

Watermark location via back-lighting modelling and verso registration

Jamal Said · Hazem Hiary

Received: 18 September 2014 / Revised: 10 February 2015 / Accepted: 23 February 2015 /
Published online: 5 March 2015
© Springer Science+Business Media New York 2015

Abstract We consider the location of paper watermarks in documents that present problems such as variable paper thickness, stain and other damage. Earlier work has shown success in exploiting a computational model of backlit image acquisition – here we enhance this approach by incorporating knowledge of surface verso features. Robustly removing recto features using established techniques, we present a registration approach that permits similarly robust removal of verso, leaving only features attributable to watermark, folds, chain lines and inconsistencies of paper manufacture. Experimental results illustrate the success of the approach.

Keywords Paper watermark · Image registration · Back-lighting · Recto and verso removal

1 Introduction

Ancient documents may contain watermarks. A watermark is a hidden pattern embedded intentionally in paper texture during its manufacture, which can be recognised as various shades of lightness or darkness when viewed by transmitted light; refers to thickness variations in paper [4]. They are not always obvious on casual inspections because of the disparity of their visibility. Thence, sometimes they may require more effort in order to be extracted due to existence of some kind of interference. Often, this interference can be presented in writings on front (recto) and back (verso) of a single paper that may overlap the watermark patterns.

Since ancient documents are considered as valuable historical artifacts, many documents of interest have been kept in private collections; so it can be difficult for researchers and scholars to have access to those collections. Various efforts have been developed in order to reproduce and exploit watermarks to assist in studying them. However, just depending

J. Said · H. Hiary (✉)
The University of Jordan, Amman 11942, Jordan
e-mail: hazemh@ju.edu.jo

on reproduction techniques is usually not enough due to obstructions left on paper that interfere the watermark patterns. Thence, the need for locating and extracting their designs with minimal interference is raised [5, 17].

Watermarks in papers provide and convey great historical information; they are very useful in paper examination because they were mainly used for dating, identifying paper sizes, determining paper usage and paper quality, trademarks of the paper-makers, paper mills, and locations [18]. They can also reveal facts and realities about historical connections relevant to the past by tracing and studying them. Sometimes, they can be used to correct errors in dating documents specially when observing twin watermarks [27].

Watermarks have attracted researchers for centuries as being a human interest through identifying and classifying them. Therefore, there is a growing need among librarians and antiquarians to make easier access to documental heritage by the public and scholars, to avoid damaging the original documents and that because most of them are too fragile [30]. Old manuscripts are ruining over time because of being affected by natural processes. Thence, reproduction techniques, e.g. back-lighting, that capture the complete representation and concern with the hidden information embedded in the paper in addition to the paper surface has been widely applied as one of the preservation methods by creating digital copies of them that can remain for a limitless time.

Back-lighting reproduction technique is characterised by being digital and can easily be available in the hands of individual scholars because it is cheaper, safer, and easier to use than other reproduction techniques [15]. Since it requires a camera, archiving scenes using multispectral imaging, e.g. a chromatic camera, that captures colored scans of paper materials has an incalculable value in preserving old manuscripts for the future and enriching their digital documentation. Due to the technological advancement, the technical specifications of digital cameras are evolved over time in terms of image sampling and quantization that allows capturing higher quality images with more details. Which in turn, it becomes an important tool for analysis and documentation of old manuscripts.

Since back-lighting is being digital, it allows further image processing techniques to be applied easily by the means of digital computers in order to highlight watermark patterns and remove interference. As observed, manipulating the digitised manuscripts to improve human perception or machine recognition by image processing techniques is being increasingly employed in the management of libraries and archives [30], and it is becoming increasingly clear that watermarks extraction from ancient manuscripts is being more acknowledged as a valid application of information technologies in the field of cultural patrimony.

This paper presents an approach composed of several digital image processing methods that can work successfully on challenging materials after being digitally reproduced to extract paper watermarks. The main focus in this work is to integrate an automatic image registration algorithm that can work successfully in aligning front and back sides of ancient manuscripts in order to attempt minimizing, as much as possible, unwanted interference that obstructs the watermark patterns caused by recto and verso writings.

This paper considers an existing successful approach to watermark location and extraction that exploits backlit images, and improves it by incorporating a verso registration phase that allows a more precise identification of paper features, of which the watermark is one.

2 Background

We present background pertaining to the research. The work presented in this paper is influenced by two categories of techniques: watermark reproduction and image registration.

2.1 Watermark reproduction

Many digital and non-digital techniques have been developed in order to reproduce paper watermarks such as manual tracing [14], rubbing [13], Dylux [12], Ilkley [24], Phosphorescence [26], Back-lighting [8, 17, 21, 28], Thermal photography [22] and Radiographic techniques (Beta-, Soft X-, and Electron-radiography) [1, 2, 6, 7, 31]. A full review of these techniques may be found in [5, 15].

Among these techniques, tracing, back-lighting and radiography are the most common. However, radiographic techniques are expensive and require precautions and special reproduction environment. Tracing is cheap, easy-to-use technique, but it is not accurate and may damage the original paper. Dylux, Ilkley, Thermography, and Phosphorescence techniques require special equipment and environment; this may prevent scholars from being granted access for libraries housing manuscript and rare book collections. Rubbing is also cheap and simple, but may damage the paper.

