

Hyperspectral Imaging for Ink Mismatch Detection

Zohaib Khan, Faisal Shafait and Ajmal Mian

School of Computer Science and Software Engineering

The University of Western Australia

35 Stirling Highway, CRAWLEY, 6009

Email: zohaib@csse.uwa.edu.au, faisal.shafait@uwa.edu.au, ajmal.mian@uwa.edu.au

Abstract—Ink mismatch detection provides important clues to forensic document examiners by identifying whether a particular handwritten note was written with a specific pen, or to show that some part (e.g. signature) of a note is written with a different ink as compared to the rest of the note. In this paper, we show that a hyperspectral image (HSI) of handwritten notes can discriminate between inks that are visually similar in appearance. For this purpose, we develop the first ever hyperspectral image database¹ of handwritten notes in various blue and black inks, comprising a total of 70 hyperspectral images each in 33 bands of the visible spectrum. In an unsupervised clustering scheme, the spectral responses of inks fall into separate clusters to allow segmentation of two different inks in a questioned document. The same method fails to segment inks correctly when applied to RGB scans of these documents, since the inks are very hard to distinguish in the visible spectral range. HSI overcomes the shortcomings of RGB and allows better discrimination between inks. We further evaluate which subset of bands from HSI is most useful for the purpose of ink mismatch detection. We hope that these findings will stimulate the use of HSI in document analysis research, especially for questioned document examination.

I. INTRODUCTION

Natural materials absorb, transmit and reflect incident light in a particular manner. This property of a material which gives its due color, is often termed as its *spectral response*. Human eye is sensitive to light in the visible range and is able to well distinguish between different material colors [1]. However, humans are metameric to certain colors, i.e. they are unable to distinguish between two very close colors however, they are different in terms of their spectral information.

Analysis of inks is of critical importance to questioned document examination. The outcome of ink analysis potentially leads to determination of forgery, fraud, backdating and ink age. One of the most important tasks is to discriminate between different inks. There are two main approaches to distinguish inks, destructive and non-destructive examination. Chemical analysis such as Thin Layer Chromatography (TLC) [2] has been used to separate constituents of a mixture of inks. The separation of inks is achieved via capillary action. There are a few drawbacks to this approach. First, the examination is destructive, time consuming and sensitive to temperature. Second, the test is essentially qualitative and various measures need to be taken to quantify the results.

Instead, a non-destructive examination technique such as one proposed with spectral imaging has more potential. Hyperspectral imaging has recently emerged as an efficient non-destructive tool for detection, enhancement [3], comparison

and identification of forensic traces [4]. Such systems are critical for forensic document examiners in differentiating writing inks. However, the task is laborious and time consuming because the examiner needs to manually observe the document under each wavelength of light and make decisions based on qualitative analysis. The number of different wavelengths to observe could range from a few to hundreds, depending on the spectral resolution of the imaging system. Hence, automatic ink analysis can play a key role in supporting efficient questioned document examination.

Brauns et al. [5] developed a hyperspectral imaging system to detect forgery in potentially fraudulent documents in a non-destructive manner. Their imaging system was based on an interferometer which relies on moving parts for frequency tuning and therefore slows the acquisition process. Their work serves as a proof of concept for the identification of writing inks. Qualitative results on a small sample dataset showed that the ink spectra can be separated into different classes in an unsupervised manner. On the contrary, our system is based on an electronically tunable filter which is fast, accurate and has no moving parts. Moreover, we quantitatively analyze the problem of ink mismatch detection.

A more sophisticated hyperspectral imaging system was developed at the National Archives of Netherlands for the analysis of historical documents in archives and libraries [6]. The system provided high spatial and spectral resolution from near-UV through visible to near IR range. The only limitation of the system was its extremely slow acquisition time (about 15 minutes) [7]. Contrary to that, our proposed system captures hyperspectral images in only a fraction of that time.

