Automated Forgery Detection in Multispectral Document Images using Fuzzy Clustering

Muhammad Jaleed Khan[1], Adeel Yousaf[1], Khurram Khurshid[1], Asad Abbas[2], Faisal Shafait[3]
[1]Department of Electrical Engineering, Institute of Space Technology, Islamabad 44000, Pakistan.
[2]School of Electrical Engineering & Computing, The University of Newcastle, NSW 2308, Australia.
[3]School of Electrical Engineering & Computer Science, National University of Sciences & Technology, Islamabad 44000, Pakistan.
mjk093@gmail.com, adeelyousaf1993@gmail.com, khurram.khurshid@ist.edu.pk, asad.abbas@uon.edu.au, faisal.shafait@seecs.edu.pk

***Abstract*— Multispectral imaging allows for analysis of images in multiple spectral bands. Over the past three decades, airborne and satellite multispectral imaging have been the focus of extensive research in remote sensing. In the recent years, ground based multispectral imaging has gained an immense amount of interest in the fields ranging from computer vision and medical imaging to art, archaeology and computational forensics. The rich information content in multispectral images allows forensic experts to examine the chemical composition of forensic traces. Due to its rapid, non-contact and non-destructive characteristics, multispectral imaging is an effective tool for visualization, age estimation, detection and identification of forensic traces in document images. Ink mismatch is a key indicator of forgery in a document. Inks of different materials exhibit different spectral signature even if they have the same color. Multispectral analysis of questioned documents images allows identification and discrimination of visually similar inks. In this paper, an efficient automatic ink mismatch detection technique is proposed which uses Fuzzy C-Means Clustering to divide the spectral responses of ink pixels in handwritten notes into different clusters which relate to the unique inks used in the document. Sauvola's local thresholding technique is employed to efficiently segment foreground text from the document image. Furthermore, feature selection is used to optimize the performance of the proposed method. The presented method provides better ink discrimination results than state-of-the-art methods when tested on publicly available UWA Writing Inks Dataset.**

Keywords*— Multispectral Imaging; Forgery Detection; Ink Mismatch Detection; FCM Clustering; Feature Selection.

I. Introduction

Each material exhibits a property to transmit, reflect and absorb light in a specific manner, giving that material its unique spectral response. Human eye can easily discriminate different colors [1] but cannot distinguish among colors that are close together in the visual spectrum. The spectral information of visually similar colors is very useful in their discrimination because objects with similar colors but different composition of materials exhibit different reflectance properties which cannot be recognized with a naked eye. Multispectral imaging is a modern spectral sensing technique which captures images using several well defined optical bands in the broad spectral range. It was initially used in satellite and airborne platforms for remote sensing applications but during the recent years, it has found its applications in different industries as well such as food quality and safety inspection, medical diagnosis and image guided surgery, defense and homeland security and forensic examination of artworks and documents.

Multispectral imaging offers added potential to forensic experts for analysis of forensic traces. Due to the increasing portability and speed of multispectral imaging systems, it can be used for quick investigation at crime scenes for viewing and interpreting forensic traces in the actual context. In the recent literature, studies are reported on analysis of several forensic traces such as finger prints, inks, firearm propellants, drugs, dentin, bruises, blood stains, hair, condoms, tapes, paints and fibers [3]. Multispectral imaging has emerged as an effective non-destructive tool for improving readability [2], ink aging and forensic examination [3] of documents. Ink analysis is an important tool in forensic examination of document images. It provides significant information which helps in determination of ink age, backdating, fraud and forgery. Ink mismatch is a key indicator of forgery in questioned documents. A document containing text written with more than one inks can be a forged document. Forensic specialists explore the ink type in an addressed report, if the note consists of more than one kind of inks; it depicts that some manipulations have been made in that document. Ink discrimination techniques can broadly be divided into two categories: destructive and non-destructive techniques. Thin layer chromatography is a destructive ink examination technique [4] which has been utilized to separate mixed inks using capillary action. Thin layer chromatography is an effective method for ink mismatch detection but it is a qualitative test and requires certain measures to quantify the results. Moreover, it is very time consuming, sensitive to temperature and destructive.