Our proposed approach is based on back-lighting reproduction technique because it differs from the other techniques in that it is purely digital; that allows to highlight watermark patterns and remove interference caused by writing ink (on both sides of paper) using computers. As a result, acquired digital images can be compared, processed, stored, retrieved and accessed easily. Also because it is simple, requires relatively low cost equipment, produces good image quality, it does not require darkroom conditions, and considered as a safe solution for capturing all details in paper. This makes it easier to preserve and store them in digital archives to be accessed remotely.

During the recent years, much of the previous works has been conducted on watermark extraction after being reproduced by manipulating images digitally trying to isolate clean watermark representations and to improve their appearance. Most commonly used digital image processes are mathematical morphology, histogram enhancement, edge detection, image segmentation, region extraction and image subtraction. However, the majority of these works lack the orientation of automatic parameters estimation [17]. The aim behind these works was to build watermark databases [21] and web archives such as [16]. A review of previous works related to digital manipulation of watermark images can be found in [15].

The Combination of back-lighting technique for digitization and applying digital image processing algorithms becomes an efficient method for watermark extraction. This combination has widely used in several research works as in [8, 15, 21, 28].

The work presented in [15] is divided into two approaches. The first, a bottom-up approach presented a prototype to extract paper watermarks using a sequence of image processing algorithms. The approach pre-processes images to remove interferences and highlight the watermark, followed by segmentation, which achieves localization and extraction of watermark patterns and chain lines (caused by wires placed vertically along the mould during paper production). They evaluated this approach with human opinion. The extracted watermark designs were exported in vector form. Their system gives effective results with the minimum interference compared to others' work. This approach used only the transmitted image for processing. Although, it successfully locates different kinds of watermarks in several data sets but it was limited to specific types of other data sets. These

data sets are characterised by thin pen strokes, thin and uniform paper, and clear watermark designs. The results were in low interference and a strong watermark signal [17].

The second is a top-down approach; it is a model-based technique to locate watermarks in difficult manuscripts. This approach serves as watermark image retrieval utility; it managed to remove recto material successfully, and a statistical approach was developed to locate watermark fragments from a known lexicon. Results show a very good record of retrieval. Web archives are available on-line of the tested manuscripts as a result of this work. This approach requires both reflected and transmitted recto images for each page [5].

2.2 Image registration

Registration has been widely studied in a number of domains and modalities: In this work we wish to co-register reflected verso scans with semi-processed recto scans. The objective is to estimate a transform that compensates for geometric distortions, and then use it to register the two images [9, 10]. In document processing, this task has hitherto often been performed manually [36].

Existing image registration techniques may be categorised as intensity-based or feature-based [38]. These methods find the pairwise (point-by-point) correspondence between base and target images, following which a spatial transformation is determined to map source to target. An extensive survey can be found in [11].

Double-sided manuscripts are often affected by bleed-through and show-through effects [25, 29, 32, 33], and many approaches have been proposed to reduce this interference to improve both human and machine readability. According to the information used, these are categorised as blind and non-blind separation [34]:

- Blind segmentation attempts to clean the front side of a document without referring to the verso. Usually, bleed-through interference is regarded as background noise and removed using threshold-like techniques. These methods are ineffective for seriously damaged documents, where bleed-through intensity is comparable to that of foreground texts. Thresholding methods then usually fail.
- Non-blind approaches use information from both sides of a document, but require accurate alignment of recto and verso images. Most non-blind approaches rely on manual alignment, which is slow, impractical, imprecise and involves human interactions.

Perfect registration of recto and verso images of a page is difficult for several reasons [35]. Firstly, registration is between verso features and their partially shown bleed-through visible in the recto. Secondly, complicated local deformations such as warped or uneven surfaces could be caused by the bounding effect or the unevenness of aging paper. Thirdly, background noise due to decolorization or stains can also affect the registration result.

Manual matching is wearisome and time consuming when a large collection of documents is to be processed. Approaches for automatic registration such as area-based techniques using image patches and standard image similarity metrics have been proposed [35].

Wang et al. [34] proposed a two-stage hierarchical alignment technique that can efficiently and accurately align the two sides of documents in order to remove bleed-through. Their approach first coarsely aligns the images using a pair of anchors extracted from recto and verso images. The alignment is then fine-tuned using block matching and RBF-based interpolation techniques. The method is fully automated and runs significantly faster than other reported alignment methods.

Tonazzini et al. [30] proposed a system which includes a fully unsupervised registration method that can co-register any number of recto and verso channel maps of multispectral scans. They used pixel-based area methods such as Fourier-Mellin transforms and parameter optimization.

Bianco et al. [3] presented a procedure aimed at improving the readability of ancient degraded documents that includes recto-verso registration based on the Fourier-Mellin transform. Although this method has been demonstrated to be the most reliable and suitable one for their application, tests have shown its limits in the presence of deformations caused by folding or crumpling.

Wang and Tan [36] presented a non-rigid registration method for restoring manuscripts from bleed-through distortion. They make use of the gradient maps of images and writing patterns. To describe the registration transform, a mapping function consisting of a global affine and local B-splines is defined and then estimated by optimizing a cost function which takes into account image similarity and transformation smoothness.

The common link between these approaches and ours is that in the transmitted recto scans of our data set, the back-lighting process produces an effect which is similar to the bleed-through or show-through effect when the ink appears from the verso to the recto of the page.