In this paper, we present an efficient hyperspectral imaging system for writing inks mismatch analysis. Our work is based on the assumption that same inks exhibit similar spectral responses whereas different inks show dissimilarity in their spectra. We assume that the spectral responses of inks are independent of the writing styles of different subjects. Thus, unlike works that identify hand writings by the texture [8] or ink-deposition traces [9], our work solely focuses on the spectral responses of inks for writing ink discrimination.

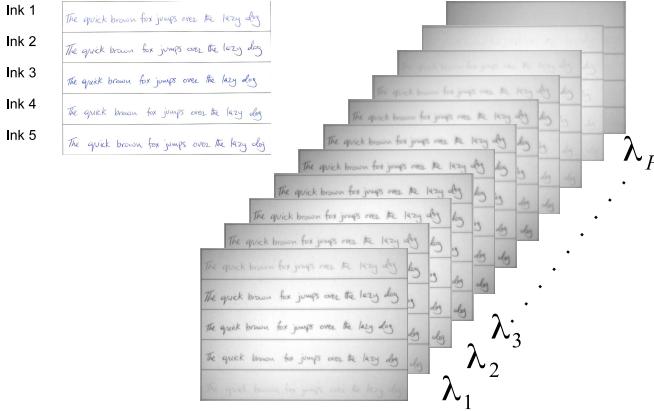
II. WRITING INK HYPERSPECTRAL IMAGE DATABASE

A. Acquisition Setup

Unlike, flatbed document scanners that operate in a line scan manner, our setup is a camera-captured imaging system. The system consists of a monochrome machine vision camera with a native resolution of 752×480 pixels. In front of the camera is a focusing lens (1:1.4/16mm) followed by a Liquid

¹UWA Writing Ink Hyperspectral Image Database
<http://www.csse.uwa.edu.au/%7Eajmal/databases.html>

Blue Inks



Black Inks

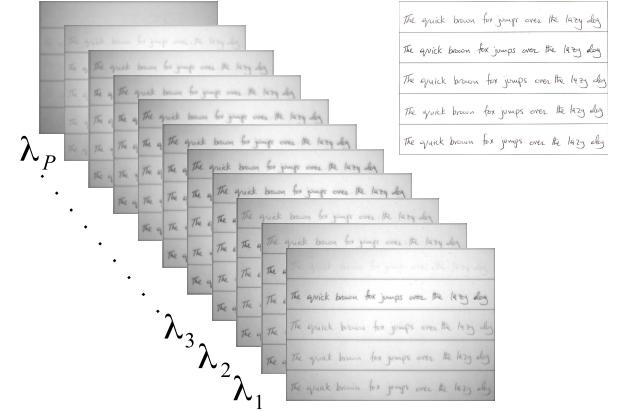


Fig. 1. Hyperspectral images of writing inks. In each line, the subject writes the same sentence with a different ink inside a rectangle. The images $\lambda_1, \lambda_2, \dots, \lambda_P$ represent the bands of the hyperspectral image. Observe how different inks exhibit variation of response in different bands.

Crystal Tunable Filter (LCTF) which is tunable in the range of 400-720 nm. The average tuning time of the filter is 50 ms. The filter bandwidth, measured in terms of the *Full Width at Half Maximum (FWHM)* is 7 to 20nm which varies with the center wavelength. This results in sufficiently narrow-band images. The scene is illuminated by halogen lamps on both sides of the target. For spectral response calibration, the white patch of a color checker was utilized as a reference.

The target is captured in the visible range (400-720nm at steps of 10 nm) that results in a 33 band hyperspectral image. It takes less than 5 seconds to sequentially capture 33 bands of a hyperspectral image. The only downside of the system is the significantly low filter transmission, and camera sensor sensitivity at shorter wavelength, which are compensated by an automatic exposure pre-calibration. Moreover, in a camera based imaging system, the illumination is non-uniformly distributed over the target.