Non-destructive ink analysis techniques are more promising as compared to the destructive techniques. The modern non-destructive ink examination methods use multispectral image analysis which has proved its effectiveness in many areas such as object detection and image enhancement [5]. Multispectral image analysis has also been used for identification of forensic traces [6]. Easton et al. [7] introduced multispectral imaging in document image analysis for the first time. Christens-Barry et al. [8] developed EurekaVision, a multispectral imager for acquisition of multispectral document images. Such systems are very useful for forensic specialists in ink analysis. However, ink analysis using band-by-band examination of multispectral images through visual assessment is very difficult and time consuming due to the fact that the documents need to be manually observed by the examiner under each wavelength of light. The examiner needs to make decisions based on the qualitative study of the multi-spectral image. The spectral

978-1-5386-3346-5/18 $31.00 © 2018 IEEE
DOI 10.1109/DAS.2018.26

resolution of a multispectral imager determines the number of wavelengths which can be observed. Automatic ink analysis is essentially helpful for forensic experts as it supports efficient and effective analysis of questioned documents. Over the past few years, a number of methods for automatic ink mismatch detection have been proposed. Khan et al. [9] were able to distinguish two inks mixed in an equal amount in a multispectral document using supervised clustering. Luo et al. [10] leveraged the minute differences between quality of inks in original and fake document images using localized multispectral image analysis.

In this paper, we propose an efficient automatic ink mismatch detection technique using multispectral image analysis. Ink pixels are segmented out using local thresholding and Fuzzy C-Means Clustering (FCM) is used to divide the spectral response vectors of ink pixels into different clusters which relates to different inks used in the document. Experiments are carried out on ink combinations with different mixing ratios and encouraging results are obtained. The proposed method gives a better discrimination between inks in questioned documents as compared to the former methods. Moreover, the effectiveness of feature selection in clustering of ink pixels is also discussed. The rest of this paper is organized as follows: Section II explains the use of multispectral imaging for automated forgery detection in document images, the proposed method is explained in detail in Section III, the experimental results are presented in Section IV and Section V concludes the paper.

II. MULTISPECTRAL IMAGING FOR FORGERY DETECTION

Multispectral cameras can capture reflectance, luminescence, photoluminescence and transmission in various parts of the electromagnetic spectrum such as ultraviolet, visible, infrared and thermal range. A multispectral image is an image stack with each image captured at a narrow spectral band and thus it contains spatial as well as spectral information as shown in Figure 1 using pseudo-colors. The spatial information relates to the geometric relationship of the image pixels to each other while the spectral information relates to the variations within image pixels as a function of wavelength. A multispectral image has two spatial dimensions (S_x and S_y) and one spectral dimension (S_λ). The spectral response at the spatial location (*x*,*y*), an RGB image and a grayscale image rendered from the image stack are also shown in Figure 1. Spectral response of a single pixel in a multispectral image provides information about its constituents and surface of the material. Multispectral imaging was initially developed for remote sensing purposes but it has found numerous applications in ground based imaging in the recent years, including agricultural and water resources control [11-12], military defense, art conservation and archeology [13-14], medical diagnosis [15-16], analyses of crime scene details [17-18], document imaging [19], forensic medicine [20], food quality control [21-22] and mineralogical mapping of earth surface [23]. The broad spectral information provided by multispectral images combined with signal processing allows for identification of the underlying material in images. Materials which are clearly visible in a specific band and obscured in the other bands are reflective within the same band and thus easily detectable by the multispectral imaging system [24]. Selection of appropriate bands in the multispectral data of historic texts and manuscripts allows for identification of inks and pigments for dating of manuscripts [25] and recovery of erased and overwritten scripts [26].

