

Towards Automated Ink Mismatch Detection in Hyperspectral Document Images

Asad Abbas*, Khurram Khurshid† and Faisal Shafait‡

*School of Electrical Engineering and Computing, The University of Newcastle, Callaghan NSW 2308, Australia

†Department of Electrical Engineering, Institute of Space Technology, Islamabad 44000, Pakistan

‡School of Electrical Engineering and Computer Science, NUST, Islamabad 44000, Pakistan

*asad.abbas@uon.edu.au, †khurram.khurshid@ist.edu.pk, ‡faisal.shafait@seecs.nust.edu.pk

Abstract—Hyperspectral imaging helps in identifying patterns and objects in an observed hyperspectral scene on the basis of their unique spectral signatures; such identification is otherwise difficult using regular imaging. Recently, ink mismatch detection analysis based on hyperspectral imaging has shown enormous potential in distinguishing visually similar inks. Such analysis provides significant information to forensic document examiners to determine the authenticity of the questioned documents. However, major challenge still exist in disproportionate ink mismatch detection because it is inherently an unbalanced clustering problem. The presented approach deals with ink mismatch detection in unbalanced clusters by using hyperspectral unmixing scheme. It identifies the spectral signatures (endmembers) of the inks and their corresponding proportions (abundances). Our results show that HySime outperforms other methods in signal subspace estimation. Hyperspectral unmixing is done by using minimum volume enclosing simplex algorithm. Efficacy of the proposed approach is demonstrated by successfully distinguishing varying disproportionate ink datasets generated from UWA database and results are compared with existing state of the art methods in hyperspectral ink mismatch detection field. We expect that these finding will further encourage the use of hyperspectral imaging in document analysis, particularly towards automated questioned document examination.

Keywords—Hyperspectral document images; hyperspectral unmixing; forgery detection; ink mismatch detection;

I. INTRODUCTION

Analysis of inks is of critical importance to address many important concerns about the questioned documents. It leads to the determination forgery, backdating and fraud, thus playing a vital role in establishing the authenticity of the documents. This requires distinguishing between different colors and types of inks. However humans are able to see in the limited electro-magnetic spectrum and are able to successfully distinguish between colors which have different spectral responses in that range [1]. However human eye cannot discriminate between visually similar inks that have unique spectral signatures.

Traditionally ink analysis techniques are broadly divided into two main approaches, destructive and non-destructive analysis. Chemical Solution based document analysis techniques such as Thin Layer Chromatography (TLC) [2] are used by forensic documents experts for ink mismatch detection. Such techniques are based on the fact that different inks used in documents throughout the history in documents have

their own unique chemical composition and have their own distinctive way of reacting with different substances depending upon the reaction environment. Moreover these chemical solutions based techniques also helped in enhancing the extremely deteriorated documents for improved readability. But there are few disadvantages of such methods. Firstly, although applied with all precautions, these techniques are still destructive in nature, and the harms to the important documents are irreversible. They are time consuming and sensitive to temperature changes as well. Secondly, such methods are qualitative in nature and it's difficult to quantify results and often needs large amount of measurements, which are often impossible to take in many scenarios due to its destructive nature.

To overcome such constraints a nondestructive document examination system such as hyperspectral imaging (HSI) has more potential. Hyperspectral document imaging is a powerful non-destructive tool for gathering information about the documents which is not available in the visible spectrum of light. It allows recording and analyzing of the documents in hundreds of narrowly spaced spectral bands, thus revealing the hidden details in the scene of interest without getting in direct contact with it. Hyperspectral imaging has developed as an effective nondestructive tool for enhancing readability of extremely deteriorated documents [3], ink aging and in forensic document analysis [4]. One of the first works in multispectral imaging was performed by Easton et al. [5]. A spectral imaging system was developed by Christens-Barry et al. for studying printed maps and cultural heritage objects [6]. Hedjam et al. [7] proposed algorithm capable of improving the visual quality of degraded image based on multispectral imaging. Hedjam et al. [8] also proposed a mathematical model for improving the readability of extremely deteriorated text. Hollaus et al. [9] introduced a method for enhancing degraded and ancient writings captured by multispectral imaging system on the basis of spectral as well as spatial information. Brauns et al. [10] designed a hyperspectral imaging technique based on Fourier transform for the non-destructive analysis of potentially fraudulent documents.