We also collected RGB scanned images at resolutions of 150 and 300 dpi using a flatbed scanner. These RGB images provided baseline information for comparison with HSI. For a fair comparison between RGB and HSI, it is important that their spatial resolutions are similar so that any differences can be attributed to the spectral dimension. This is the reason for selecting low resolution in RGB scans. One advantage of a flatbed scanning system is that the illumination is uniformly distributed over the imaging surface.

All efforts were made to avoid prolonged exposure to ambient/daylight by keeping the samples under cover in dark. This is because different inks are likely to undergo a transformation in their spectral properties due to chemical reactions induced by light. Such an occurrence is, although, favorable in distinguishing two different inks, however, it is somewhat destructive in nature and would bias our analysis. Moreover, all samples were simultaneously collected from the subjects so that their effective age is the same.

B. Database Specifications

The hyperspectral ink database has a total of 70 hyperspectral images of a hand-written note in 10 different inks by 7 subjects. All subjects were instructed to write the sentence,

'The quick brown fox jumps over the lazy dog', once in each ink on a white paper. The pens included 5 varieties of blue ink and 5 varieties of blank ink. It was ensured that the pens came from different manufacturers while the inks still appeared visually similar. Fig. 1 shows hyperspectral images of 5 blue and 5 black inks of the same subject. The corresponding RGB images are also shown on the outer sides.

III. UNSUPERVISED CLUSTERING OF INK PIXELS

Consider a hyperspectral image $\mathcal{I} \in \mathbb{R}^{M \times N \times P}$, where M, N are the number of pixels in each spatial dimension and P is the number of bands in the spectral dimension ($P = 33$). There are two main tasks in ink clustering from a hyperspectral image. The first objective is to compute an $M \times N$ binary mask which associates each pixel to the foreground or background. The foreground represents the ink pixels and the background represents the blank areas of the page. This foreground/background segmentation is achieved using image thresholding. The second objective is to subsequently establish the class membership of each foreground pixel with ink K . This is achieved using unsupervised clustering of spectral responses.

A. Ink Pixels Segmentation

In the first step, the handwritten text requires to be segmented from the blank paper area. The task can be accomplished via image thresholding. However, global image thresholding methods, such as the Otsu's [10] are not sufficient because of the illumination variation over the document, due to camera based imaging. The non-uniform illumination pattern can be observed in Fig. 1. We therefore resort to local threshold based binarization method such as Sauvola's with an efficient integral image based implementation [11]. The thresholding is applied to a single selected channel of the hyperspectral image (640 nm). For thresholding RGB images, we use Otsu's method which is sufficient given there is no illumination variation in the scanned images.

B. Clustering of Ink Pixels

Data clustering is a useful technique to group similar data items in an unsupervised manner [12]. Our data matrix

$X = \{x_1, x_2, x_3, \dots, x_n\} \in \mathbb{R}^{n \times p}$ is the matrix of n spectral response vectors $x \in \mathbb{R}^p$ of the foreground pixels of image \mathcal{I} . To reduce the effect of spatially non-uniform illumination, we normalize each spectral response by dividing it with its 2-norm $\hat{x}_n = x_n/\|x_n\|$. Our objective is to ascertain the label $y = 1, 2, \dots, k \in \mathbb{Z}_+^n$ of each pixel. K-means is a partitional clustering algorithm which divides the n samples into $k \leq n$ groups. It minimizes the squared error between a cluster centroid and its members via the following criterion

$$\arg \min_{\mathcal{C}} \sum_{i=1}^k \sum_{\hat{x}_j \in \mathcal{C}_i} \|\hat{x}_j - \mu_i\|^2, \quad (1)$$

where \hat{x}_j is the j^{th} spectral response sample in the i^{th} cluster \mathcal{C}_i . $\|\cdot\|$ is the squared error between cluster members \hat{x}_j and its centroid μ_i . The number of clusters is k which relates to the number of mixed inks. As the number of mixed inks is unknown, k can theoretically lie in the range $[1, \infty]$. For the sake of this analysis, we fix $k = 2$, i.e., we assume that there are two inks in the image. An implication of this assumption could be that an image with more than two inks would still be grouped into two clusters. At such instances, manual intervention of the document examiner who selects sufficiently smaller areas of interest could result in correct segmentation of the inks. On the other hand, presence of a single ink can be indicated by formation of mixed clusters, providing no definite segmentation.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

