary endmembers depicted in 3.2.3. Similarly, the Gram matrix function shows that there are six distinct green pigments used to represent the green pigments over the entire Selden Map.

Then, we use K-means classification to classify the spectra of the 60 endmembers into six clusters. Results are shown in Fig. 4 (a-f). Notice that these six classes are distinct from each other. Fig. 4(g) shows the averaged



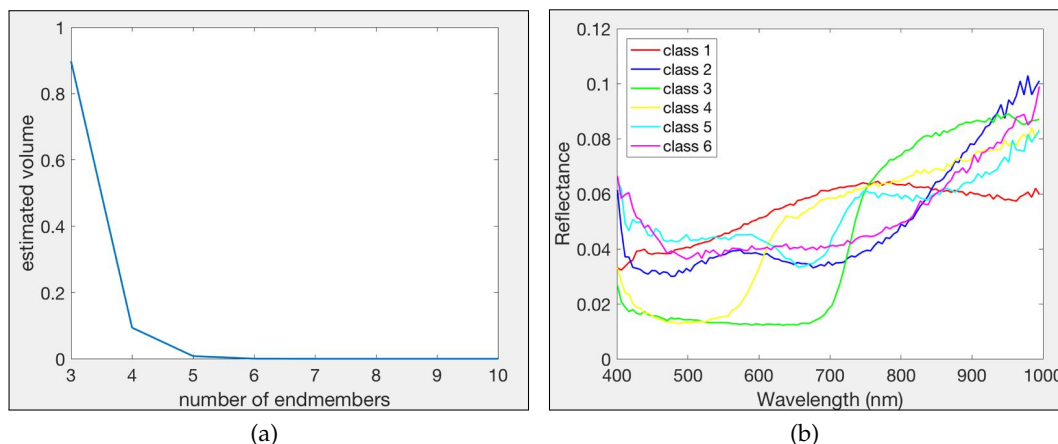


Figure 3. Material diversity plot for chip 1 of the Selden Map in (a), with x axis: number of endmembers and y axis: estimated convex hull volume. Endmember plot for chip 1 in (b), with x axis as wavelength in nm and y axis as reflectance.

spectra within each class in (a-f). These six distinct spectra represents the global distinct green pigments in the Selden Map.

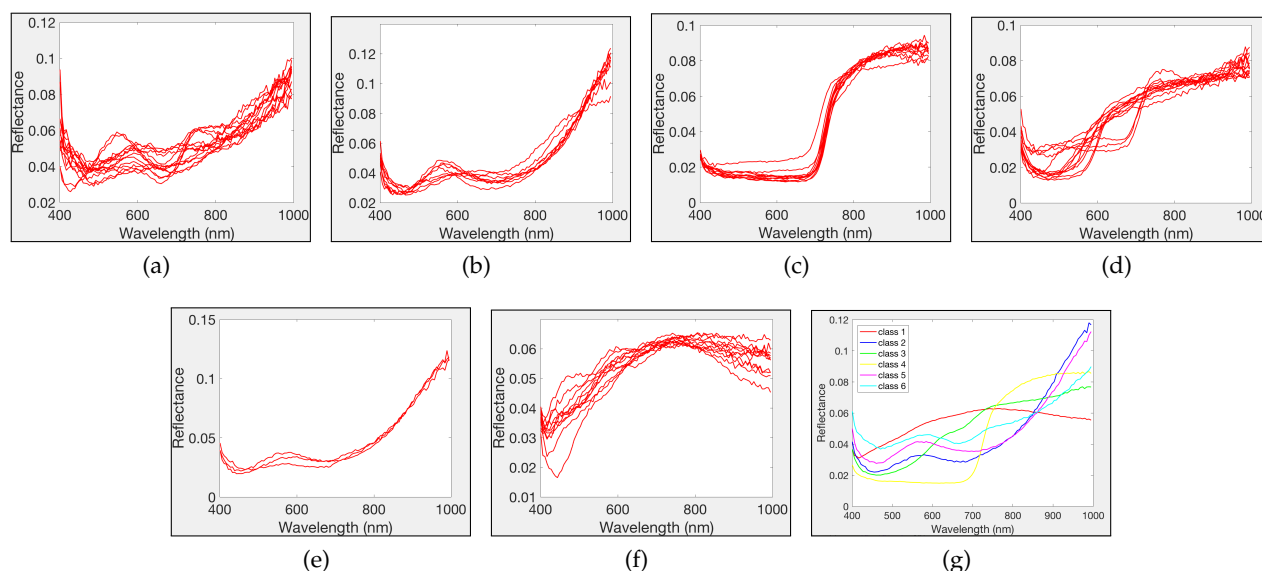


Figure 4. K-means classification results of 60 primary endmembers of the green pixels in (a-f) and averaged spectra of above six classes (g) representing six global endmembers in the map.

### 3.2.5 Classify the green pixels per chip

Now we use spectral unmixing techniques to classify the green pixels based on the six global endmembers per chip. Using either FUMI or NNLS, the output is an abundance matrix, whose size is 6 (rows)  $\times$  p (columns), where p is the number of pixels. Each column gives the proportion of the six endmembers in one pixel. Here, instead of simply choosing the maximum proportion within each column as the class of the pixel, we use K-means to classify the abundance matrix into six classes. We also expand the abundance matrix into a 21 (rows)  $\times$  p (columns) matrix which also includes the differences between each row of the abundance matrix and then use K-means to classify the new "abundance+difference" matrix. Compared to the original 6 (rows)  $\times$  p (columns)



matrix, the new matrix contains more descriptors for each pixel, which enables to identify subtle differences between the pixels, for example, the waves in the ocean. The output of the K-means classification result is the label of green pigment. After classification, we map the results back to the image for visual interpretation. The final results, after stitching all the chips, are shown in Fig. 5 (a-d). Fig. 5 (e) and (f) are for comparison.