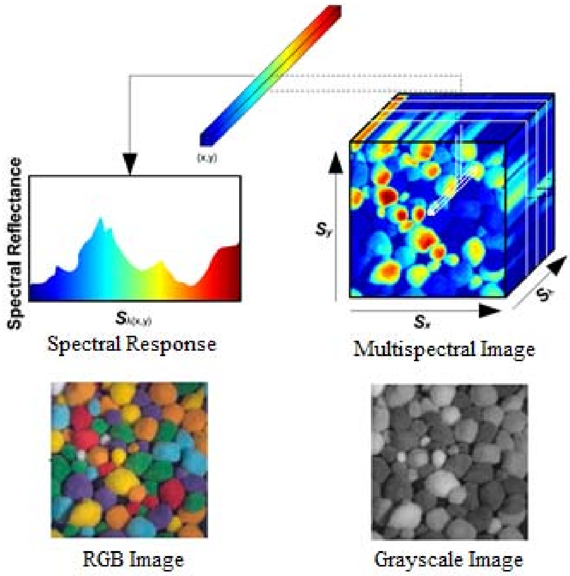


Fig. 1. A multispectral image represented as an image stack, spectral response at the spatial location (x,y), an RGB image and a grayscale image rendered from the image stack.

Over the past few years, various methods based on multispectral image analysis have been proposed for automated forgery detection in documents [27-29]. The use of multispectral image analysis have had a significant impact on forgery detection as compared to the traditional techniques. A multispectral imaging system proposed by Brauns et al. [30] detects forgery in a non-destructive manner in potentially fake documents using an interferometer which uses different moving parts for frequency tuning and subsequently moderates or slows down the acquisition procedure. This work verifies the concept to distinguish between writing inks. Subjective outcomes on a very small database demonstrated that the spectra of ink can be isolated into various classes in an unsupervised way. An advanced and more complex multispectral imaging framework for examination of ancient documents was created at the National Archives of Netherlands [31] which provided high spatial as well as high spectral resolution over a broad spectral range (near UV to near IR). The system was very robust and effective but its acquisition time is very long [32]. Texture [33] and ink-deposition traces [34] have also been employed for writing analysis in document images. Different inks were successfully classified after obliteration of the text in a document [28].

Hedjam et al. [35] used multispectral image analaysis for restoration of historical documents, and focused on the practical situations where the ink is undetectable by human eyes and proposed an advanced technique [36]. Hollaus et al. [37] utilized spectral as well as spatial data in their strategy for enhancing degraded images. Silva et al. [38] concentrated on inks of various ages and devised a chemometric based strategy for forgery identification using multispectral image analysis. Morales et al. [39] developed an embedded system for ink examination using least square Support Vector Machine classifier and multispectral image analysis. A multispectral document analysis system is presented by Khan et al. [9] for detection of different inks in a forged document. By gathering the spectral responses of inks under examination in 33 spectral bands, the technique can recognize the different inks used in a handwritten document. This technique can be holistically applied to handwritten notes and assumes that a similar measure of content has been composed with each of the two inks to be differentiated. However, practical scenarios where documents are forged by modifying a small part of the document may not be handled with that method proposed due to the assumption of presence of two different inks in a document in equal amount. Luo et al. [10] combined detection and unsupervised learning and leveraged the minute differences between quality of inks in original and fake document images using localized multispectral image analysis. Abbas et al. [40] proposed hyperspectral unmixing for discrimination between inks from different pens in a document. Iqbal et al. [28] used hyperspectral unmixing for automatic signature extraction from hyperspectral document images.

III. Proposed Document Forgery Detection Technique

In this section, the proposed FCM based forgery detection technique is presented. Technical specifications of the dataset and the pre-processing steps performed on the dataset are explained in detail followed by multispectral mixing of ink samples for experiments and the use of FCM in the proposed technique.

A. Database

The publicly available UWA Writing Inks Database [9] has been used in this work. The database contains 14 multispectral images of size 752x480 pixels, each with 33 bands in the visible range (400~720nm at steps of 10 nm). A single document, shown in Figure 2, consists of five lines of the English language phrase “The quick brown fox jumps over the lazy dog” written with five different pens with either blue or black ink by seven different authors. Each pen came from a different brand to ensure subtle variations within their inks even if they have the same color. These subtle variations allow us to make use of the multispectral property of these inks, because despite of their similar appearance to naked eye, they have their own unique spectral signatures. Extra precautions were taken to avoid chemical transformation of the inks, due to prolonged exposure in open atmosphere, which may affect the original spectral responses of the inks. All samples were written by the authors in a single session to avoid the effect of ink aging. A multispectral image written with five different blue inks is shown in Figure 3 along with the corresponding RGB image. Figure 4 shows the discrimination of inks offered by multispectral imaging. A selected number of bands at specific wavelengths for two different blue inks (for the word ‘fox’) are shown. Only the pixels belonging to the writing pixels are shown and the pixels of the background are masked out.