We produce mixed writing ink images from single ink notes by joining equally sized image portions from two inks written by the same subject. This makes roughly same proportion of the two inks under question. The blue and black ink samples are separately dealt with in all experiments. This is because it is very unlikely to use such analysis to distinguish blue and black inks, because they can be easily distinguished by visual examination. We also choose not to mix inks among different subjects as it is likely to least affect the results of our analysis. The reason is that we disregard the spatial variation induced by different handwriting styles and only make use of the spectral responses. In our analysis, 5 different inks, taken 2 at a time results in 10 ink combinations, for blue and black color each. In the following experiments, C_{ij} denotes the combination of ink i with ink j .

B. Evaluation Metric

We assess the segmentation accuracy in terms of intersection/union metric, which measures the number of correctly labeled pixels of an ink divided by the number of pixels labeled with that ink in either the ground truth labeling or the predicted labeling [13]. Synonymously, the accuracy is given as

$$\text{Accuracy} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives} + \text{False Negatives}}$$

The segmentation accuracy is averaged over seven samples for each ink combination C_{ij} . It is important to note that according to this evaluation metric, the accuracy of a random

guess (in a two class problem) will be 1/3. This is different to common classification accuracy metrics where the accuracy of a random guess is 1/2. This is because our chosen metric additionally penalizes false negatives which is favorable to observe in a segmentation problem.

C. Experiments and Analysis of Results

We begin with analyzing the effect of different spatial resolutions on segmentation performance in RGB images. Fig. 2 shows the average segmentation accuracy at different resolutions for all ink combinations. We observe the relationship between the image resolution and accuracy. For most of the blue ink combinations, the choice of resolution is 300 dpi. A resolution of 150 dpi seems to be too small but gives an accuracy comparable to 300 dpi in a few ink combinations. Surprisingly, for black ink combinations, no conclusive evidence is available to support any resolution as the accuracy is within the range of (0.3,0.4). This means that RGB does not carry enough information to differentiate between different black inks and the output is close to a random guess. As a final choice, we resort to use the 300 dpi RGB images for further analysis.

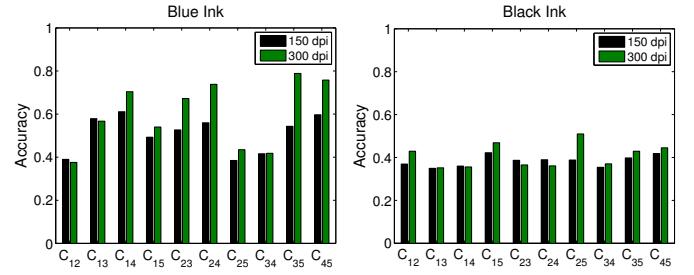


Fig. 2. The effect of spatial resolution on ink segmentation from RGB images.

We now observe how hyperspectral information can be beneficial in discrimination of inks. We compare the segmentation performance of HSI with RGB in Fig. 3. As expected, HSI significantly improves over RGB in most of the ink combinations. This results in most accurate clustering of ink combinations C_{12} , C_{14} , C_{15} , C_{25} , C_{35} and C_{45} . In case of black inks, ink 1 is highly distinguished from all other inks resulting in the most accurate clustering for all combinations C_{1j} . However, it can be seen that for a few combinations, HSI does not show a remarkable improvement. Instead, in some cases, it is less accurate compared to RGB. These results encourage us to further look at HSI in detail in order to take advantage of the most informative bands.