3 Manuscripts and digitization

The predominant data set for this work is an unusual, comprehensively scanned, 19th century Sudanese edition of the Qur’ān [19]. Each of its 346 pages contains a watermark embedded in paper texture at the side margins in each sheet (see Fig. 1).

This manuscript is one of the most complex among a range of digitised materials [16] and is challenging for several reasons: its importance as a complete double-sided handwritten historical collection of the Qur’ān, paper sheets and writing on both recto and verso sides are thick, the background is not uniform, and the watermark patterns are not clear. Nevertheless, it does not suffer from bleed-through or show-through effects. The manuscript has wire watermarks with two shapes: double-headed eagle and moonface-within-shield designs (see Fig. 2).

As shown in Fig. 3, the digitised images normally consist of the paper (in the center) with a border region due to the lighting sheet during the digitization.

Beside the ‘Mahdiyya’ Copy of the Qur’ān, other data sets were experimented in this paper; their full description is available at [16]. These data sets are:

- An Islamic Prayer (Kitāb Durrat ‘iqd al-naḥr fī ‘asrār ḥizb al-baḥr). This manuscript consists of 32 folios; it includes the tre lune (three moons) watermark.
- The ‘West African’ copy of the Qur’ān. This copy consists of 332 folios, and also includes the tre lune watermark.

Figure 4 illustrates sample data taken from the Islamic Prayer, the tre lune watermark appears in the middle of the transmitted recto scan.

Precise details of data capture are given elsewhere [15]; each sheet delivers three images;

1. R_R : reflected recto – a scan of the page.
2. R_T : transmitted recto – a backlit version of the same page.
3. V_R : reflected verso – a scan of the other side of the page.



Fig. 1 Transmitted recto scan, enlarged to show watermark area (enhanced for display)

Each image is of resolution 3040×2160 in 24-bit RGB. R_R and R_T are coregistered perfectly, but V_R is not. R_T reveals features internal to the paper and inscribed on the verso. While there is a clear correspondence between features visible in V_R and (some) features visible in reflected form in R_T , various physical features of the paper such as folds or damage may cause these not to be exact.

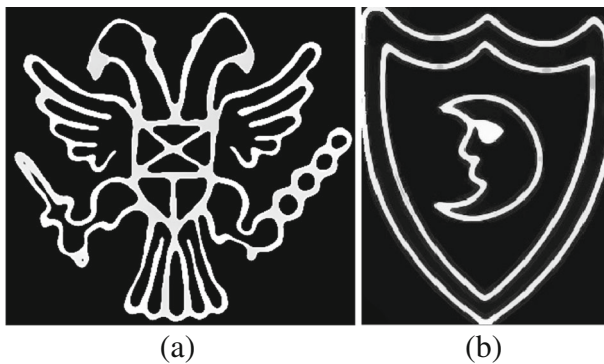


Fig. 2 Rough sketches of (a) the double-headed eagle watermark and (b) the moonface-within-shield countermark



Fig. 3 Sample page from the ‘Mahdiyya’ copy of the Qur’an, (a) reflected recto scan, (b) transmitted recto scan (c) reflected verso scan

4 Algorithm

In earlier work we have presented a model based on top-down approach to recto removal [5]. While this operates with considerable success, verso features remain that were then identified with a simple thresholding approach that was often too crude. Here, we present an improvement to this phase which firstly permits highly quality verso removal, and secondly permits an enhancement of the faint and often partial watermark features. An overview of the whole process is shown in Fig. 5: recto removal, backlit-verso registration, verso removal, image grouping (arithmetic mean) and watermark location.

4.1 Recto removal

We use the approach of Boyle and Hiary [5]. Reasonable assumptions permit a piecewise linear approximation to the back-lighting effect to be derived robustly. The algorithm partitions RGB data channels of R_R image into a number of clusters that contains pixels of a uniform intensity using k-means method which is controlled by some global parameter.

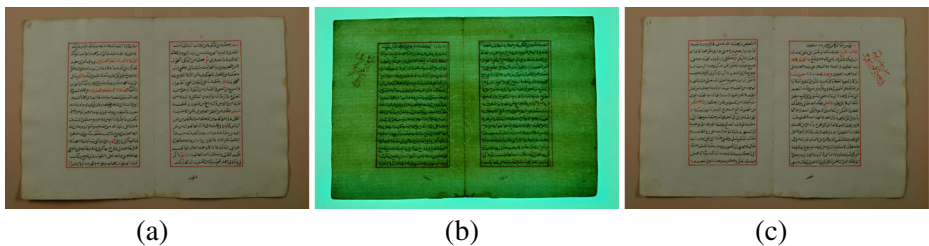


Fig. 4 Sample page from the Islamic Prayer, (a) reflected recto scan, (b) transmitted recto scan (c) reflected verso scan

