Recto-verso registration, enhancement and segmentation of ancient documents

G. Bianco, F. Bruno
Dipartimento di Meccanica
Università della Calabria
Rende (CS), Italy
gianfranco.bianco@unical.it
f.bruno@unical.it

A. Tonazzini, E. Salerno
ISTI
Consiglio Nazionale delle Ricerche
Pisa, Italy
anna.tonazzini@isti.cnr.it
emanuele.salerno@isti.cnr.it

E. Console
Tea s.a.s.
Catanzaro, Italy
elena@teacz.com

***Abstract*—This paper presents some experiments on different image processing techniques combined and adapted to define a procedure aimed at improving the readability of ancient degraded documents. The focus is on documents affected by the bleed-through effect, which significantly reduces readability. The availability of the recto and verso scans of any single page is assumed. The proposed procedure is composed of three steps: recto-verso registration, image enhancement through statistical decorrelation and, finally, image segmentation. As the registration step can strongly influence the results obtained during the subsequent phases, we carried out a comparison of different registration algorithms, in order to evaluate their overall effectiveness on recto-verso processing.**

Keywords: image enhancement, registration, binarization, ancient documents.

I. INTRODUCTION

The virtual restoration of ancient documents is being more and more acknowledged as a valid application of information technologies in the field of cultural heritage. It can serve various purposes: in some cases, it is necessary to restore the original aspect of the document as it appeared before the deterioration, without causing losses of important pieces of information. In other cases, the priority goes to the improvement of readability, so that altering the aspect of the document can be allowed. This happens, for example, when the document must be transcribed, or processed by an OCR device. It is apparent, however, that these two goals are intimately related, since virtual restoration is often a preliminary step to improve or even make possible an effective segmentation.

There are many scientific works concerning the application of image processing algorithms for processing ancient documents. Some of the techniques dedicated to virtual restoration of ancient documents aim to correct the problem of bleed-through, i.e. when the ink transudes from the back side (*verso*) to the front side (*recto*) of the page, thus creating an interference pattern that affects readability. Removing these interferences is not trivial, since their intensity is often close to the one of the main text. On the basis of a single-side scan, the techniques attempted to reduce the interference are based on adaptive thresholding or segmentation-classification [1], [2], [3], [4]. Especially for gray-level images, techniques that exploit both sides of a page promise to perform better [5], [6]. Any such technique must rely on a preliminary registration of the two sides. Along this line, this paper proposes and tests a composite procedure to improve the readability of recto-verso pairs of documents affected by bleed-through, starting from the raw grayscale scans. This procedure consists in cascading three steps: recto-verso registration, image enhancement, and image segmentation.

Registration of recto-verso images is a very challenging task, since the recto and verso sides of a page are usually different not only with respect to image intensity, but also topographically. Moreover, the common features on which registration can be based are often very sparse: a correspondence between pixels can only be identified in the restricted regions where a pattern present in one side also appears in the opposite side. We implemented and tested some image registration algorithms recently proposed in the literature [7], in order to choose the most suitable for this application.

Among the many image enhancement algorithms available, we focused on multiview image decorrelation techniques. These have been previously proposed to separate background artifacts from foreground text in documents, by exploiting multiple observations. In particular, our specific technique is based on blind source separation [8], and uses symmetric orthogonalization algorithms to decorrelate the two sides.

As far as segmentation is concerned, we implemented and tested the algorithm proposed by Sauvola and Pietaksinen [9], since it is reputed one of the most efficient segmentation algorithms to be applied to ancient documents [10].

All the algorithms have been implemented in a Matlab environment; a graphic interface to test their performances has also been realized.

II. RELATED WORK

In this section, we briefly describe the three image processing techniques used in our procedure.

A. Image Registration

Image registration consists in aligning two or more images of the same scene, acquired under variable optical systems, viewpoints, or sensors. The process finds a correspondence between a "reference image" and the so-called "sensed image",

This work has been supported by the Project "Restarts" (www.restarts.org), financed by the Program POR Calabria 2000-2006 Misura 3.16.d.

and to overlap the two images through a geometric transformation [7].

A first classification of the registration techniques provides two types of methods:

- Area-based methods: use of the whole image to estimate the transformation parameters
- Feature-based methods: use of extracted features to estimate the registration parameters.

A further distinction concerns the way of extracting the information needed to compute the registration parameters (manual, semi-automatic and automatic methods).