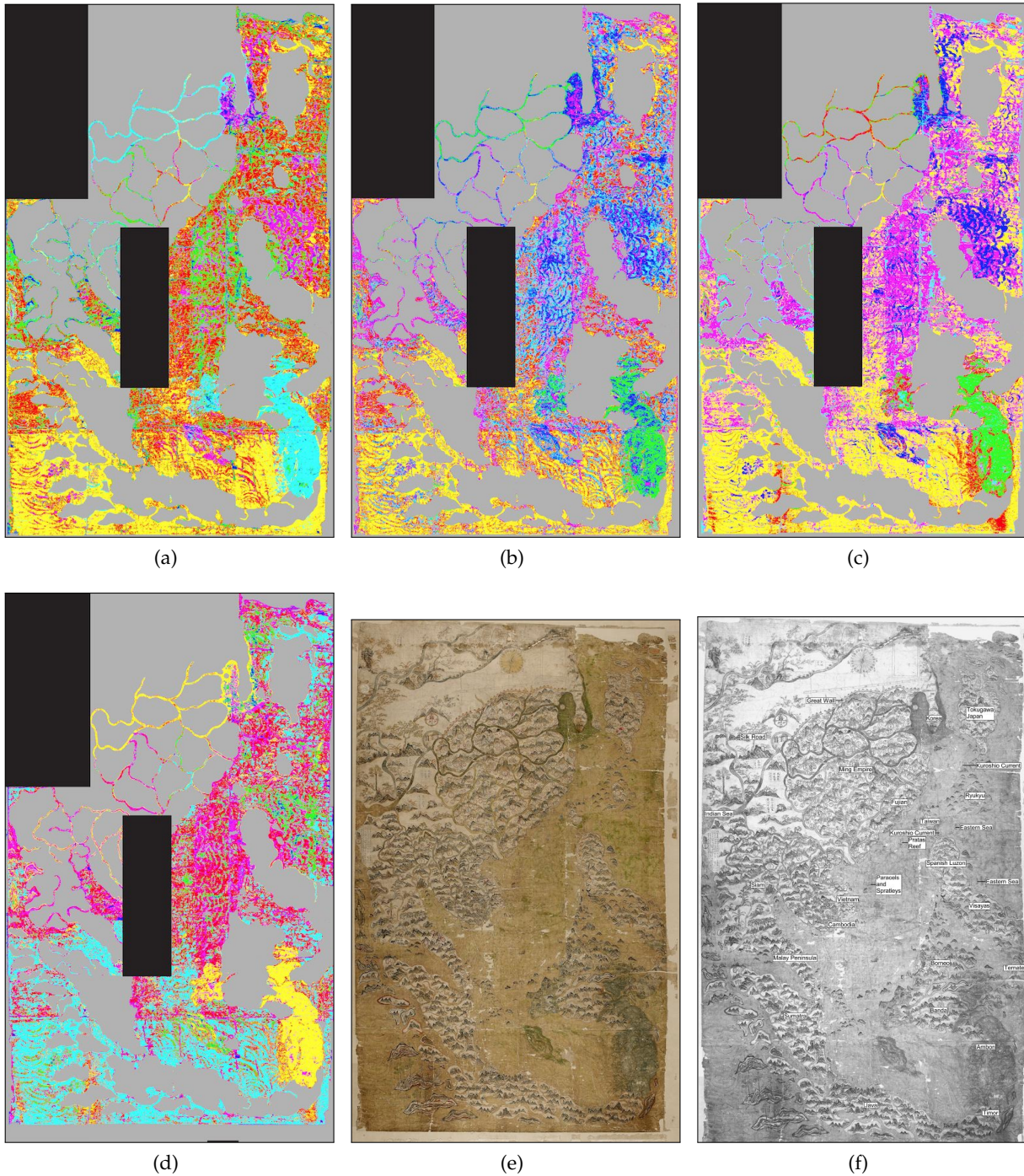
From Fig. 5, notice that all the classification results in (a-d) show a clear separation between the north and south of the rivers inside the Ming Empire. Specifically, the green pigments of the rivers in the north side is the same as the pigments used in Ambon and west of the Banda. The boundary pigments of Korea are the same as the rivers in the north of Ming Empire, but they are different from those inside Korea. The pigments inside Korea is similar to the green pigments in the green land between Java and Banda. The green waves of the Eastern Sea can be easily noticed and there is a clear separation between the waves and their background. Also notice that there is a folding line in the south of this map, where even though the pigments in either side of the folding line have some similarities, there is a clear faded line shown in all classification results.

The Selden Map was frequently displayed for its unusual characters and by the early 20th century, it had deteriorated and had been repaired and mounted as a wall hanging.<sup>49</sup> By the 1970s, the map was significant damaged according to the conservation records.<sup>49</sup> It is impossible for us to know exactly how many green pigments were used in this map and we do not have ground truth for it.