A more sophisticated hyperspectral imaging system for quantifying and monitoring aging processes of documents was developed by National Archives of Netherlands [11].

High spatial and spectral resolution images were captured from near ultraviolet to infrared range but documents requires near 15 minutes of exposure to be captured [12]. Thus slow acquisition time limits the use of the system when large amount of images are needed to be captured. Other hyperspectral imaging based forensic analysis systems like ChemImage [13] and Foster & Freeman [14] allows manual comparisons of ink samples. Hammond [15] proposed lab color mode technique to differentiate black ballpoint pen inks. Such manual analysis techniques have certain drawbacks as each hyperspectral image can range up to several hundred of spectral bands depending on the spectral resolution of the imaging sensor. So manually reviewing the document under each wavelength of light (i.e. each spectral band) is dreary process as well as it will require great effort by the analysis to settle on a choice on the premise of such subjective investigation.

Morales et al. [16] proposed an approach for ink analysis in pen verification and handwritten documents using Least Square SVM classification. Silva et al. [17] developed a non destructive method to detect fraud in documents based on different chemometric techniques.

Khan et al. [18] proposed a joint sparse band selection based hyperspectral imaging document analysis technique to distinguish different metameric inks. However this work uses equal ink proportions for conducting experiments and highly disproportionate cases were less distinguishable and have low accuracy, limiting its usability in most practical cases.

In this paper, an ink analysis technique based on hyperspectral unmixing is proposed for ink mismatch detection. Our main focus is to distinguish visually similar inks which are mixed in varying proportions to form an unbalance clustering problem. Hyperspectral subspace identification by minimum error algorithm (HySime) [19] is used to estimate the number of inks (or endmembers) present in the image. Moreover, for hyperspectral unmixing, minimum volume enclosing simplex (MVES) algorithm [20] is employed on a dimensionally reduced hyperspectral image cube, which is obtained by weighted affine set fitting [21]. It should be noted that our work is generalized for ink mismatch detection of different mixing proportion of several visually similar inks. Furthermore, unlike Khan et al. [18] no assumptions are made regarding the total number of inks present in the document or their respective mixing proportions. The results of this hyperspectral unmixing approach were also compared with Khan et. al [18] for varying ink proportions as both approaches uses same writing ink hyperspectral database.

II. WRITING INK HYPERSPECTRAL IMAGE DATABASE

A. Database Specification

UWA writing ink hyperspectral image database [18] is used in this work having the spectral resolution of 33 bands and spatial resolution of 752 x 480 pixels in the visible range

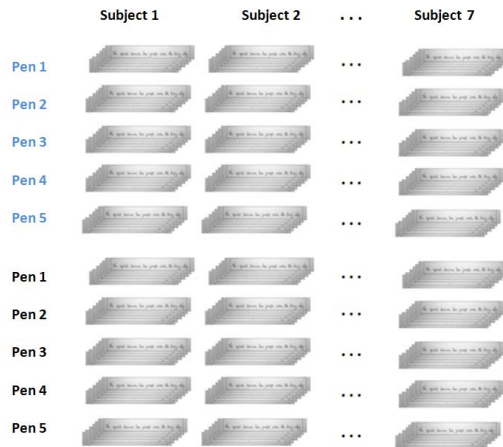


Figure 1: Writing Ink database showing 70 HSI written by 7 subjects with 5 unique blue and black pens

of 400-720nm at steps of 10nm. The database comprises of 70 hyperspectral images of handwritten notes from seven subjects, all the subjects were asked to write the English language phrase “*The quick brown fox jumps over the lazy dog*” with 10 different inks including five blue and five black pens on white paper. Moreover each pen came from different brand to make sure that they have subtle variations within their ink composition even if they have visually same color. Such variations allow us to make use of the hyperspectral property of these inks, because despite of their similar appearance to naked eye, they have their own unique spectral signatures. Fig 1 shows 70 HSI having handwritten phrase “*The quick brown fox jumps over the lazy dog*” written with 5 different pens of black and blue ink by a 7 subjects.