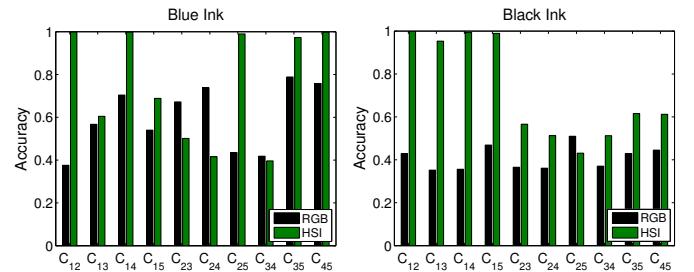


Fig. 3. Comparison of RGB and HSI image based segmentation accuracy.

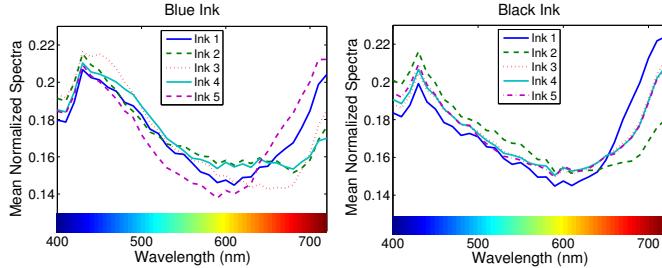


Fig. 4. Spectra of the blue and black inks under analysis. Note that at some ranges the ink spectra are more distinguished than others.

Fig. 4 shows the average normalized spectra of all blue and black inks, respectively. It was achieved by computing the average of the spectral responses of each ink over all samples in the database. It can be observed that the spectra of the inks are distinguished at different ranges in the visible spectrum. In order to evaluate the contribution of sub-ranges, we spectrally divide the hyperspectral data and perform separate segmentation experiments. We empirically divide the visible spectrum into three ranges and name them as low-visible (400nm-500nm), mid-visible (510nm-590nm) and high-visible range (600nm-720nm), respectively. These ranges roughly correspond to the blue, green and red and have been empirically selected because no clear sub-categorization of the visible spectrum, other than this, is defined in the literature. A close analysis of variability of the ink spectra in these ranges reveals that most of the differences are present in the high-visible range, followed by mid-visible and low-visible ranges.

Fig. 5 shows the results of separate experiments in low-visible, mid-visible and high-visible range. Note that for most of the ink combinations, the high-visible range is most accurate, followed by the mid-visible and the low-visible range respectively. Observe that the black ink combinations C₃₄, C₃₅ and C₄₅ are more distinguished in the low-visible range. This can be verified from the ink spectra in Fig. 4. The black inks 3, 4 and 5 are almost similar in the high-visible range, while they show a difference in the low-visible range.

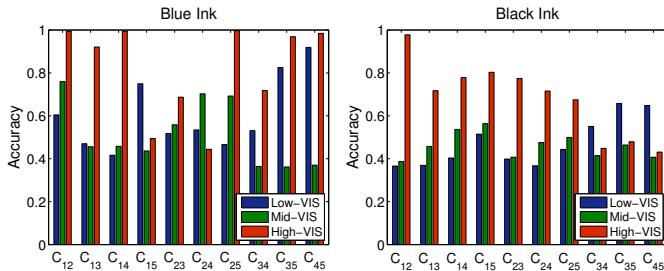


Fig. 5. HSI wavelength range analysis. Observe that the high-VIS range performs better than the mid-VIS and low-VIS ranges.

After clustering of inks in an unsupervised manner using all bands of the HSI, we now improvise our approach by a supervised feature selection. Feature selection is an effective tool to select a small number of variables that optimize certain performance criteria [14]. To this point, we do not have any knowledge of the capability of individual bands of the HSI. Now our aim is to seek a subset of bands which are sufficient to discriminate between different inks. We choose a forward feature selection approach to select a subset of bands. A leave one ink combination out strategy is adopted, in a 10 fold

TABLE I. RESULTS OF FEATURE SELECTION IN LEAVE-1-OUT CROSS VALIDATION. ALL BANDS ARE PRESENTED IN DECREASING ORDER OF IMPORTANCE.