As far as we know, previous study of the Selden Map suggested that:

1. The ocean in the map was flooded in an uneven greenish wash. The waves are those patterns with cloud-like billows. It might have been blue before the copper pigment oxidised.<sup>50</sup>
2. Large areas of the map (the ocean) were painted with green. Most of it was in a brownish green color but some regions were drawn with brighter green pigments.<sup>10</sup>
3. The ocean of the map is colour-washed in a yellowish green with conventional and faint waves. Around Korea in the north, to the south and west of Borneo, darker green pigments were used. In the Borneo area, the new sea color blots out what was formerly a region of land.<sup>7</sup>
4. In the southeastern corner of the map, the darker green might be a correction that removes a former connection between Timor and Ambon.<sup>7</sup>
5. In 2016, three green pigments were examined by a X-ray fluorescence (XRF) Spectrometer including 1) indigo, 2) copper green (a basic copper chloride and possibly malachite) and 3) orpiment.<sup>10</sup>

The classification results enriched previous work by providing a spatial classification pattern of all the green pigments and these results give historians more insight on the creation, revision of the map and even the transformation of the continent.



### 3.2.6 Spectral angle spatial patterns

Due to the absence of ground truth data, in this research, we use four different ways to classify the HSI of the Selden Map. Notice that the four classification results are very close, but not exactly the same, which is expected. However, it is desirable to validate the classification results as best as possible. Here, we develop a spectral angle spatial pattern method. After classification, we first compute the average spectrum within each class and take that spectrum as the representative of this class. Then we replace each pixel's value by its spectral angle to this class's representative spectra and map the results back to the image. In this way, the spectral angle spatial patterns of each class convey how close the pigments are within each class to the mean of that class.

Figs. 6, 7, 8 & 9 shows the spatial patterns of the green pigments in six classes using four different methods: 1) NNLS + abundances, 2) NNLS + abundances + differences, 3) FUMI + abundances, 4) FUMI + abundances + differences. The color bar depicts the spectral angle ranging from 0 to 10 degrees. Notice that almost all the patterns are mainly composed of blue and cyan, which indicate that most of the spectral angles are around 2-3 degrees indicating the class label is generally good and there is not much variability in the class. Fig. 9(b) has some red points with large spectral angle, but there are very few points in this class compared to the other classes and we will ignore that. Since spectral angle is globally small in each class, we conclude that our classification results are acceptable.

## 4. SUMMARY

This research provides a novel application for pigment analysis of a medieval map using hyperspectral imagery. The Selden Map of China was likely created in 17th century and the rediscovery of the Selden Map of China in the Bodleian Library in 2008 provides an opportunity to reassess the history of Chinese cartography. The Selden Map was imaged using a hyperspectral imaging system, which provides both spatial information and spectral information of this map for the first time. The goal of this research is to estimate the material diversity of the green pigments of the Selden Map and propose a new spectral angle spatial pattern method to validate the classification results. The Gram matrix and MaxD techniques are used to estimate the dimensionality and compute the endmembers of each HSI chip. After using spectral unmixing methods to classify pigments based on six global endmembers, we map them back into the original image so that global patterns are identified. Four different techniques based on NNLS and FUMI convey similar classification results, which enrich previous study on the green pigment analysis of the Selden Map. To validate the classification results, we generate spectral angle spatial patterns within each class. Since the dominant spectral angle are smaller than 3 degrees, we can confidently say that our classification results are acceptable.

The conclusion of this research is that there are six distinct green pigments present in the Selden Map. In addition, there is a clear separation inside the rivers of the Ming Empire. The green pigments of the rivers in the north side is the same as the pigments used in Ambon and west of the Banda. The green pigments inside Korea is similar to the land between Java and Banda. The green waves in the Eastern Sea can be easily noticed and there is a folding line in the south of this map. Apparently, there is some spatial relationship to the distribution of the green pigment classes that could be related to the timeline of the construction and correction of the Selden Map. This methodology can be generalized to a novel pigment analysis tool for historical geographers and cartographic historians to analyze the material diversity of cultural heritage artifacts by using hyperspectral imaging and consequently, contribute to the codicological study of the cultural heritage.

## ACKNOWLEDGMENTS

The authors wish to acknowledge Dr. Robert Batchelor for his insightful comments into the history, significance, and questions about the Selden map. We also wish to acknowledge Mr. Nick Millea at the Bodleian Library, Oxford University for his assistance with this project.