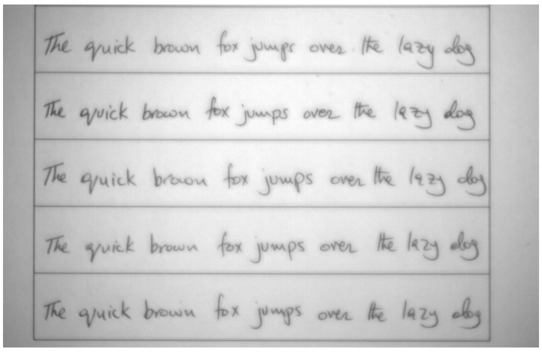


Fig. 2. A single band of a multispectral image with five textlines of the English phrase written with five different pens.

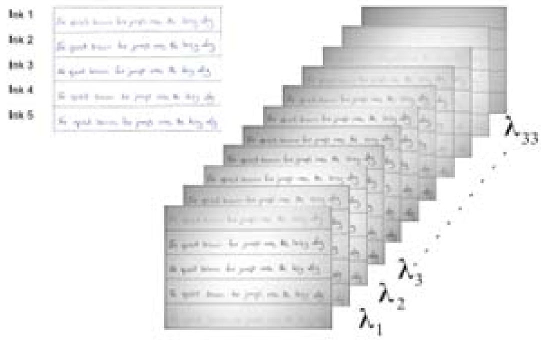


Fig. 3. A multispectral image stack written with five different blue inks and the corresponding RGB image.

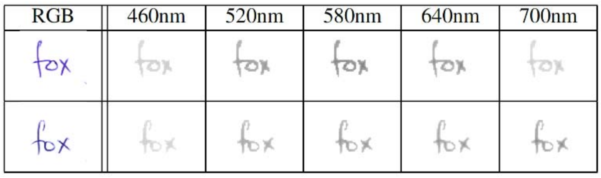


Fig. 4. Ink pixels in the selected bands at specific wavelengths for two different blue inks showing the ink discrimination offered by multispectral imaging.

B. Pre-processing

The aim of pre-processing is to make the multispectral images ready for further experimentation. The handwritten text lines need to be segmented from each band in the image using image thresholding to get rid of the blank paper area. The document images in the dataset [9] contain non-uniform illumination over the image because they have been captured using a camera. Global thresholding techniques such as Otsu binarization [41] are unable to give satisfactory results while the local thresholding techniques such as Sauvola [42] effectively segment the text in images with non-uniform illumination. In the proposed technique, we compute an $M{\times}N$ binary mask using Sauvola’s local thresholding method which represents the ink pixels in the corresponding multispectral image. After segmentation, five multispectral data cubes, one for each text-line of the English phrase are obtained from a single multispectral data cube along with their binary masks as shown in Figure 5. Finally, we are left with 70 multispectral images, each with a single text-line "A quick brown fox jumps over the lazy dog" written by a single type of pen.

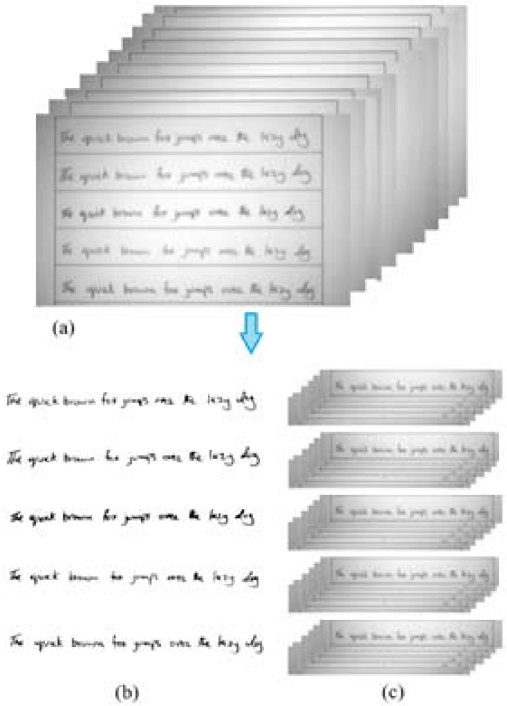


Fig. 5. (a) A multispectral data cube consisting of 33 spectral bands (b) 5 binary masks corresponding to 5 Text Lines (c) 5 Multispectral data cubes of 33 spectral bands each, corresponding to 5 Text Lines in (a)