Fold	Ink	
	Blue	Black
1	720,550,490,520,560	710,440,520,700,550
2	720,550,490,520,560	700,530,550,460,720,520,450
3	710,560,490,550,520	700,530,550,720,440,520,710
4	720,550,490,520	700,530,520,460,720,440
5	710,550,560,500,720,480,520	700,530,520,460,720,440,500
6	720,500,560,550,520	700,520,530,460,720,550
7	720,490,550	700,530,550,720,450,520,460
8	720,550,490,520,560	720,520,450,690,530,550
9	720,550,490,520,560	720,520,450
10	720,550,490,520,560	700,530,550,720,450,520,460

cross validation to avoid bias of a particular ink combination in the selected features. In the first step of feature selection, the performance of each individual band on the segmentation of inks is computed. Then, starting with the best band, at each following step, the remaining bands are added (with replacement) one at a time and the best combination of bands is selected. The process continues until adding another band decreases accuracy from the previous step. The same process is repeated in a 10 fold cross validation, each time, learning features from 9 ink combinations and testing on the left out ink combination.

Table I gives the selected bands after the feature selection procedure. It can be seen that the selected bands are fairly consistent and stable. We first analyze results of blue ink. One observation is that the 720nm band (which is in the high-visible range) plays the most important role in distinguishing inks. The second band (550nm) is generally selected from the mid-visible range. The third band (490nm) is selected from the low-visible range. Thus the first three bands are far apart, providing complementary information. Thereafter, different bands are selected from all ranges. We can correlate these findings with the ink spectra in Fig. 4 and observe how all spectra are fairly separated at these wavelengths. In case of black inks, again, the most important band (700 nm) falls in the high-visible range. The next two in the mid-visible range and then in the low-visible range. In both blue and black inks, it is interesting to note that rarely a bands is selected in the range 560-680nm, which is the initial part of high-visible range. This indicates that there is low discriminatory information in this sub-range.

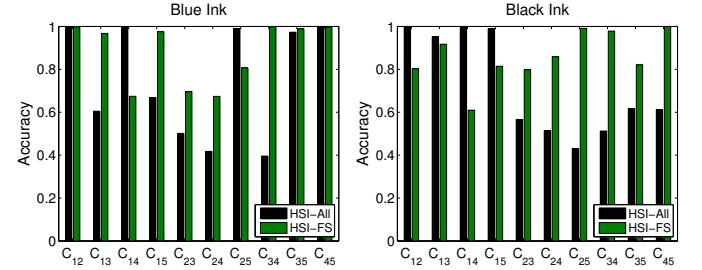


Fig. 6. Results of band selection and comparison with results using all bands.

We now compare the results of using all bands to that of using only the selected bands as shown in Fig. 6. In this figure for HSI-FS, the accuracy of each combination is based on the features selected in Table I by leaving that ink combination out. The feature selection shows its superiority in ink segmentation for most of the ink combinations. Only for a few combinations,

Original Image		
Ground Truth		
Result (RGB)		
Result (HSI-All)		
Result (HSI-FS)		

Fig. 7. Example test images. We purposefully selected two hard cases so that the capability of RGB and HSI based segmentation is visually appreciable.

it shows worse performance. It is likely that the algorithm discontinued at a local maxima and could not further explore additional bands for improved accuracy.

The band selection gives an insight into how a customized multispectral imaging device for writing ink analysis can be made with a smaller number of bands. Bianco et al. [15] developed a multispectral imaging device by combining an imaging sensor with an automatic filter wheel. They empirically selected six different filters (3 visible and 3 NIR) for the prototype device. Such devices may benefit from the findings of the proposed study in the selection of an optimal combination of filters.

Finally, we qualitatively analyze some results on example images of blue and black ink combinations. In Fig. 7, the original images shown are a combination of two blue inks (C_{34}) and black inks (C_{45}), respectively. We show RGB images for better visual appearance. The ground truth images are shown in pseudo-colors, where green represents the first ink pixels and red represents the second ink pixels.