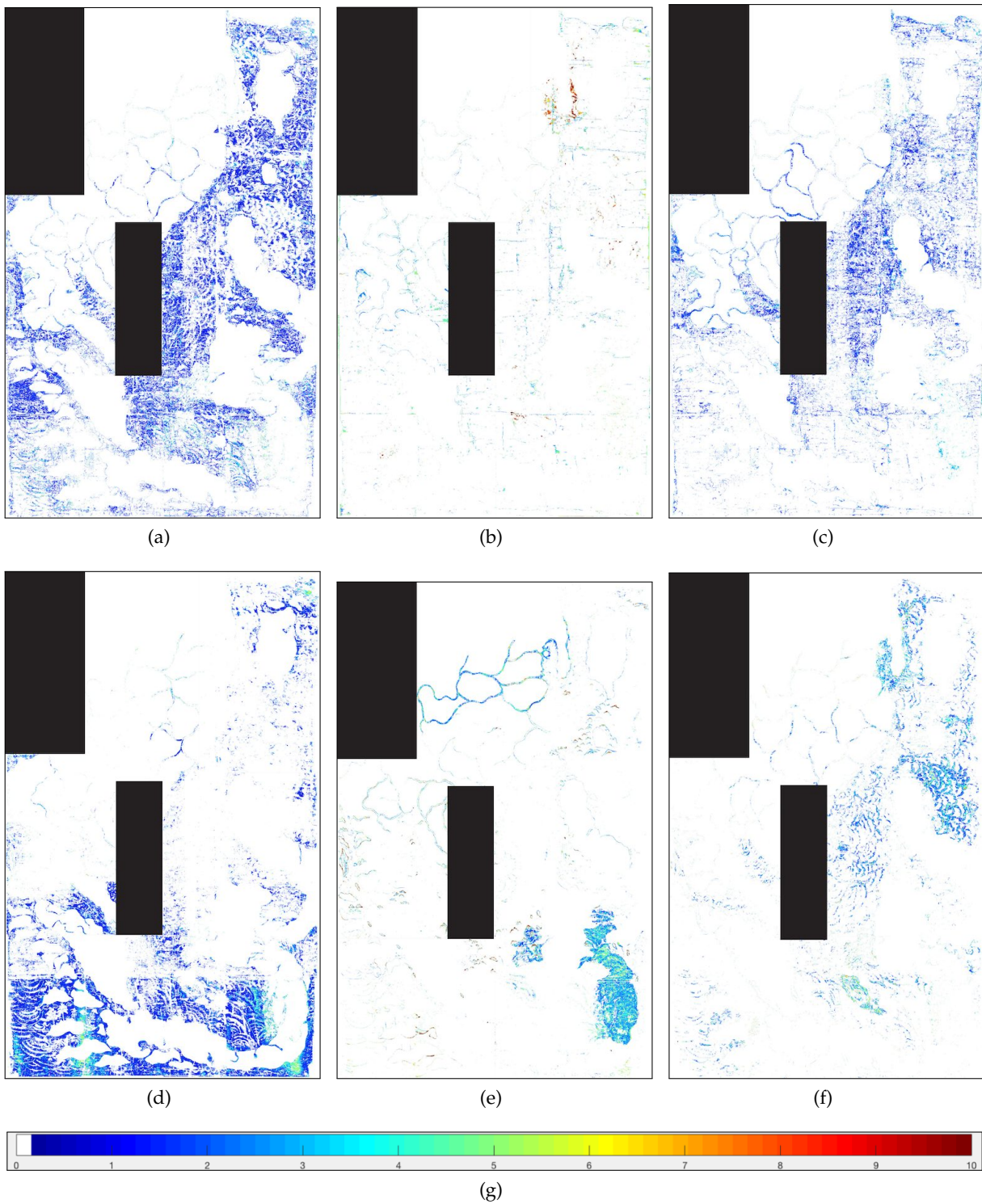


Figure 6. Spatial pattern of the spectral angles using NNLS+abundances within each class. (a) to (e) represent class 1 to 6 of the green pigment classification. (g) is the color bar ranging from 0 to 10.