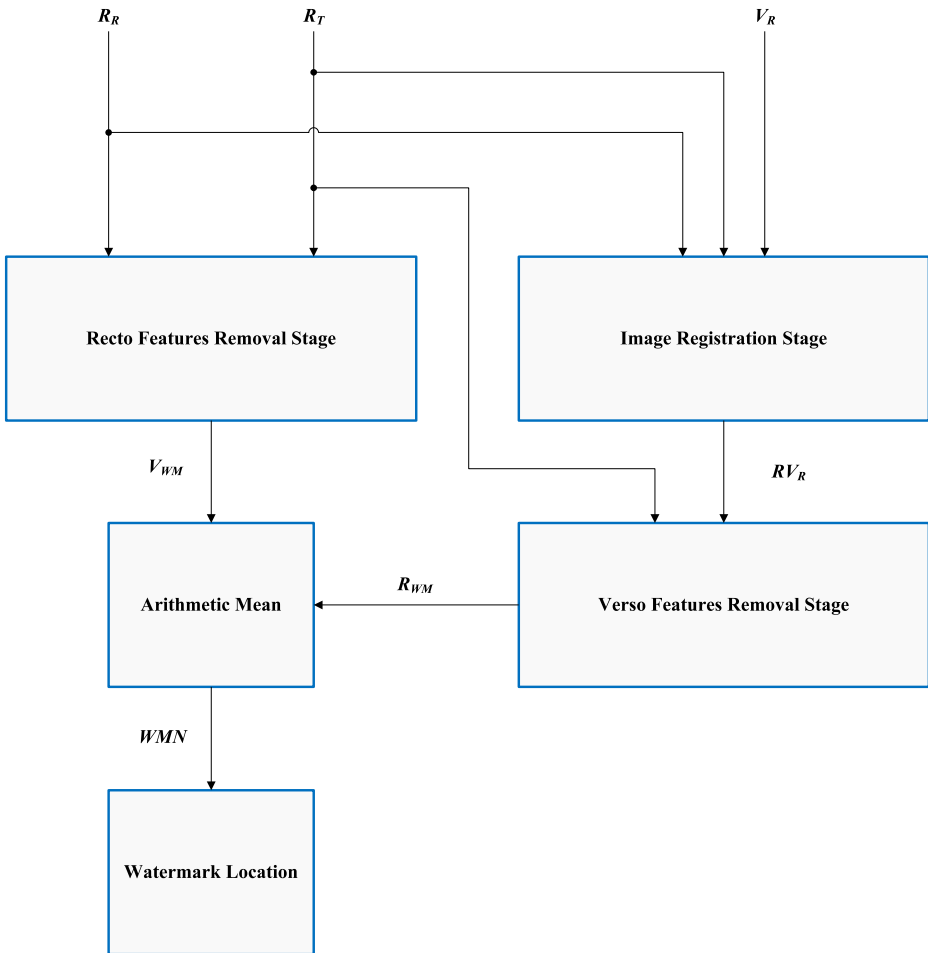


Fig. 5 Flow chart of watermark location: R , recto data; R_R , reflected recto image; V_R , reflected verso image; R_V_R , registered reflected verso image; V , verso data; R_{WM} , watermark and (parts of) recto information; V_{WM} , watermark and verso information; WMN , watermark data and noise

Then, for each cluster it computes a particular transform matrix (A) according to (1) that approximates the intensity effect of back-lighting; this linear transform provides a good approximation to the image \hat{R}_R that would result from back-lighting R_R in the absence of any verso, watermarking or other paper irregularity or feature: $D = (R_T - \hat{R}_R)$ then provides an image V_{WM} containing watermark and verso information;

$$A = [(\rho_p, \gamma_p, \beta_p) - (\mu_\rho, \mu_\gamma, \mu_\beta)] [(r_p, g_p, b_p) - (\mu_r, \mu_g, \mu_b)]^{-1} \tag{1}$$

where

$$\begin{aligned}
 (\mu_r, \mu_g, \mu_b) &= \text{mean}(r_p, g_p, b_p) : p \in R_R \\
 (\mu_\rho, \mu_\gamma, \mu_\beta) &= \text{mean}(\rho_p, \gamma_p, \beta_p) : p \in R_T
 \end{aligned}$$

and seeking a linear relationship for A

$$(\rho_p, \gamma_p, \beta_p) \approx A((r_p, g_p, b_p) - (\mu_r, \mu_g, \mu_b)) + (\mu_\rho, \mu_\gamma, \mu_\beta)$$

assuming that (r, g, b) is a vector in R_R , (ρ, γ, β) is a vector in R_T and p is pixels restricted to each cluster (not the whole image). The iterative refinement approach of (2) is applicable to each such cluster;

$$\hat{D} = \{p : |D_p| < T\}$$

$$A_{new} = [(\rho_p, \gamma_p, \beta_p) - (\mu_\rho, \mu_\gamma, \mu_\beta)][(r_p, g_p, b_p) - (\mu_r, \mu_g, \mu_b)]^{-1}, p \in \hat{D} \quad (2)$$

where D is an image in which pixels are given by the difference between their detected back-lit intensity (in R_T), and the intensity we might expect given the corresponding location in R_R . $|D_p|$ is a measure of the magnitude of the difference vector at p . This process is illustrated in Fig. 6, and an example output of the process is shown in Fig. 7.

4.2 Verso removal

The removal of recto inscription described in Section 4.1 is very successful, but leaves significant traces of the verso. We suggest that if a reflected copy of the verso scan, $\overline{V_R}$, is registered with R_T , then the process may be repeated to derive another image R_{WM} giving watermark and (parts of) recto information.