Registration is normally needed for both multichannel images of a single side of a document and for recto-verso images. Registering single-side images acquired by multispectral cameras is often necessary to deal with factors such as acquisition by optical filters with different refractive indices, filter changing, autofocusing, operator errors, etc. The images acquired by recto-verso scanning, in turn, are normally misaligned since it is impossible to position the front and back sides exactly in the same way. Additional distortions that can cause misalignments are possibly due to page folding.

In this study, we assume that the images are acquired by the same sensor, and in absence of page folding. This hypothesis allows us to align recto and verso scans through an affine transformation. Many works can be found in the literature that deal with image registration, but very few of them are specifically devoted to recto-verso documents [11], [6].

Our work is focused on the area-based automatic methods, which offer the notable advantage of limiting the user's intervention. In particular, we have been analyzing the cross-correlation-based methods, which do not require particular features to be extracted, but make use of reduced image windows (templates) to compute the transformation parameters. Among the methods we have implemented in our study, one computes the normalized cross-correlation in the spatial domain [12]. Another one computes the normalized cross-correlation in the frequency domain [13], but is only able to evaluate translation parameters. Another frequency-domain method we implemented, the Fourier-Mellin method [14], allows us to compute scale, rotation and translation parameters. The translation parameters are first computed, then the rotation and scaling parameters are computed by phase correlation in the log-polar space [15].

We also tested the method described in [6], specifically designed for double-sided document images. This method is based on the optimization of the parameters of an affine transformation, computed by minimizing the sum of the squared differences between corresponding pixels in the recto and verso sides, respectively. We experimented all these four methods on recto-verso images of documents affected by bleed-through to choose the most suitable for our application.

B. Multiview Image Decorrelation

A specific research line for artifact reduction in ancient documents focuses on applying blind source separation (BSS) algorithms, viewing the foreground text and the interferences as individual patterns that overlap in the document, and relying on multiple observations for their separation. In [16], a linear, instantaneous mixture model for the data is proposed, and an independent component analysis strategy is exploited to analyze multispectral single-side scans with several overlapping information layers. In [17], this approach is extended to reduce the bleed-through interference in registered recto-verso grayscale scans. For the peculiar symmetries characterizing the mixing operator in this case, decorrelation methods such as whitening or symmetric orthogonalization seem to be particularly suitable and effective for artifact reduction, and have already given promising results through fast and efficient algorithms.

In synthesis, let $\mathbf{x}(x,y)$ be the vector image containing the recto and verso scans of the document. The size of this vector ranges from 2, for grayscale recto-verso scans, to $2n$, for the cases where each side has been acquired by an n-channel multispectral sensor. Each component in $\mathbf{x}$ can be thought of as a linear combination of the interference-free recto and verso patterns at a specific channel. For each channel, we have two maps representing the recto and verso appearances. Let the 2-vector map $\mathbf{x}^{(i)}(x,y)$ be the pair of maps acquired at the i-th channel, and $\mathbf{s}^{(i)}(x,y)$ be the pair of the interference-free patterns at the same channel. As detailed in [17], an estimate of $\mathbf{s}^{(i)}(x,y)$ can be obtained by decorrelating the components of $\mathbf{x}^{(i)}(x,y)$, that is

$$\hat{\mathbf{s}}^{(i)} = \mathbf{V}_{\mathbf{x}^{(i)}} \Lambda_{\mathbf{x}^{(i)}}^{-\frac{1}{2}} \mathbf{V}_{\mathbf{x}^{(i)}}^{T} \mathbf{x}^{(i)} \qquad (1)$$

where $V_x^{(i)}$ is the matrix of the eigenvectors of the covariance matrix of $\mathbf{x}^{(i)}$, and $\Lambda_x^{(i)}$ is the corresponding matrix of eigenvalues. We have shown in [17] that this is only one of the possible choices to decorrelate the components of $\mathbf{x}^{(i)}(x,y)$, corresponding to applying a linear self-adjoint operator to the data vector.

C. Segmentation

Image segmentation techniques are usually applied to grayscale images, and aim at individuating a threshold grey value, which allows to separate the textured background from the foreground text, i.e. to isolate the written words from the artifacts. The output of segmentation is then a binary image, where text pixels and background pixels are labeled. The segmentation techniques can be of two types: global and local. The global techniques are applied to the whole image without distinction, while the local ones apply different types of thresholding, according to the various regions of the image.