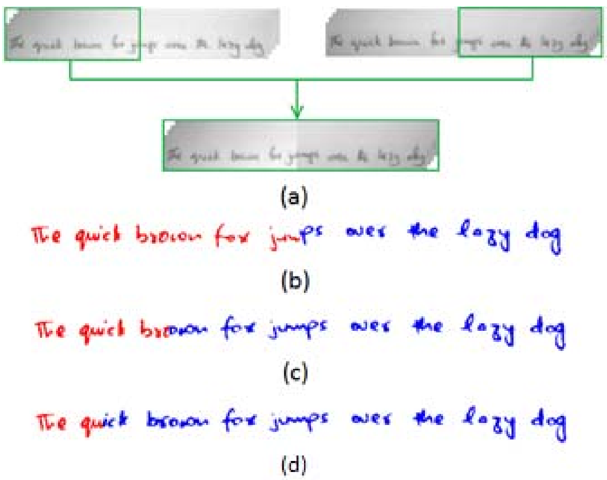


Fig. 6. (a) Multispectral mixing process, (b-d) Combinations of two distinct ink samples with the same color and written by the same author with mixing ratios (b) 1:1, (c) 1:3 and (d) 1:7.

C. *Multispectral Mixing*

We need to distinguish between visually similar inks from different pens in the same document based on their unique spectral signatures. For this purpose, we merge image portions in different ratios from two different ink samples with the same color and written by the same author to produce different mixed ink combinations for experimental analysis as illustrated in Figure 6(a). The mixed combinations of ink samples are summarized in Table I. The combinations are made with mixing ratios 1:1, 1:3 and 1:7 to analyze the performance of the algorithm for different scenarios. Combinations of two inks with different mixing ratios are shown in Figure 6(b-d). Different inks are colored differently for the sake of visualization.

TABLE I. COMBINATIONS OF TWO DIFFERENT INKS WITH SAME COLOR AND SAME WRITER.

Combination	Ink Type 1	Ink Type 2
A	1	2
B		3
C		4
D		5
E	2	3
F		4
G		5
H	3	4
I		5
J	4	5

D. *Fuzzy Clustering*

Clustering is an unsupervised learning technique which groups data on the basis of some similarity metrics such as distance, intensity, and connectivity. Fuzzy clustering techniques such as Fuzzy C-means Clustering (FCM) can retain more information from data as compared to hard clustering techniques such as K-Means clustering [43].

A multispectral document image $I \in \mathbb{R}^{M \times N \times B}$ is taken as the input where M and N represent the spatial resolution and B is the number of spectral bands in the image. Each Spectral Response Vector (SRV) $v \in \mathbb{R}^b$ is divided with its 2-norm to minimize the effects of varying illumination across the image.

$$\widehat{v_k} = \frac{v_k}{\|v_k\|}$$

The matrix of k SRVs of the ink pixels in the document image I as the data matrix for clustering,

$$V = \{v_1, v_2, \dots, v_k\} \in \mathbb{R}^{k \times b}$$

We use FCM to divide the ink pixels into $n \in [1, k]$ number of clusters, which is equal to the number of distinct inks, used in the document. Being a soft clustering technique, FCM starts with randomly assigning weights to each pixel. The weights indicate the extent to which the pixel belongs to each cluster. FCM iteratively computes the new centroids of each cluster and updates the cluster weights of pixels using the new centroids till the algorithm converges. The centroid c_n is the weighted mean of members of the nth cluster. FCM minimizes the following objective function to converge:

$$\underset{C}{argmin} \sum_{i=1}^{k} \sum_{j=1}^{n} w_{i,j} \|\widehat{v_i} - c_j\|^2$$

where $w_{i,j}$ represents the weight of j^{th} cluster for i^{th} pixel. Presence of different inks in a document indicates a forgery. In an unforged document containing a single ink, mixed clusters would be formed which will not provide any precise segmentation. In practical scenarios, the number of inks used in a questioned document is not known. The proposed algorithm can be applied several times with different number of clusters *n* depending on the nature of document. The output with the best segmentation would indicate the correct assumption of the number of inks in the document.