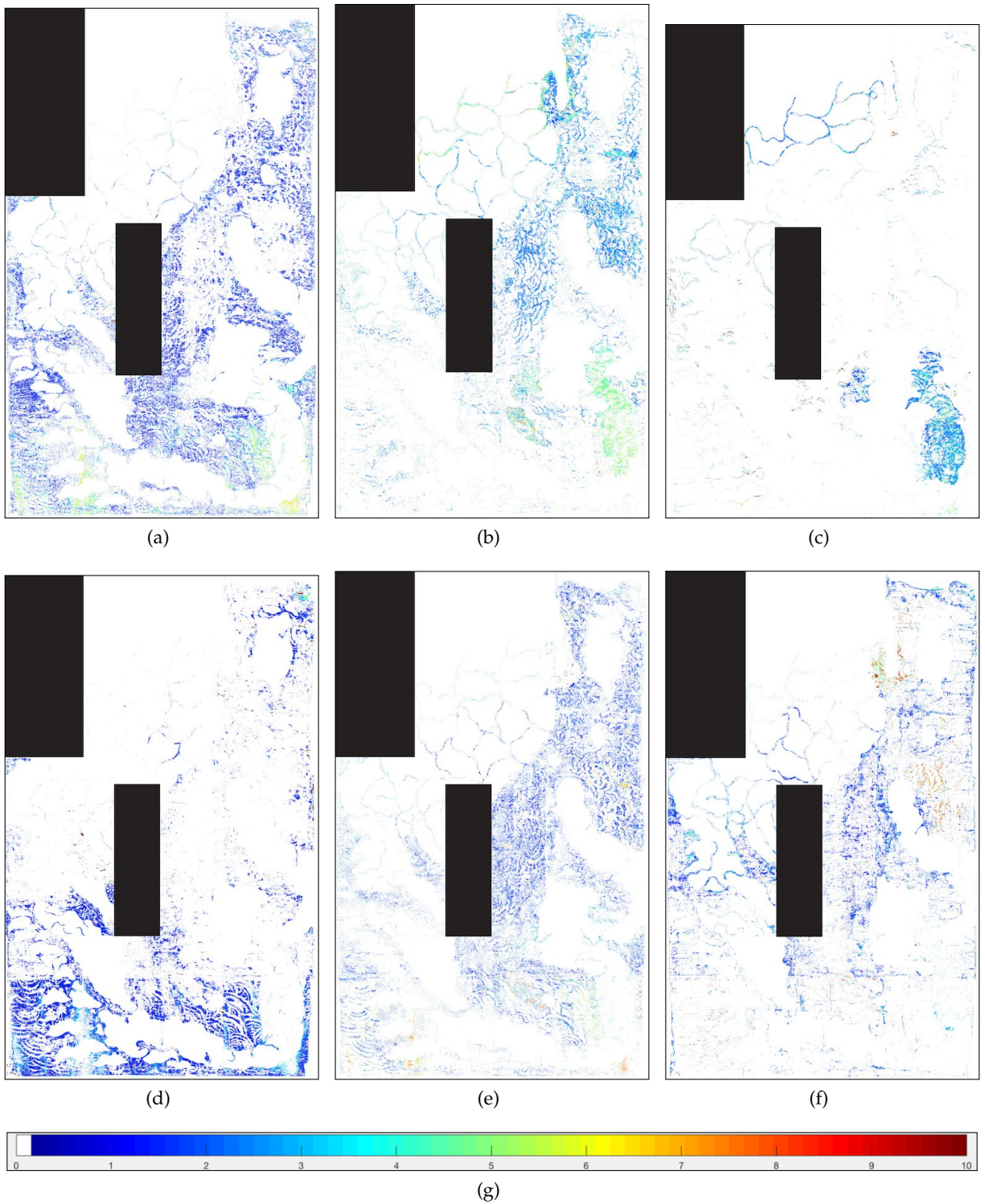


Figure 7. Spatial pattern of the spectral angles using NNLS+abundances+differences within each class. (a) to (e) represent class 1 to 6 of the green pigment classification. (g) is the color bar ranging from 0 to 10.

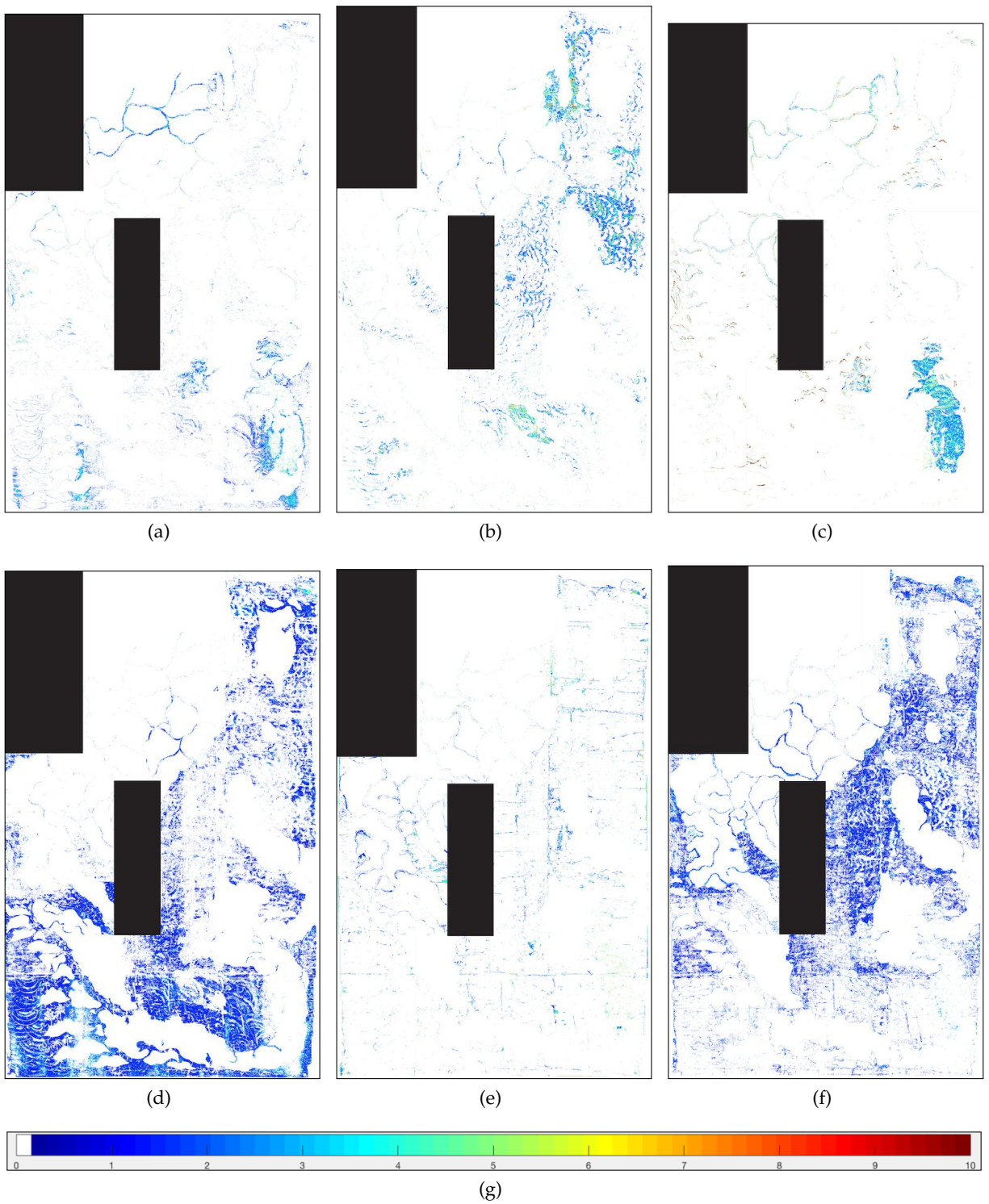


Figure 8. Spatial pattern of the spectral angles using FUMI+abundances within each class. (a) to (e) represent class 1 to 6 of the green pigment classification. (g) is the color bar ranging from 0 to 10.