4.2.1 Registration

When the verso is scanned it is in a similar position only to the recto. It is reasonable to expect that as a first approximation an affine transform will align $\overline{V_R}$ with R_T . This we

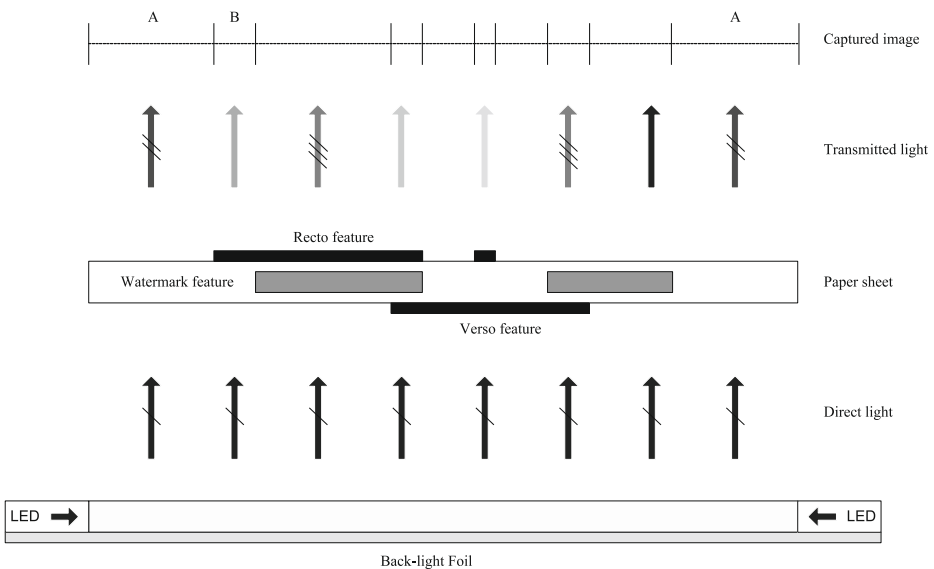


Fig. 6 The model of back-lighting. Low intensity is generally caused by a combination of recto and verso inked regions. Illumination from below is indicated by *up-arrows*, and the sensed image is at the top. *Vertical lines* along the image indicate points at which the received signal may change: for example, at A, blank featureless paper is being detected; at B recto data inscribed on paper is detected [5]

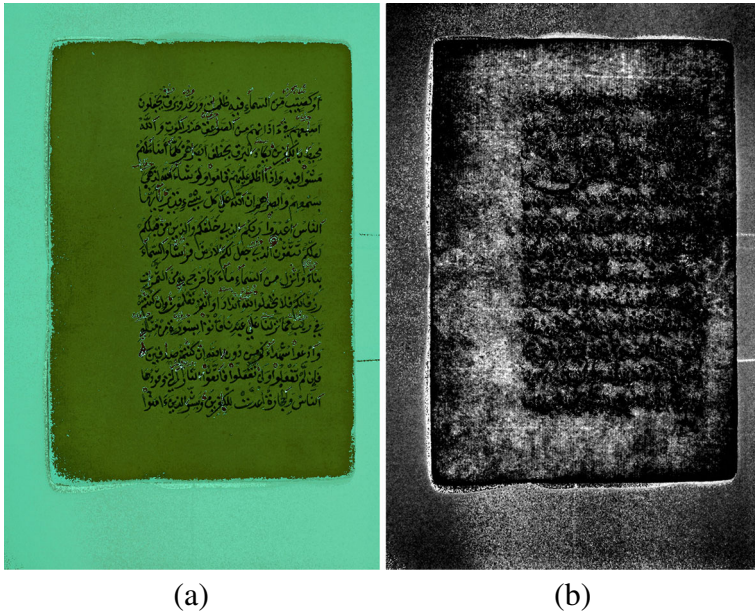


Fig. 7 (a) \hat{R}_R – an image derived from R_R that simulates back-lighting, (b) the result of a differencing R_T and \hat{R}_R (enhanced for display)

perform by binarising $\overline{V_R}$ (to V_R^B) and R_T (to R_T^B) to extract stroke (and other information) – we would then expect foreground in R_T^B to be a superset of that in V_R^B .