IV. Experimental Analysis

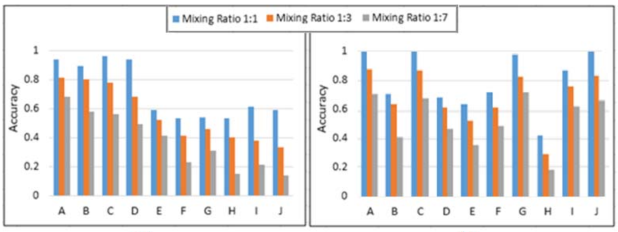
We perform experimental analysis on document images with different ratios of two visually similar but spectrally distinct inks. We evaluate the proposed technique in terms of accuracy, which is defined as the number of correctly labeled pixels of an ink divided by the number of pixels labeled with that ink in either predicted labeling or ground truth labeling [45].

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

The results of experiments on black and blue ink combinations are shown in Figure 7(a) and Figure 7(b) respectively. Most accurate results for black inks were obtained for combinations A, B, C, D and I whereas most accurate results for blue inks were obtained for combinations A, C, G, I and J. It can be seen that accuracy decreases with an increase in the portion of one ink over the other in the ink combination. Thus, some noise removal techniques need to be applied to images with a higher ratio of one ink over the second ink.

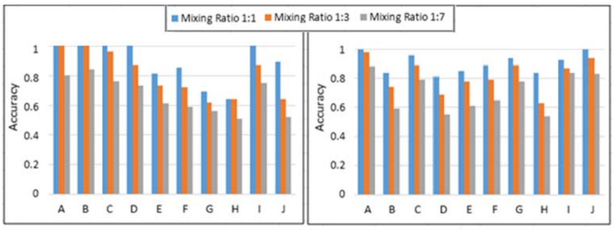
Promising clustering results for some ink combinations such as black ink combinations F, G and H and blue ink combinations E and H could not be obtained. Therefore, we need to select bands with the highest information content to improve our results. The high visible band gives the best discrimination followed by the mid and low visible band [9]. Feature selection effectively selects the best subset of feature to optimize the classification [28]. Stable and consistent bands with the highest discriminatory information are selected and used for classification on the ink combinations shown in Table I. Figure 8(a) and Figure 8(b) show the remarkable improvement in the accuracy after employing feature selection. An overall accuracy of 68% was obtained using FCM which is better than the former methods proposed in [9] and [10] on the same dataset. After employing feature selection in the proposed technique, an overall accuracy of 76% was reported which proves the efficacy of the proposed technique for ink mismatch detection. Blue inks reported better results as compared to black inks because the spectral responses of blue inks are easily separable in the dataset.

We see that majority of the ink pixels are correctly clustered and feature selection has been shown to be a very useful step in clustering of ink pixels. The proposed method gives a better discrimination between inks in questioned documents as compared to the former methods. The experiments were performed on a machine with Intel Core i3 @ 2.40GHz with 6.00GB RAM using MATLAB R2017a. The execution time per hyperspectral document image was 4 seconds.



(a) (b)

Fig. 7. Segmentation accuracy achieved by FCM on (a) black ink combinations and (b) blue ink combinations with different mixing ratios.



(a) (b)

Fig. 8. Segmentation accuracy achieved by FCM and Feature Selection on the (a) black ink combinations and (b) blue ink combinations with different mixing ratios.