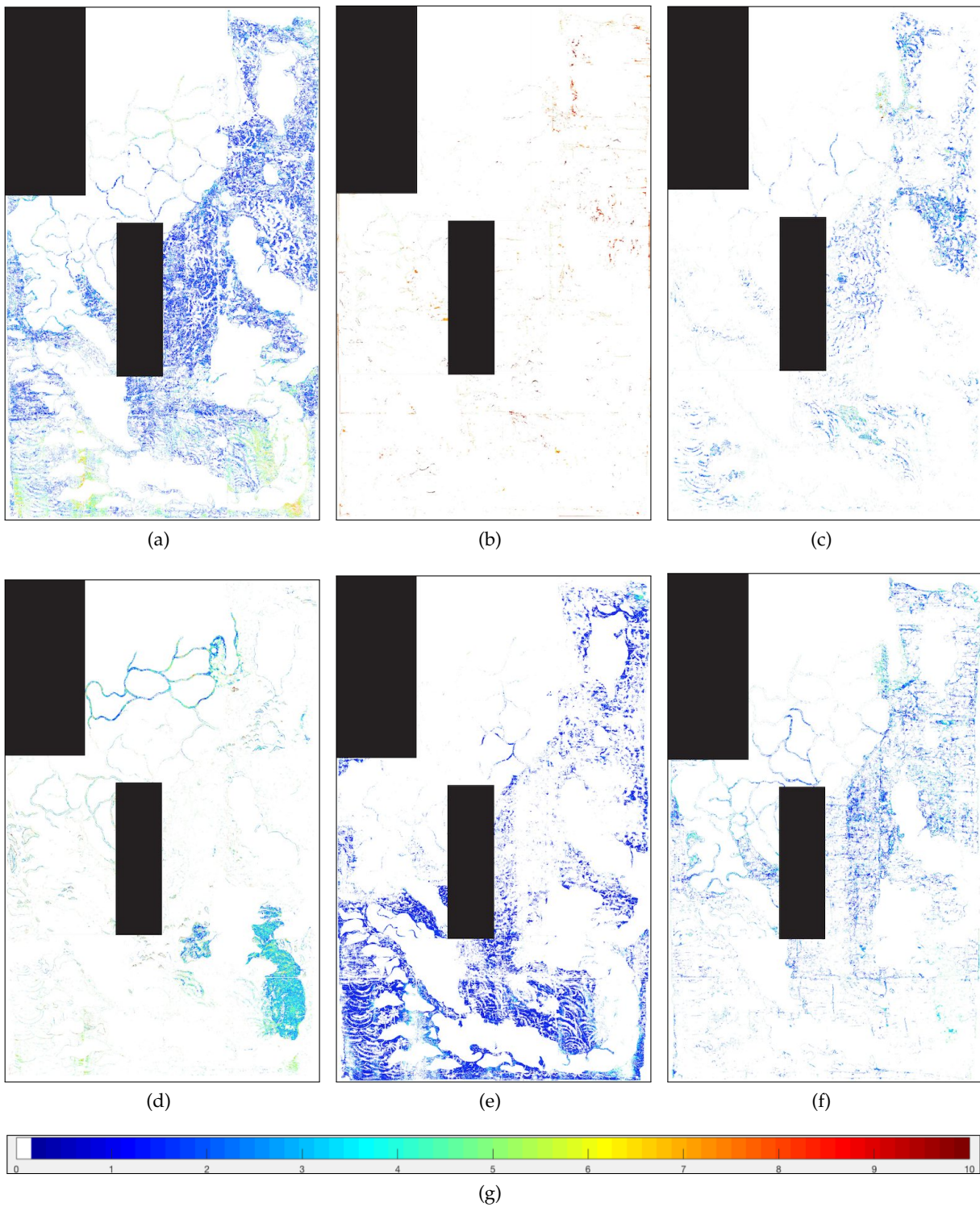


Figure 9. Spatial pattern of the spectral angles using FUMI+abundances+differences within each class. (a) to (e) represent class 1 to 6 of the green pigment classification. (g) is the color bar ranging from 0 to 10.