B. Experimental Setup

To analyze and measure the accuracy of our ink mismatch detection approach, we generated various mixed ink HSI datasets from UWA writing ink hyperspectral image database [18]. We produced mixed ink HSI of various proportions by merging two, three or four hyperspectral database images at a time written by the same subject in different proportions. Same color ink samples were merged according to the desired ratios to generate ground truth hyperspectral data cubes. Fig. 2 demonstrates how 1:1 ratio dataset (D1) is generated by merging two halves of blue ink hyperspectral images in equal ratios. Further more no black and blue ink samples were mixed with each other, as it doesn’t serves the purpose of mixing in practical situation, as such mixture of inks can be distinguished easily in visual manner. Additionally, we merged ink samples taken from same subject to negate any spatial variations induced by individual handwriting styles, so that ink mismatch detection can only be made using unique spectral signatures of

different inks.

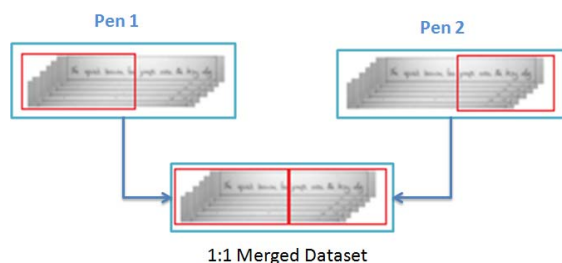


Figure 2: A sample from Dataset D1

In our analysis, five same color hyperspectral ink images (i.e. either 'blue' or 'black'), taken 2, 3 and 4 at a time from UWA database were merged in different proportions. We get total of 10 ink combinations, for each blue and black color ink combination for all cases, except when 4 inks were mixed at a time; we have only 5 ink combinations per ink. Approximate mixing ratios of these ink combinations are given as:

- Two same color inks were merged in ratios of 1:1 (Dataset D1), 1:3 (Dataset D2), 1:7 (Dataset D3), 1:15 (Dataset D4) and 1:31 (Dataset D5).
- Three same color inks were merged in ratios of 1:1:1 (Dataset D6)
- Four same color inks were merged in ratios of 1:1:1:1 (Dataset D7)

Fig. 3 shows ground truth images of dataset D1 to D7, which are labeled in different colors to identify the constituent inks in the note. It should be noted that our generated dataset D1 from UWA database has same ink mixing proportions as described by Khan et al. [18] and is also used for comparison purposes. And to analyze and compare the effect of varying ink proportions on ink mismatch detection accuracy with Khan et al. [18], we also generated same disproportionate ink cases in ratios of 1:8, 2:7, 3:6, 4:5, 5:4, 6:3, 7:2 and 8:1 by five different inks, taken two at a time which results in ten ink combinations, as for blue and black color each, as

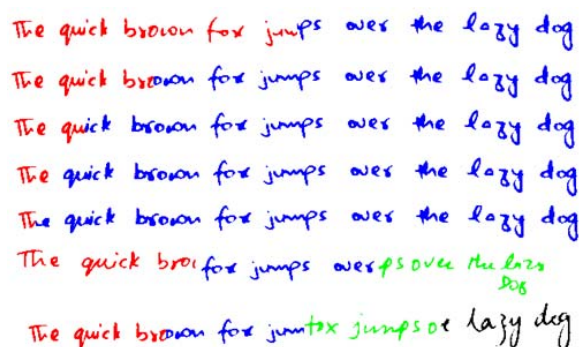


Figure 3: Ground Truth Sample Images of Datasets D1 to D7 respectively

discussed in [18].

III. METHODOLOGY

As hyperspectral images have a mixed pixel nature, spectrum of each pixel is a combination of distinctive materials. A powerful and proficient technique is needed to unmask the details of materials and underlying objects. In this work hyperspectral unmixing is used for analysis of inks in hyperspectral data cubes. Hyperspectral unmixing [22] is a process of decomposing the observed spectrum of a hyperspectral scene into a collection of endmembers and their relative proportions (or abundances). It is vital in identifying individual objects from a hyperspectral scene. There are three major steps in hyperspectral unmixing:

- Dimension Reduction
- Endmember Extraction
- Inversion

Dimension reduction is helpful for reducing the overall complexity of the given dataset and facilitates in the subsequent processes namely the endmember extraction and inversion. Typically maximum noise fraction (MNF) [23] and principal component analysis PCA [24] algorithms along with model estimation algorithms such as hyperspectral signal subspace identification by minimum error (HySime) [19] and virtual dimensionality reduction (VD) [25] are used for dimension reduction.