In [10], a performance analysis on several thresholding algorithms is reported. A specific test related to the separation of the text from the background was carried out using 40 documents with different fonts, sizes and typefaces. The document degradations were simulated by blur and speckle noise, as suggested by the degradation models proposed in [18]. Among the different local-based methods examined, the one that showed the best performance for document binarization was the method proposed in [9] by Sauvola and Pietaksinen.

In this algorithm, a threshold $t(x,y)$ is computed for each pixel, which depends on two quantities, $m(x,y)$ and $s(x,y)$, representing the local mean and standard deviation, respectively, calculated on a square window 10-20 pixels wide:

$$t(x,y) = m(x,y)\left[1 + k\left(\frac{s(x,y)}{R} - 1\right)\right] \qquad (2)$$

where R is the maximum value of the standard deviation (R=128 for a grayscale document), and k is a parameter that takes positive values in the range [0.2, 0.5]. Using this formula, the threshold is calculated according to the contrast in the area surrounding each pixel. When the contrast in a small neighborhood of the pixel is high, $s(x,y) \approx R$ and $t(x,y) \approx m(x,y)$. Conversely, when the contrast is low, $t(x, y)$ becomes smaller than the mean value, thus allowing the relatively dark regions in the background to be removed successfully. Parameter k controls the value of the threshold in the local window so that the higher the value of k, the smaller is the threshold. A value of k = 0.5 is suggested in [9], but some experiments reported in [19] put in evidence that k = 0.34 gives better results. Anyway, the algorithm is not very sensitive to the value of k used.

III. Procedure Description

The procedure we describe here can be applied for two different purposes:

a) Virtual restoration: the goal is to bring the document back to its original conditions, eliminating the degradations as much as possible; as already mentioned, image enhancement based on BSS techniques, in combination with multiple scans, have been successfully proposed to this end;

b) Readability improvement: the target here is to improve readability, even altering the appearance of the document; for a wide range of degradations, this can be obtained by just segmenting each side individually.

However, it is easy to see that these two seemingly different goals are intimately related. Indeed, it is obvious that the higher is the contrast between the text and the disturbances, the more the segmentation is effective. Our experiments have highlighted that, in presence of a strong bleed-through, the segmentation algorithms are not very effective, since there is not a sufficient contrast between the text in the recto and the strokes interfering from the verso. The procedure we propose assumes that, in such cases, the data can be pre-processed by decorrelation algorithms, which allow the interferences to be reduced. In this way, the performance of the segmentation algorithm is improved.

The first step of our procedure (Fig. 1) is acquisition, that is, the scanning of the two document sides. After this, the goal of the procedure must be defined. If a virtual restoration is required, a registration of the available observations and a subsequent decorrelation are carried out. Conversely, if an improvement of readability is required, it is necessary to evaluate if, in each side, there is enough contrast between the main text and the bleed-through pattern. In this case, segmentation can be directly performed on each side individually. Otherwise, a registration and a decorrelation of the two sides must be executed before segmentation. If the segmentation is not satisfactory, the procedure allows us to go back to registration and decorrelation, and then to repeat the segmentation.

It is worth noting that registration and decorrelation, that is, the image enhancement part of the procedure, can also be applied to RGB or multispectral acquisitions, thus preserving the information available on the document in its entirety [20]. Briefly, in the RGB case, the three pairs of corresponding channels in the recto and verso images are separately aligned and decorrelated, and then, for each side, the three processed RGB components are recomposed, to give the restored document sides in color. This is an interesting feature when virtual restoration is addressed. However, when readability improvement is the goal, since the subsequent segmentation step is typically applied to grayscale images, it is necessary to perform an RGB merge or just choose a single channel. Doing this before or after registration and decorrelation is a matter of choice and of computational cost.

We implemented the procedure in a Matlab environment and integrated it in a graphic interface that assists the user in the elaboration (Fig. 2).