The clustering based on RGB images is unable to group similar ink pixels into same clusters. Instead, a closer look reveals that all the ink pixels are falsely grouped into one cluster whereas most of the boundary pixels are grouped into another cluster. This means that RGB is not sufficient to discriminate inks.

On the other hand, segmentation based on HSI is much more effective compared to RGB. Although, it is not a perfect segmentation result, but, compared to RGB images, it is more reasonable. Finally, we see how the segmentation accuracy improves by using only the selected features in HSI-FS. The selected features exhibit a clear advantage over using all the features. It can be seen that the majority of the ink pixels are correctly grouped according to ground truth segmentation.

V. CONCLUSION

Hyperspectral imaging is of critical value in supporting ink examination. We collected a database of hyperspectral images of different blue and black writing inks. We analyzed and compared the performance of RGB images to HSI images. We also evaluated how a subset of useful bands from hyperspectral images can be effective in ink segmentation. Despite the limitations of our hardware in terms of signal to noise ratio, encouraging results of ink segmentation were achieved.

This proves the reliability of hyperspectral information in questioned ink analysis.

ACKNOWLEDGMENT

This research was supported by ARC Grant DP110102399.

REFERENCES

- [1] E. H. Land, J. J. McCann *et al.*, “Lightness and retinex theory,” *Journal of the Optical society of America*, vol. 61, no. 1, pp. 1–11, 1971.
- [2] V. Aginsky, “Forensic examination of “slightly soluble” ink pigments using thin-layer chromatography,” *Journal of Forensic Sciences*, vol. 38, pp. 1131–1131, 1993.
- [3] S. Joo Kim, F. Deng, and M. S. Brown, “Visual enhancement of old documents with hyperspectral imaging,” *Pattern Recognition*, vol. 44, no. 7, pp. 1461–1469, 2011.
- [4] G. Edelman, E. Gaston, T. van Leeuwen, P. Cullen, and M. Aalders, “Hyperspectral imaging for non-contact analysis of forensic traces,” *Forensic Science International*, vol. 223, pp. 28–39, 2012.
- [5] E. B. Brauns and R. B. Dyer, “Fourier transform hyperspectral visible imaging and the nondestructive analysis of potentially fraudulent documents,” *Applied spectroscopy*, vol. 60, no. 8, pp. 833–840, 2006.
- [6] R. Padoan, T. A. Steemers, M. Klein, B. Aalderink, and G. de Bruin, “Quantitative hyperspectral imaging of historical documents: technique and applications,” *ART Proceedings*, 2008.
- [7] M. E. Klein, B. J. Aalderink, R. Padoan, G. De Bruin, and T. A. Steemers, “Quantitative hyperspectral reflectance imaging,” *Sensors*, vol. 8, no. 9, pp. 5576–5618, 2008.
- [8] K. Franke, O. Bunnemeyer, and T. Sy, “Ink texture analysis for writer identification,” in *Proc. IEEE Workshop on Frontiers in Handwriting Recognition*, 2002, pp. 268–273.
- [9] K. Franke and S. Rose, “Ink-deposition model: The relation of writing and ink deposition processes,” in *Proc. IEEE Workshop on Frontiers in Handwriting Recognition*, 2004, pp. 173–178.
- [10] N. Otsu, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, no. 285–296, pp. 23–27, 1975.
- [11] F. Shafait, D. Keysers, and T. M. Breuel, “Efficient implementation of local adaptive thresholding techniques using integral images,” *Document Recognition and Retrieval XV*, pp. 681 510–681 510–6, 2008.
- [12] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: a review,” *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999.
- [13] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The PASCAL Visual Object Classes Challenge 2012.”
- [14] L. C. Molina, L. Belanche, and À. Nebot, “Feature selection algorithms: A survey and experimental evaluation,” in *Proc. International Conference on Data Mining*, 2002, pp. 306–313.
- [15] G. Bianco, F. Bruno, and M. Muzzupappa, “Multispectral data cube acquisition of aligned images for document analysis by means of a filter-wheel camera provided with focus control,” *Journal of Cultural Heritage [Preprint]*, 2012.