## REFERENCES

- [1] Bai, D., Messinger, D. W., and Howell, D., "A pigment analysis tool for hyperspectral images of cultural heritage artifacts," in [*SPIE Defense+ Security*], 101981A–101981A, International Society for Optics and Photonics (2017).
- [2] Bai, D., Messinger, D. W., and Howell, D., "Hyperspectral analysis of cultural heritage artifacts: pigment material diversity in the gough map of britain," *Optical Engineering* **56**(8), 081805–081805 (2017).
- [3] Messinger, D. W., Ziemann, A., Basener, B., and Schlamm, A., "Metrics of spectral image complexity with application to large area search," *Optical Engineering* **51**(3), 036201–1–036201–9 (2012).
- [4] Schlamm, A., Messinger, D., and Basener, B., "Geometric estimation of the inherent dimensionality of single and multi-material clusters in hyperspectral imagery," *Journal of Applied Remote Sensing* **3** (2009).
- [5] Lawson, C. L. and Hanson, R. J., [*Solving least squares problems*], vol. 15, Siam (1995).
- [6] Zare, A. and Gader, P., "Pattern recognition using functions of multiple instances," in [*Pattern Recognition (ICPR), 2010 20th International Conference on*], 1092–1095, IEEE (2010).
- [7] Batchelor, R., "The selden map rediscovered: a chinese map of east asian shipping routes, c. 1619," *Imago Mundi* **65**(1), 37–63 (2013).
- [8] Batchelor, R. K., [*London: The Selden Map and the Making of a Global City, 1549-1689*], University of Chicago Press (2014).
- [9] Nie, H., "The selden map of china: a new understanding of the ming dynasty," (2014).
- [10] Kogou, S., Neate, S., Coveney, C., Miles, A., Boocock, D., Burgio, L., Cheung, C. S., and Liang, H., "The origins of the selden map of china: scientific analysis of the painting materials and techniques using a holistic approach," *Heritage Science* **4**(1), 28 (2016).
- [11] Easton, R. L., Barry, W. A. C., and Knox, K. T., "Ten years of lessons from imaging of the archimedes palimpsest," *Commentationes Humanarum Litterarum*. **129**, 5–34 (2011).
- [12] Easton, R. L. and Noel, W., "Infinite possibilities: Ten years of study of the archimedes palimpsest," *Proceedings of the American Philosophical Society* **154**, 50–76 (2010).
- [13] Barry, W. A. C., Boydston, K., and Easton, R. L., "Some properties of textual heritage materials of importance in spectral imaging projects," *Commentationes Humanarum Litterarum*. **129**, 35–50 (2011).
- [14] Barry, W. A. C., Boydston, K., and Easton, R. L., "Some properties of textual heritage materials of importance in spectral 27 imaging projects," *Proceedings of Eikonopoiia, Eikonopoiia, Digital Imaging of Ancient Textual Heritage: Technological Challenges and Solutions, Eikonopoiia* **1**, 27–38 (2010).
- [15] Easton, R. L., Knox, K. T., and Barry, W. A. C., "Some properties of textual heritage materials of importance in spectral 27 imaging projects," *Proceedings of the European Signal and Image Processing Conference, Eurasp*, 1440–1444 (2011).
- [16] Knox, K. T., Easton, R. L., Barry, W. A. C., and Boydston, K., "Recovery of handwritten text from the diaries and papers of david livingstone," *Proceedings of the SPIE, Electronic Imaging* **7689** (2011).
- [17] Easton, R. L. and Kelbe, D., "Statistical processing of spectral imagery to recover writings from erased or damaged manuscripts," *Manuscript Cultures*. **7**, 35–46 (2014).
- [18] Easton, R. L. and Noel, W., "Multispectral imaging of the archimedes palimpsest," *Gazette du Livre Medieval*. **45**, 39–49 (2004).
- [19] Easton, R. L., Sacca, K., Heyworth, G., Boydston, K., Duzer, C. V., and Phelps, M., "Rediscovering text in the yale martellus map, spectral imaging and the new cartography," *7th IEEE International Workshop on Information Forensics and Security, 7th IEEE International Workshop on Information Forensics and Security* (2015).
- [20] Goltz, D., Attas, M., Young, G., Cloutis, E., and Bedynski, M., "Assessing stains on historical documents using hyperspectral imaging," *Journal of cultural heritage* **11**(1), 19–26 (2010).
- [21] Legrand, S., Vanmeert, F., Van der Snickt, G., Alfeld, M., De Nolf, W., Dik, J., and Janssens, K., "Examination of historical paintings by state-of-the-art hyperspectral imaging methods: from scanning infra-red spectroscopy to computed x-ray laminography," *Heritage Science* **2**(1), 13 (2014).
- [22] Fischer, C. and Kakoulli, I., "Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications," *Studies in Conservation* **51**(sup1), 3–16 (2006).