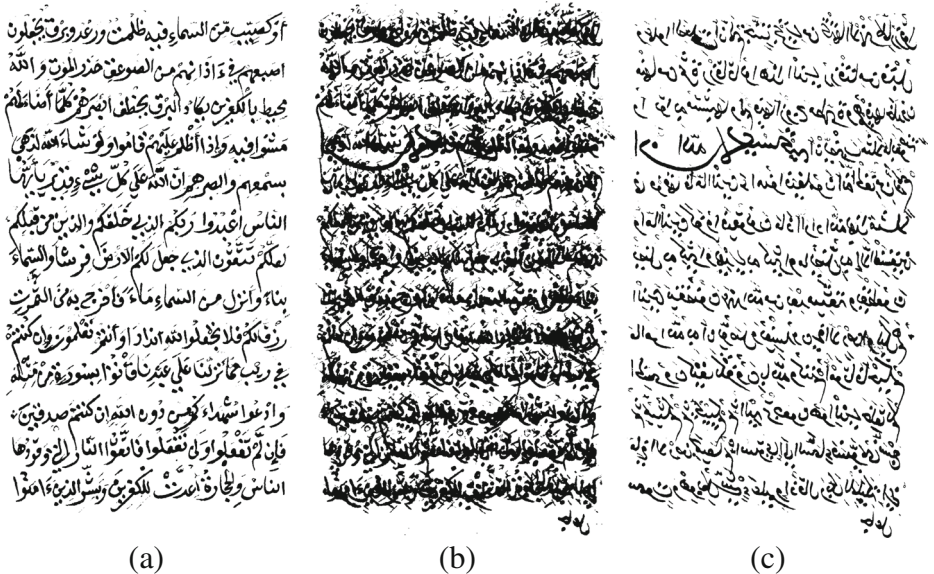
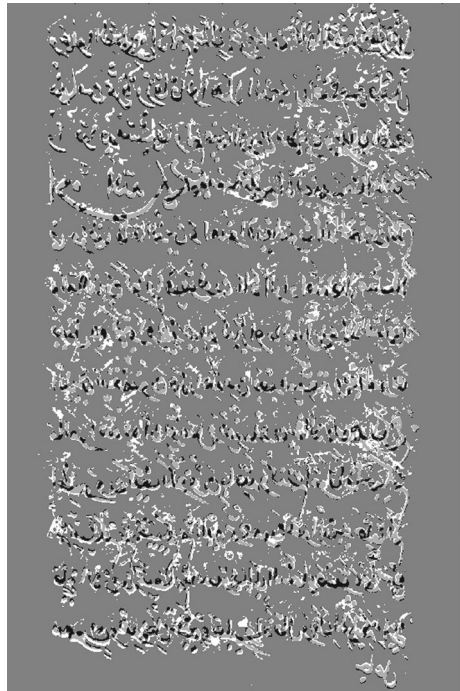


Fig. 8 Binarised versions of (a) R_R (R_R^B), (b) R_T (R_T^B) and (c) V_R (V_R^B)

Fig. 9 Example of an alignment; source and target images are illustrated in *black* and *white*



We perform the binarising process by selecting the red channel, which provides maximal contrast, contrast stretching, then applying Otsu's global image thresholding method [23]. The results of this process are illustrated in Fig. 8.

Now R_T^B will consist of features attributable to the recto, verso, watermark, other marks of manufacture, damage, noise etc. We might expect $R_T^B \oplus R_R^B$ then to be composed of parts of verso and features attributable to paper. We compute this image and subject it to simple morphological opening which has the useful effect of removing residual recto features. It is the output of this operation to which we seek to register \overline{V}_R^B .

We seek a simple affine transform to do this registration – this will have 2 parameters (translation and rotation) since there will be no detectable change in scale. For a matching metric we use normalised 2-D cross correlation [20], and the approach of Wolberg [37].

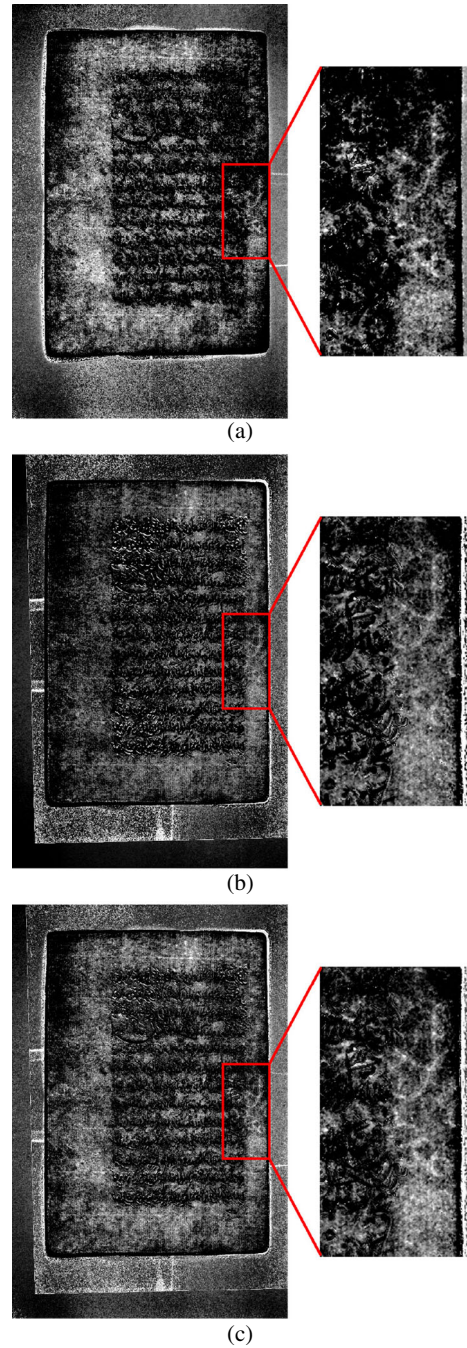
4.2.2 Verso feature removal

The image \overline{V}_R^B is transformed (using NN interpolation) to bring it into alignment with the R_T scan. We then reuse the approach of Boyle and Hiary [5] to derive the difference image R_{WM} . It is now possible to average the images V_{WM} and R_{WM} to improve the SNR of the watermark (and other paper effect) signal. This image is subjected to the existing technique [5, 15] to identify occurrences of known watermarks.

5 Results and discussion

The system has been implemented in MATLAB and tested against all of the three data sets described in Section 3. A total of over 700 pages of various watermarks were processed.