Endmember extraction is a process to determine the chemical species (also called endmember or materials) that contribute to the measured spectra. Many endmember extraction algorithm exists including pixel purity index (PPI) [26], convex cone analysis [27] and vertex component analysis (VCA) [28]. Most of the endmember extraction algorithms need prior information about the number of endmembers present in the scene. In this work HySime [19] is used for estimation of number of endmembers present in the hyperspectral scene.

Inversion is the final process in hyperspectral unmixing; it involves estimation of the abundances associated with endmembers. In general, unmixing algorithms falls into two broad categories. Some of these unmixing methods are able to estimate the endmembers spectral signatures and their corresponding abundances simultaneously, for example, iterated constrained endmembers (ICE) [29], non-negative matrix factorization (NMF) [30], minimum volume transform (MVT) [31] and joint Bayesian approach (JBA) [32]. While the second type of unmixing algorithms can either only extracts the endmembers present in the scenes or estimate their abundances, like fully constrained least squares (FCLS) [33]. We used Minimum volume enclosing simplex (MVES) algorithm [20] which can estimate the endmembers spectral signatures and their corresponding abundances simultaneously.

A. Estimation of Number of Endmembers Using HySime

Correct estimation of number of endmembers in an observed hyperspectral scene is vital for effective and successful unmixing of hyperspectral images. However due to inevitable presence of noise and outliers it has become quite a challenging task. Thus the key to solve this problem lies in the correct estimation of noise. Hyperspectral subspace identification by minimum error algorithm (HySime) [19] is used in this work for estimating the number of endmembers present in hyperspectral scene.

The HySime algorithm consists of two main parts, namely noise estimation and signal subspace identification. The main aim of this method is to obtain the subset of eigenvectors that effectively describes the signal subspace in minimum mean squared error sense. Reader is referred to [19] for more detailed discussion on HySime.

B. Minimum Volume Enclosing Simplex Algorithm

Minimum volume enclosing simplex (MVES) [20] algorithm is employed on a dimensionally reduced hyperspectral data obtained by affine set fitting method [21]. One of the limitation of MVES is that it needs number of endmembers ‘N’ present in hyperspectral scene to be known a priori. In our case we estimated ‘N’ beforehand using HySime [19] as discussed in section III-A. MVES algorithm successively leads to the unique identification of endmember signature matrix \mathbf{A} and their relative abundances $\mathbf{s}[n]$. MVES algorithm uses convex analysis concepts like convex hull and affine hull [34]. The unmixing problem of finding the minimum volume simplex enclosing all the dimensionally reduced pixel vectors are described as optimization problem in [20].

C. Post Processing

After hyperspectral unmixing, post processing is needed to get the desired results. Since endmember estimation algorithm (HySime) tends to overestimate the number or materials (endmembers) present in the hyperspectral scene, so its imperative that we manually discard those abundance maps which are not coherent, i.e. those which spread throughout the line. Finally, a thresholding scheme is applied on useful abundance maps to get the final results.

IV. RESULTS

A. HySime:Noise Estimation and Eigen Analysis

Noise estimates are obtained using multiple regression based approach and it effectively removes the noise from the useful information. We used eigen decomposition based approaches like MNF [23] and PCT [24] on our generated datasets and compared the results with HySime. The results show that in case of PCT and HySime, first few components contains the most of the spectral energy, while in case of MNF, there is gradual decrease in magnitudes, thus showing the presence of noise along with the signal sources. Table I

shows the first 10 eigenvalues acquired by PCT, HySime and MNF and their related cumulative percentage of spectral energy for a sample hyperspectral image from Dataset 7. Eigenvalues are in arranged in descending order and to get a better picture of the eigenvalues, the cumulative percentage is calculated, so that the variance exhibited by first few eigenvalues can be studied.

From Table I, we can analyze that by using Hysime algorithm we can achieve more than 99.99% of total spectral energy in the first 8 eigenvalues, as compared to the spectral energy given by first eight eigenvalues of PCT and MNF, which is 99.96% and 32.11% respectively. So it is clear from this observation that in case of HySime algorithm maximum amount of energy is contained in the first few eigenvalues.