V. Conclusion

Multispectral imaging is an effective tool for non-destructive and non-contact examination of forensic traces. Ink mismatch detection is a key step in forgery detection. The spectral information in multispectral documents images is useful for distinguishing between visually similar inks. We used Fuzzy C-Means Clustering (FCM) for dividing the spectral response vectors of ink pixels into different clusters for ink mismatch detection. Experiments were carried out on ink combinations with different mixing ratios, i.e. 1:1, 1:3 and 1:7. The proposed method gives a better discrimination between inks in questioned documents as compared to the former methods. We also found the effectiveness of feature selection in clustering of ink pixels. We expect that these findings will further encourage the use of multispectral image analysis along with state-of-the-art clustering and classification methods in document analysis, particularly for automated questioned document examination.

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] S. J. 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.

[3] 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, no. 1, pp. 28–39, 2012.

[4] V. N. Aginsky, "Forensic examinaton of "slightly soluble" ink pigments using thin-layer chromatography," *Journal of Forensic Sciences*, vol. 38, pp. 1131–1131, 1993.

[5] 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.

[6] 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.

[7] R. L. Easton Jr, K.T. Knox, and W.A. Christens-Barry, ”Multispectral imaging of the Archimedes palimpsest”. 32nd IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Oct 2003, pp. 111-116.
[8] W. A. Christens-Barry, K. Boydston, F. G. France, K. T. Knox, R. L. Easton Jr, & M. B. Toth, ”Camera system for multispectral imaging of documents”. *Proc SPIE Sensors, Cameras, and Systems for Industrial/Scientific Applications X*, pp. 724908-724908, 2009.
[9] Z. Khan, F. Shafait, A. Mian, ”Hyperspectral Imaging for Ink Mismatch Detection”, 12th Int Conf on Document Analysis and Recognition (ICDAR), Aug 2013, pp. 877-881.
[10] Z. Luo, F. Shafait and A. Mian, "Localized forgery detection in hyperspectral document images," 13th International Conference on Document Analysis and Recognition (ICDAR), Nancy, France, Aug 2015, pp. 496-500.
[11] M. Govender, K. Chetty, and H. Bulcock, “A review of hyperspectral remote sensing and its application in vegetation and water resource studies,” Water Sa, vol. 33, no. 2, 2007.
[12] E. Adam, O. Mutanga, and D. Rugege, “Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review,” Wetlands Ecology and Management, vol. 18, no. 3, pp. 281– 296, 2010.
[13] C. Fischer and I. Kakoulli, “Multispectral and hyperspectral imaging technologies in conservation: current research and potential applications,” Studies in Conservation, vol. 51, no. Supplement-1, pp. 3–16, 2006.
[14] H. Liang, “Advances in multispectral and hyperspectral imaging for archaeology and art conservation,” Applied Physics A, vol. 106, no. 2, pp. 309–323, 2012.
[15] O. Carrasco, R. B. Gomez, A. Chainani, and W. E. Roper, “Hyperspectral imaging applied to medical diagnoses and food safety,” in AeroSense 2003. International Society for Optics and Photonics, 2003, pp. 215– 221.
[16] M. A. Afromowitz, J. B. Callis, D. M. Heimbach, L. A. DeSoto, and M. K. Norton, “Multispectral imaging of burn wounds: a new clinical instrument for evaluating burn depth,” Biomedical Engineering, IEEE Transactions on, vol. 35, no. 10, pp. 842–850, 1988.
[17] J. Kuula, I. Pol¨onen,¨ H.-H. Puupponen, T. Selander, T. Reinikainen, T. Kalenius, and H. Saari, “Using vis/nir and ir spectral cameras for detecting and separating crime scene details,” in SPIE Defense, Security, and Sensing. International Society for Optics and Photonics, 2012, pp. 83 590P–83 590P.
[18] R. L. Schuler, P. E. Kish, and C. A. Plese, “Preliminary observations on the ability of hyperspectral imaging to provide detection and visualization of bloodstain patterns on black fabrics,” Journal of forensic sciences, vol. 57, no. 6, pp. 1562–1569, 2012.
[19] 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.
[20] 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, no. 1, pp. 28–39, 2012.