- [23] Rosi, F., Miliani, C., Braun, R., Harig, R., Sali, D., Brunetti, B. G., and Sgamellotti, A., "Noninvasive analysis of paintings by mid-infrared hyperspectral imaging," *Angewandte Chemie International Edition* **52**(20), 5258–5261 (2013).
- [24] Dooley, K. A., Lomax, S., Zeibel, J. G., Miliani, C., Ricciardi, P., Hoenigswald, A., Loew, M., and Delaney, J. K., "Mapping of egg yolk and animal skin glue paint binders in early renaissance paintings using near infrared reflectance imaging spectroscopy," *Analyst* **138**(17), 4838–4848 (2013).
- [25] Delaney, J. K., Zeibel, J. G., Thoury, M., Littleton, R., Palmer, M., Morales, K. M., Rie, E. R. d. l., and Hoenigswald, A., "Visible and infrared imaging spectroscopy of picasso's harlequin musician: mapping and identification of artist materials in situ," *Applied spectroscopy* **64**(6), 584–594 (2010).
- [26] Delaney, J. K., Thoury, M., Zeibel, J. G., Ricciardi, P., Morales, K. M., and Dooley, K. A., "Visible and infrared imaging spectroscopy of paintings and improved reflectography," *Heritage Science* **4**(1), 6 (2016).
- [27] Liang, H., "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," *Applied Physics A* **106**(2), 309–323 (2012).
- [28] Daniel, F., Mounier, A., Pérez-Arategui, J., Pardos, C., Prieto-Taboada, N., de Vallejuelo, S. F.-O., and Castro, K., "Hyperspectral imaging applied to the analysis of goya paintings in the museum of zaragoza (spain)," *Microchemical Journal* **126**, 113–120 (2016).
- [29] Baronti, S., Casini, A., Lotti, F., and Porcinai, S., "Multispectral imaging system for the mapping of pigments in works of art by use of principal-component analysis," *Applied optics* **37**(8), 1299–1309 (1998).
- [30] Attas, M., Cloutis, E., Collins, C., Goltz, D., Majzels, C., Mansfield, J. R., and Mantsch, H. H., "Near-infrared spectroscopic imaging in art conservation: investigation of drawing constituents," *Journal of Cultural Heritage* **4**(2), 127–136 (2003).
- [31] Melessanaki, K., Papadakis, V., Balas, C., and Anglos, D., "Laser induced breakdown spectroscopy and hyper-spectral imaging analysis of pigments on an illuminated manuscript," *Spectrochimica Acta Part B: Atomic Spectroscopy* **56**(12), 2337–2346 (2001).
- [32] Balas, C., Papadakis, V., Papadakis, N., Papadakis, A., Vazgiouraki, E., and Themelis, G., "A novel hyper-spectral imaging apparatus for the non-destructive analysis of objects of artistic and historic value," *Journal of Cultural Heritage* **4**, 330–337 (2003).
- [33] Klein, M. E., Aalderink, B. J., Padoan, R., De Bruin, G., and Steemers, T. A., "Quantitative hyperspectral reflectance imaging," *Sensors* **8**(9), 5576–5618 (2008).
- [34] Casini, A., Lotti, F., Picollo, M., Stefani, L., and Buzzegoli, E., "Image spectroscopy mapping technique for noninvasive analysis of paintings," *Studies in conservation* **44**(1), 39–48 (1999).
- [35] Cucci, C., Delaney, J. K., and Picollo, M., "Reflectance hyperspectral imaging for investigation of works of art: old master paintings and illuminated manuscripts," *Accounts of chemical research* **49**(10), 2070–2079 (2016).
- [36] Polak, A., Kelman, T., Murray, P., Marshall, S., Stothard, D. J., Eastaugh, N., and Eastaugh, F., "Hyperspectral imaging combined with data classification techniques as an aid for artwork authentication," *Journal of Cultural Heritage* (2017).
- [37] Camastra, F., "Data dimensionality estimation methods: A survey," *Pattern Recognition* **36**, 2945–2954 (2003).
- [38] Canham, K., Schlamm, A., Ziemann, A., Basener, B., and Messinger, D. W., "Spatially adaptive hyper-spectral endmember selection and spectral unmixing," *IEE Trans. on Geoscience and Remote Sensing* **49**(11) (2011).
- [39] Kirby, M., [Geometric Data Analysis: An Empirical Approach to Dimensionality Reduction and the Study of Patterns], John Wiley & Sons, Inc., New York, NY, USA (2000).
- [40] Jolliffe, I., [Principal Component Analysis.], Springer-Verlag New York, New York (2002).
- [41] Ziemann, A. K., Messinger, D. W., and Basener, W. F., "Iterative convex hull volume estimation in hyper-spectral imagery for change detection," *Proc. SPIE* **7695**, 76951I–76951I–9 (2010).
- [42] Yao, F. and Qian, Y., "Band selection based gaussian processes for hyperspectral remote sensing images classification," *2009 16th IEEE International Conference on Image Processing (ICIP)* , 2845–2848 (Nov 2009).
- [43] Pettis, K. W., Bailey, T. A., Jain, A. K., and Dubes, R. C., "An intrinsic dimensionality estimator from near-neighbor information," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-1**, 25–37 (Jan 1979).

- [44] Heylen, R., Parente, M., and Gader, P., "A review of nonlinear hyperspectral unmixing methods," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**(6), 1844–1868 (2014).
- [45] Shure, L., "Brief history of nonnegative least squares in matlab," *Blog available at: <http://blogs.mathworks.com/loren>* (2006).
- [46] Chen, D. and Plemmons, R. J., "Nonnegativity constraints in numerical analysis," in [*The birth of numerical analysis*], 109–139, World Scientific (2010).
- [47] Jiao, C. and Zare, A., "Functions of multiple instances for learning target signatures," *IEEE Transactions on Geoscience and Remote Sensing* **53**(8), 4670–4686 (2015).
- [48] Jiao, C. and Zare, A., "Functions of multiple instances for learning target signatures." <https://github.com/GatorSense/FUMI> (2015).
- [49] Minte, R., Stiglitz, M., Sugiyama, K., and Barnard, M., "From quanzhou, china to oxford, uk: an account of the selden map of china and its conservation," *Studies in Conservation* **59**(sup1), S115–S118 (2014).
- [50] Brook, T., [*Mr. Selden's Map of China: Decoding the Secrets of a Vanished Cartographer*], Bloomsbury Publishing USA (2013).
- [51] Kruse, F. A., Lefkoff, A. B., Boardman, J. B., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., and Goetz, A. F. H., "The spectral image processing system (sips) - interactive visualization and analysis of imaging spectrometer data," *Remote Sensing of Environment*. **44**, 145–163 (1993).