B. HySime:Signal Subspace Estimation

Signal dimension is estimated by minimizing the mean square error (MSE) between the original signal and the noisy projection of it. The number of negative terms in the minimization \hat{k} dictates the signal subspace dimension. The optimum minimization for dataset 6 is given by $k=8$ after which the minimum square error curve start rising again. Fig. 4 shows the subspace dimension \hat{k} vs mean square error graph, including the noise error and projection error as a function of signal subspace dimension of a sample HSI from dataset D6.

In this sample case from dataset D6, the signal subspace dimension estimated by HySime $\hat{k} = 8$, which literally means that in the given dataset there are 8 different signal sources present, each having its own distinct spectra. But in general that is not true; most of the endmember estimation algorithms overestimate the number of spectrally distinct materials present in the hyperspectral scene due to the presence of unknown signals. This estimate ($\hat{k} = 8$) is used as an input parameter for MVES algorithm to determine the endmember signatures and their corresponding abundance maps in the next step. In our case, HySime outperformed other state-of-the-art endmember estimation algorithms. Table II shows the comparison of different endmember estimation algorithms on datasets. HySime algorithm is compared to

Table I: Eigenvalues and their corresponding Cumulative Percentage of Spectral Energy

HySime	Spectral Energy	PCT	Spectral Energy	MNF	Spectral Energy
0.42103	97.2741	0.39678	97.0557	2.1995	4.1506
0.00692	98.8735	0.00693	98.7529	2.1812	8.2666
0.00367	99.7224	0.00369	99.6554	2.1663	12.3546
0.00069	99.8834	0.00071	99.8297	2.15	16.4117
0.00038	99.9716	0.00039	99.927	2.1247	20.4211
4.68E-05	99.9824	6.11E-05	99.942	2.0927	24.3701
2.81E-05	99.9889	4.25E-05	99.9524	2.0622	28.2615
2.15E-05	99.9939	3.52E-05	99.961	2.0402	32.1115
4.80E-06	99.995	1.82E-05	99.9655	2.0265	35.9355
3.43E-06	99.9958	1.02E-05	99.968	2.0155	39.7389

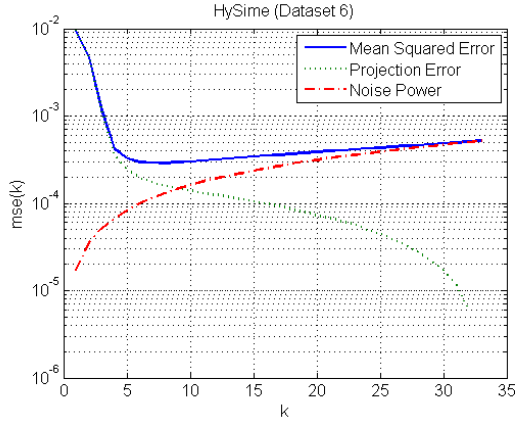


Figure 4: Number of Eigenvalues vs. Cumulative Percentage of spectral energy respectively

GENE-AH [21], O-GENE [35] and HFC [25] methods. HySime outperformed other algorithms

Table II: Estimated Numbers of Endmembers Comparison (Mean of Over 50 Independent Runs for Sample HSI from Datasets)

Datasets	HySime Mean	GENE-AH Mean	O-GENE Mean	HFC Mean
D1	8	27.46	23.77	12
D2	7	26	22	14
D3	7	23	21	12
D4	7	27	23	12
D5	7.05	28	22	11
D6	8	27	24	20
D7	8	27	24	15

C. MVES: Endmember Extraction and Inversion

Minimum volume enclosing simplex algorithm is used for spectral unmixing i.e. to estimate the endmember signature matrices and their corresponding abundance maps. The HySime estimate \hat{k} as discussed in section IV-B, is used as an input parameter for MVES algorithm to determine the endmember signatures and their corresponding abundance maps. As for dataset D7 the mean \hat{k} value was 8, so by using $N = 8$ as input parameter in MVES we get 8 corresponding endmember signature matrices and their relative abundances. Fig. 5 shows useful abundance maps for a sample HSI taken from dataset D7. Fig. 6 shows the abundance maps which are manually discarded in post processing step as discussed in section III-C.