[21] A. Gowen, C. O’Donnell, P. Cullen, G. Downey, and J. Frias, “Hyperspectral imaging–an emerging process analytical tool for food quality and safety control,” Trends in Food Science & Technology, vol. 18, no. 12, pp. 590–598, 2007.
[22] Y.-Z. Feng and D.-W. Sun, “Application of hyperspectral imaging in food safety inspection and control: a review,” Critical reviews in food science and nutrition, vol. 52, no. 11, pp. 1039–1058, 2012.
[23] R. N. Clark and G. A. Swayze, “Mapping minerals, amorphous ma-terials, environmental materials, vegetation, water, ice, and snow, and other materials: The usgs tricorder algorithm,” in Summaries of the Fifth Annual JPL Airborne Earth Science Workshop, vol. 1. JPL Publication, 1995, pp. 39–40.
[24] M. Attas, E. Cloutis, C. Collins, D. Goltz, C. Majzels, J. R. Mansfield, H. H. Mantsch, “Near-infrared spectroscopic imaging in art conservation: investigation of drawing constituents,” J. Cult. Herit. 4 (2003) 127–136.
[25] K. Melessanaki, V. Papadakis, C. Balas, D. Anglos, “Laser induced breakdown spectroscopy and hyperspectral imaging analysis of pigments on an illuminated manuscript,” Spectrochim. Acta Part B 56 (2001) 2337–2346.
[26] C. Balas, V. Papadakis, N. Papadakis, A. Papadakis, E. Vazgiouraki, G. Themelis, “A novel hyper-spectral imaging apparatus for the non-destructive analysis of objects of artistic and historic value,” Journal of Cultural Heritage 4 (2003), pp. 330–337.
[27] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. Prats-Montalban, A. Ferrere, Near infrared hyperspectral imaging for forensic analysis of documentforgery, Analyst 139 (2014) 5176–5184.
[28] K. Iqbal and K. Khurshid, “Automatic Signature Extraction from Document Images using Hyperspectral Unmixing,” Proceedings of the Pakistan Academy of Sciences, 54(3):257-265, 2017.
[29] Z. Khan, F. Shafait, A. Mian, “Towards automated hyperspectral document image analysis,” Proceedings of the 2nd International Workshop on Automated Forensic Handwriting Analysis:, Washington DC, USA, Aug 2013.
[30] E. B. Brauns and R. B. Dyer, “Fourier transform hyperspectral visi-ble imaging and the nondestructive analysis of potentially fraudulent documents,” *Applied spectroscopy*, vol. 60, no. 8, pp. 833–840, 2006.
[31] 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.
[32] 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.
[33] 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.
[34] 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.
[35] R. Hedjam, and M. Cheriet, ”Historical document image restoration using multispectral imaging system”. Pattern Recognition, vol. 46, no. 8, pp. 2297-2312, 2013.
[36] R. Hedjam, M. Cheriet, and M. Kalacska, ”Constrained Energy Maximization and Self-Referencing Method for Invisible Ink Detection from Multispectral Historical Document Images”. 22nd Int Conf on Pattern Recognition (ICPR), Aug 2014, pp. 3026-3031.
[37] F. Hollaus, M. Gau, and R. Sablatnig, ”Enhancement of multispectral images of degraded documents by employing spatial information”, 12th Int Conf on Document Analysis and Recognition (ICDAR), Aug 2013, pp. 145-149.
[38] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. Prats-Montalban, and A. Ferrer, Near infrared hyperspectral imaging for forensic analysis of document forgery, Analyst, vol. 139, no. 20, pp. 51765184, 2014.
[39] A. Morales, M. A. Ferrer, M. Diaz-Cabrera, C. Carmona, and G. L. Thomas, The use of hyperspectral analysis for ink identification in handwritten documents, in ICCST. IEEE, 2014, pp. 15.
[40] A. Abbas, K. Khurshid, and F. Shafait, “Towards Automated Ink Mismatch Detection in Hyperspectral Document Images”, 14th IAPR International Conference on Document Analysis and Retrieval (ICDAR), November 2017, Japan.
[41] N. Otsu, “A threshold selection method from gray-level histograms,” *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.
[42] 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.
[43] 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.
[44] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, “The PASCAL Visual Object Classes Challenge 2012.”
[45] L. C. Molina, L. Belanche, and A. Nebot, “Feature selection algorithms: A survey and experimental evaluation,” in Proc. International Conference on Data Mining, 2002, pp. 306–313.