Fig. 10 (a) Zoomed output of recto removal stage, (b) zoomed output of verso removal stage, (c) zoomed output of averaging



Efficacy of the registration procedure was evaluated against a number of manual registrations. In the absence of folding, crumpling or similar effects, we might expect it to be of high quality, especially as the parameter search can be initialised against a ‘guess’ based on

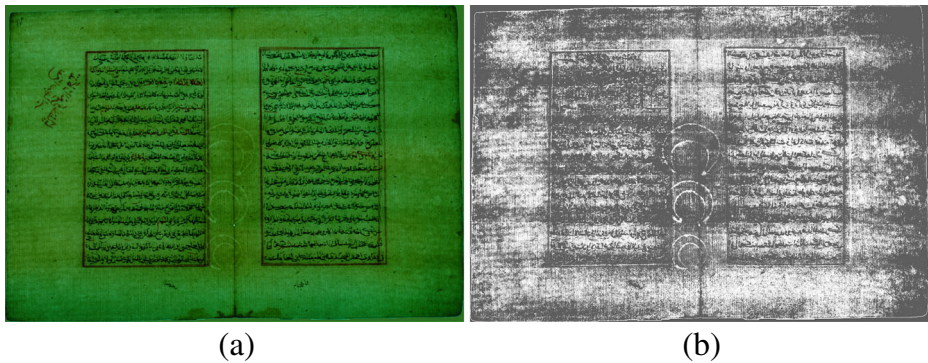


Fig. 11 (a) Transmitted recto scan, (b) output of averaging recto and verso stages (enhanced for display)

paper boundaries. In fact, a number of inaccuracies are evident: this is illustrated for one page in Fig. 9, indicating that there is scope for deriving a non-rigid transformation. This problem has been seen before [3].

With respect to paper watermark location, qualitative comparison can be performed by inspecting the images V_{WM} , R_{WM} and their average. An example is shown in Fig.10, in which the output of the enhanced algorithm is seen to be an appreciable improvement.

Another sample result of from the Islamic Prayer data set is shown in Fig. 11, which shows the output (enhanced for display) of averaging recto and verso stages of the example from Fig. 4, the tre lune watermark appears clearly in the middle.

Table 1 shows retrieval results of four design parts: a double-headed eagle watermark ‘E’, and a moonface-within-shield countermark ‘M’ used in the ‘Mahdiyya’ copy of the Qur’ān. Results represents an improvement on earlier results [5].

As can be seen from Table 1, the proposed approach outperforms that presented in [5]: this improvement is due to successful manipulation of verso features. As for the other two data sets (Islamic Prayer and ‘West African’ copy of the Qur’ān), the proposed approach achieved 100 % matching accuracy, showing that it is suitable for various data sets of different watermarks and paper structure.

Table 1 Matching results for different watermark shapes (%)

Watermark	True positive		True negative	
	Proposed work	Previous work	Proposed work	Previous work
M upper	99.0	98.8	100	100
M lower	98.7	97.7	100	100
E upper	97.8	96.5	100	100
E lower	96.0	94.3	100	100

Each shape is divided into two due to the cutting of paper

6 Conclusions

This paper has presented an enhancement of a watermark extraction algorithm which exploits knowledge of a verso scan. We have implemented a registration technique that, while simple in overlooking crumple and fold effects, allows reuse of the earlier technique to derive a second watermark signal. When this is averaged with the first, an improved algorithm results.

Experiments on three different data sets have shown encouraging results; further improvements could be made by exploiting more non-rigid registration methods. Further, the model of back-lighting is [piecewise] linear, and it is possible that a more sophisticated model would generate improved outputs.

Acknowledgments Thanks are due to the University of Leeds for making the manuscripts available online [16]. We also acknowledge the interest and support of Prof. Roger Boyle and Dr. Kia Ng (Interdisciplinary Centre for Scientific Research in Music, University of Leeds).

References

1. Aken J (2003) An improvement in Grenz radiography of paper to record watermarks, chain and laid lines. *Stud Conserv* 48(2):103–110
2. Ash NE (1982) Recording watermarks by beta-radiography and other means. In: *The Book and Paper Group Annual*, vol 1. American Institute for Conservation of Historic and Artistic Works
3. Bianco G, Bruno F, Tonazzini A, Salerno E, Console E (2009) Recto-verso registration, enhancement and segmentation of ancient documents. *Proc. IEEE 15th International Conference on Virtual Systems and Multimedia*, Vienna, Austria, pp 131–136
4. Biermann C (1996) *Handbook of pulping and papermaking*, 2nd edn. Academic Press, San Diego, California, USA, p 171
5. Boyle R, Hiary H (2009) Watermark location via back-lighting and recto removal. *Int J Doc Anal Recognit* 12(1):33–46
6. Bridgman CF, Keck S, Sherwood HF (1958) The radiography of panel paintings by electron emission. *Stud Conserv* 3(4):175–182
7. Bridgman CF (1965) Radiography of paper. *Stud Conserv* 10(1):8–17
8. Edge D (2001) The digital imaging of watermarks. *Comput Musicol* 12:261–274
9. Gonzalez R, Woods R (2008) *Digital image processing*, 3rd edn. Prentice-Hall, Upper Saddle River, NJ, USA
10. Gonzalez R, Woods R, Eddins S (2009) *Digital image processing using MATLAB*, 2nd edn. Gatesmark Publishing, USA
11. Goshtasby A (2005) *2-D and 3-D Image registration for medical, remote sensing, and industrial applications*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA
12. Gravell TL (1990) The wizard of watermarks. *Du Pont Mag* 84(1):4–6
13. Haupt W (1981) Wasserzeichenwiedergabe in schwierigen fallen. *Maltechnik-Restaur* 87(1):38–43
14. Heawood E (1950) Watermarks mainly of the 17th and 18th centuries. Paper Publications Society, Hilversum
15. Hiary H (2008) Paper-based watermark extraction with image processing. Dissertation. University of Leeds
16. Hiary H, Boyle R, Ng K Digitised Arabic texts from the University of Leeds. <http://www.comp.leeds.ac.uk/arabictxts/>. Accessed 8 February 2015
17. Hiary H, Ng K (2007) A System for segmenting and extracting paper-based watermark designs. *Int J Digit Libr* 6(4):351–361
18. Hunter D (1978) *Papermaking: The History and Technique of an Ancient Craft*. Dover Publications, New York, USA
19. Special Collections at Leeds University Library. <http://library.leeds.ac.uk/special-collections>. Accessed 8 February 2015
20. Lewis J (1995) Fast normalized cross-correlation. *Proc. Vision Interface*, Québec, Canada, pp 120–123
21. Moschini D (2000) La marca d'acqua: A system for the digital recording of watermarks. In: *Puzzles in Paper: Concepts in Historical Watermarks* (Papers from the 1996 International