D. Final Results

Segmentation accuracy is determined in terms of intersection/union metric, which calculates 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 [36]. The segmentation accuracy 'A' is averaged over 10 and 5 samples for datasets D1 to D6

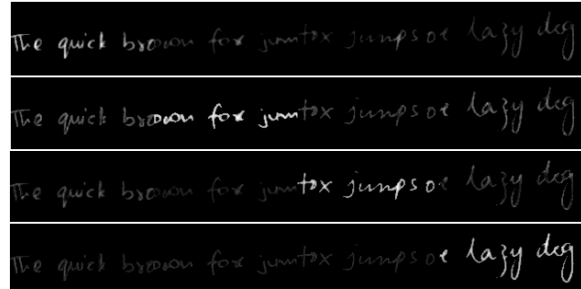


Figure 5: Abundance Maps corresponding to a sample HSI from dataset D7

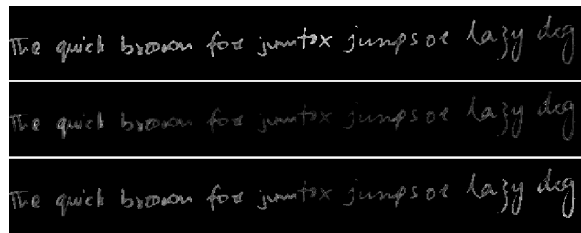


Figure 6: Four discarded abundance maps corresponding to a sample HSI from dataset D7

and dataset D7 respectively for each blue and black ink and is given as:

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

Fig. 7 shows the average segmentation accuracy for Datasets D1 to D7 for blue and black inks. Fig. 8 shows the qualitative comparisons of the sample ground truth (GT) images corresponding to dataset D1 to dataset D7 and the final result after hyperspectral unmixing.

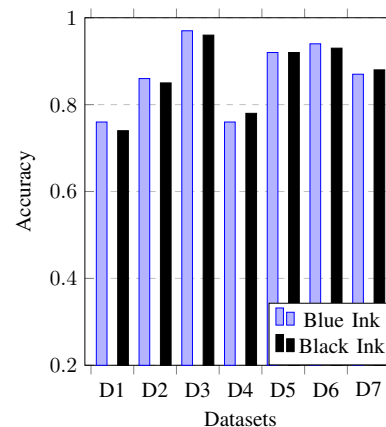


Figure 7: Average Ink Segmentation Accuracy for Datasets D1 to D7

Dataset D1	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D2	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D3	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D4	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D5	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D6	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog
Dataset D7	
Ground Truth	The quick brown fox jumps over the lazy dog
Final Result	The quick brown fox jumps over the lazy dog

Figure 8: Comparisons of Ground Truth Images (Dataset D1 to D7) with Final Results

E. Result Comparisons

Ink mismatch detection results are compared with Khan et al. [18] as both approaches use the same writing ink hyperspectral image database. To make this comparison more meaningful, we generated the same proportion ink cases as used by Khan et al. [18] and discussed in section II-B. The ink mismatch detection accuracies of cases in which inks are mixed in equal proportions are compared in Fig 9 for blue and black inks.

Results indicate that both approaches show great potential in ink mismatch detection, especially in cases where inks are mixed in equal proportions as shown in Fig.9. But we are more interested in studying the efficacy of both approaches in disproportionate ink mismatch detection cases which is inherently an unbalanced clustering problem. The effect of varying ink proportions on ink mismatch detection accuracy are compared in Fig. 10. Unlike the first comparison, these ink mixtures consist of highly unbalanced clusters. Khan et al. [18] shows satisfactory results when the ink clusters are nearly balanced but its accuracy deteriorates as we approach to highly unbalanced ink clusters as evident in highly unbalanced cases such as 1:8 and 8:1 as shown in Fig. 10.

V. DISCUSSIONS

Hyperspectral document imaging has great potential for determining the authenticity of questioned documents in forensic document examination. We proposed a hyperspectral unmixing based method for ink mismatch detection in cases where two or more inks are not mixed in equal proportions. We compared our result with Khan et al. [18] and demonstrated that our approach showed great potential in highly disproportionate ink cases, which can be seen for ink proportions 1:8 and 8:1 in Fig 10. Disproportionate ink mismatch detection is challenging because it is inherently an

unbalanced clustering problem. We used hyperspectral subspace estimation by minimum error algorithm (HySime) to estimate the number of inks (or endmembers) present in hyperspectral document image. Minimum volume enclosing simplex algorithm (MVES) is used for endmembers extraction and inversion. One of the limitations of this unmixing scheme is that as endmember estimation algorithm (HySime) tends to overestimate the number of inks (endmembers) present in the hyperspectral scene so we have to manually discard those abundance maps which are not coherent i.e. those which spread throughout the scene in post processing step. The results indicate that this scheme showed high accuracy in ink mismatch detection for varying proportion ink. We hope that the results presented in this paper will further motivate the researchers to explore new exciting challenges towards automated hyperspectral document analysis.

REFERENCES

- [1] E. H. Land and J. McCann, "Lightness and retinex theory," *JOSA*, 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. 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.
- [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, no. 1, pp. 28–39, 2012.
- [5] R. L. Easton Jr, K. T. Knox, and W. A. Christens-Barry, "Multispectral imaging of the archimedes palimpsest," in *2012 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE Computer Society, 2003, pp. 111–111.

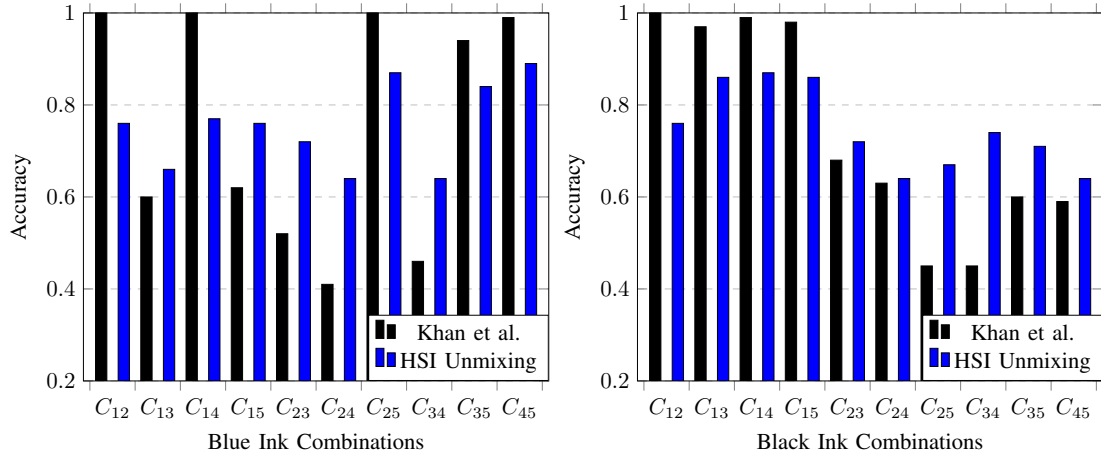


Figure 9: Comparison of ink mismatch detection accuracy of equally proportion mixed inks