- Conference on Watermarks at Roanoke, Virginia). Oak Knoll Books and the British Library, Newcastle, pp 187–192
22. Neuheuser H, Märgner V, Meinschmidt P (2005) Wasserzeichendarstellung mit Hilfe der Thermographie. *ABI TECHNIK* 25(4):266–278
 23. Otsu N (1979) A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 9(1):62–66
 24. Schoonover D (1987) Techniques of reproducing watermarks: a practical introduction. In: *Essays in Paper Analysis*. Folger Shakespeare Library, Washington, pp 154–167
 25. Sharma G (2001) Show-Through Cancellation in Scans of Duplex Printed Documents. *IEEE Trans Image Process* 10(5):736–754
 26. Small CA (2000) Phosphorescence watermark imaging. In: *Puzzles in Paper: Concepts in Historical Watermarks* (Papers from the 1996 International Conference on Watermarks at Roanoke, Virginia). Oak Knoll Books and the British Library, Newcastle, pp 169–181
 27. Stevenson A (1951) Watermarks are twins. *Stud Bibliogr* 4:57–91. Accessed 8 February 2015. <http://www.bsuva.org>
 28. Stewart D, Scharf R, Arney J (1995) Techniques for digital image capture of watermarks. *J Imaging Sci Technol* 39(3):261–267
 29. Tonazzini A, Bedini L, Salerno E (2004) Independent component analysis for document restoration. *Int J Doc Anal Recognit* 7(1):17–27
 30. Tonazzini A, Bianco G, Salerno E (2009) Registration and enhancement of double sided degraded manuscripts acquired in multispectral modality. *Proc. IEEE 10th International Conference on Document Analysis and Recognition, Barcelona, Spain*, pp 546–550
 31. Tomimasu H, Kim D, Suk M, Luner P (1991) Comparison of four paper imaging techniques: beta-radiography, electrography, light transmission, and soft X-radiography. *Tappi* 74(7):165–176
 32. Tonazzini A, Salerno E, Mochi M, Bedini L (2004) Blind Source Separation Techniques for Detecting Hidden Texts and Textures in Document Images. *Lect Notes Comput Sci* 3212:241–248. Springer Berlin Heidelberg
 33. Tonazzini A, Salerno E, Bedini L (2006) Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique. *Int J Doc Anal Recognit* 10(1):17–25
 34. Wang J, Brown M, Tan C (2008) Accurate alignment of double-sided manuscripts for bleed-through removal. *Proc. IEEE 8th IAPR International Workshop on Document Analysis Systems (DAS'8)*, Nara, Japan, pp 69–75
 35. Wang J, Brown M, Tan C (2009) Automatic corresponding control points selection for historical document image registration. *Proc. IEEE 10th International Conference on Document Analysis and Recognition, Barcelona, Spain*, pp 1176–1180
 36. Wang J, Tan C (2010) Non-rigid image registration for historical manuscript restoration. *Proc. IEEE 20th International Conference on Pattern Recognition, Istanbul, Turkey*, pp 2274–2277
 37. Wolberg G, Zokai S (2000) Robust image registration using log-polar transform. *Proc. IEEE International Conference on Image Processing, Vancouver, BC, Canada, vol 1*, pp 493–496
 38. Zitova B, Flusser J (2003) Image registration methods: a survey. *Image Vis Comput* 21(11):997–1000



Jamal Said received his B.Sc. and M.Sc. degrees in Computer Science from the University of Jordan (Amman, Jordan) in 2006 and 2011, respectively. His research interests include Digital Image Processing and Paper Watermarks.



Hazem Hiary received his PhD in Computer Science / Image Processing from the University of Leeds (Leeds, UK) in 2008, B.Sc. and M.Sc. degrees in Computer Science from the University of Jordan (Jordan) in 2001 and 2003 respectively. He is currently an Associate Professor in the Computer Science Department at the University of Jordan. His research interests include Image Processing, Pattern Recognition, Paper Watermarks, Document Analysis and Recognition, Audio Processing and Data Hiding.