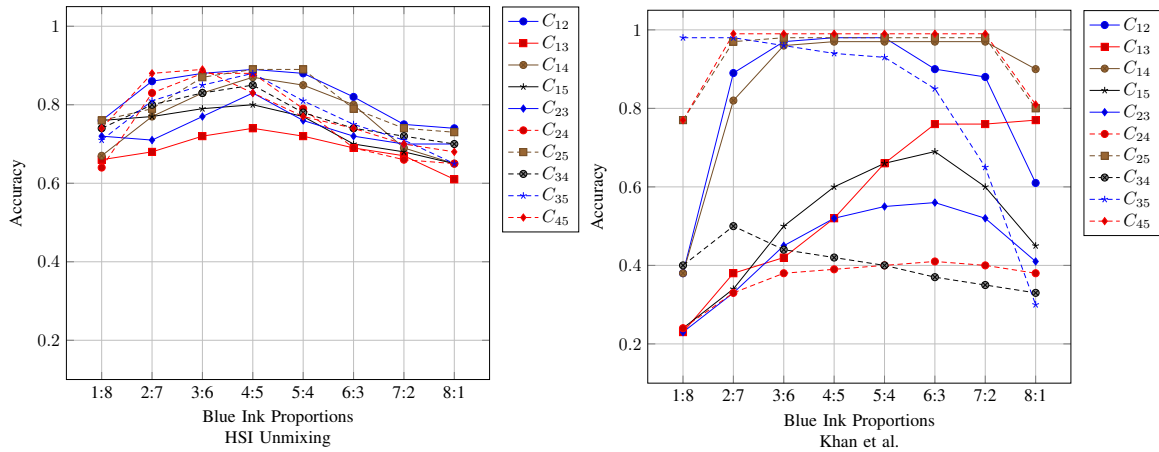


Figure 10: Comparison of ink mismatch detection accuracy of different Blue ink proportions

- [6] W. A. Christens-Barry, K. Boydston, F. G. France, K. T. Knox, R. L. Easton Jr, and M. B. Toth, "Camera system for multispectral imaging of documents," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2009, pp. 724 908–724 908.
- [7] R. Hedjam and M. Cheriet, "Historical document image restoration using multispectral imaging system," *Pattern Recognition*, vol. 46, no. 8, pp. 2297–2312, 2013.
- [8] R. Hedjam, M. Cheriet, and M. Kalacska, "Constrained energy maximization and self-referencing method for invisible ink detection from multispectral historical document images," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*. IEEE, 2014, pp. 3026–3031.
- [9] F. Hollaus, M. Gau, and R. Sablatnig, "Enhancement of multispectral images of degraded documents by employing spatial information," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*. IEEE, 2013, pp. 145–149.
- [10] 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.
- [11] 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.
- [12] —, "Quantitative hyperspectral imaging of historical documents: technique and applications," *ART Proceedings*, 2008.
- [13] ChemImage, <http://www.chemimage.com/>.
- [14] Foster and Freeman, <http://www.chemimage.com/>.
- [15] D. L. Hammond, "Validation of lab color mode as a nondestructive method to differentiate black ballpoint pen inks*," *Journal of forensic sciences*, vol. 52, no. 4, pp. 967–973, 2007.

- [16] 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 *Security Technology (ICCST), 2014 International Carnahan Conference on*. IEEE, 2014, pp. 1–5.
- [17] C. S. Silva, M. F. Pimentel, R. S. Honorato, C. Pasquini, J. M. Prats-Montalbán, and A. Ferrer, "Near infrared hyperspectral imaging for forensic analysis of document forgery," *Analyst*, vol. 139, no. 20, pp. 5176–5184, 2014.
- [18] Z. Khan, F. Shafait, and A. Mian, "Automatic ink mismatch detection for forensic document analysis," *Pattern Recognition*, 2015, (In Press).
- [19] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 46, no. 8, pp. 2435–2445, 2008.
- [20] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4418–4432, 2009.
- [21] A. Ambikapathi, T.-H. Chan, C.-Y. Chi, and K. Keizer, "Hyperspectral data geometry-based estimation of number of endmembers using p-norm-based pure pixel identification algorithm," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 51, no. 5, pp. 2753–2769, 2013.
- [22] N. Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55–78, 2003.
- [23] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 26, no. 1, pp. 65–74, 1988.
- [24] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.
- [25] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 42, no. 3, pp. 608–619, 2004.
- [26] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of aviris data," 1995.
- [27] A. Ifarraguerri and C.-I. Chang, "Multispectral and hyperspectral image analysis with convex cones," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 37, no. 2, pp. 756–770, 1999.
- [28] J. M. Nascimento and J. M. Bioucas Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 4, pp. 898–910, 2005.
- [29] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "Ice: A statistical approach to identifying endmembers in hyperspectral images," *IEEE transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, 2004.
- [30] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [31] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 3, pp. 542–552, 1994.
- [32] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4355–4368, 2009.
- [33] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 39, no. 3, pp. 529–545, 2001.
- [34] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [35] A. Ambikapathi, T.-H. Chan, C.-H. Lin, and C.-Y. Chi, "Convex geometry based outlier-insensitive estimation of number of endmembers in hyperspectral images," in *Proc. IEEE Workshop Hyperspectr. Image Signal Process*, 2013, pp. 1–4.
- [36] M. Everingham, L. Van Gool, C. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge 2012 (voc2012)," 2012.