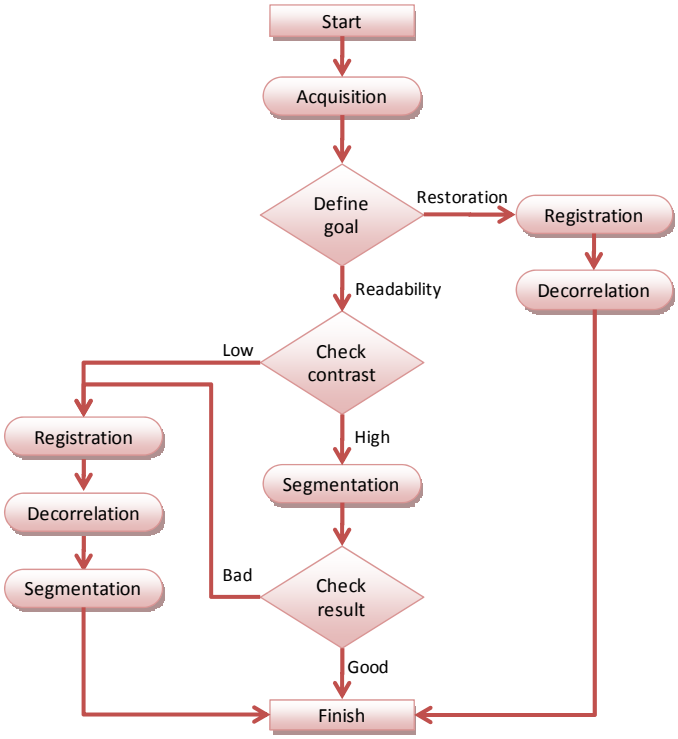


Figure 1. A flow chart representation of the proposed procedure

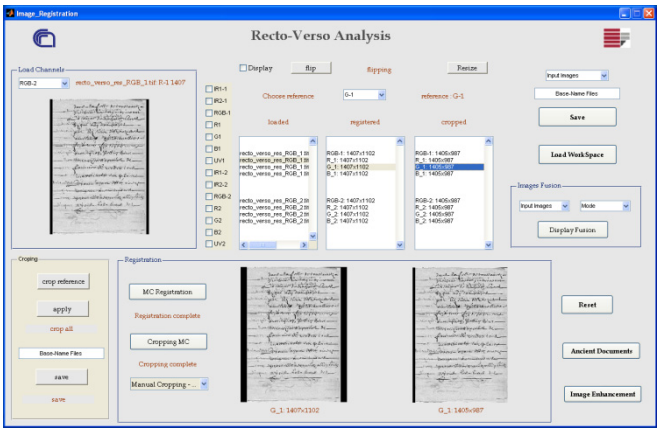


Figure 2. The graphical interface of the testing environment implemented in Matlab

IV. PROCEDURE EXPERIMENTATION

In this section, the effectiveness of the whole procedure is analyzed experimentally.

We conducted some preliminary tests on the most critical step, i.e. registration, to compare the performances of two of the algorithms implemented, namely, the Fourier-Mellin technique, and the method based on parameter optimization. We also tested the two methods based on cross-correlation, but they failed in the recto-verso registration. These algorithms, in fact, are not even able to yield a solution on recto-verso pairs, because of the significant differences between the two sides.

The Normalized Root-Mean-Square Error (NRMSE) of the registered pair was chosen to evaluate quantitatively the results of the registration [13]:

$$E^2 = 1 - \frac{\max_{u,v} |r_{fg}(u,v)|^2}{\sum_{x,y} |f(x,y)|^2 \sum_{x,y} |g(x,y)|^2} \qquad (3)$$

where $f(x,y)$ and $g(x,y)$ are the registered recto and verso images, respectively, and $r_{fg}(u,v)$ is their correlation coefficient.

The error values obtained by the two methods considered for three representative examples are reported in Table I. It can be observed that the two algorithms exhibit very similar behaviors. However, in order to test the whole procedure, we chose to employ the Fourier-Mellin method for its reliability in all the recto-verso registration experiments we carried out. Although we have been carrying out our experiments on both grayscale and color or multispectral images, since our present goal is to maximize readability, we only put our emphasis on the results that are not specific of color processing.

TABLE I. NORMALIZED MEAN-SQUARE ERROR OBTAINED USING TWO DIFFERENT ALGORITHMS ON 3 CASES

Method / Case	Fourier-Mellin	Parameter optimization
RV 1	0.1782	0.1487
RV 2	0.1846	0.1765
RV 3	0.2939	0.3095

The first experiment has been conducted on the recto-verso pair shown in Fig. 3.

An expert eye can easily appreciate the good contrast between the recto and the verso patterns on each side of the document. This could make us assume that the segmentation algorithm can produce acceptable results; however, we chose to verify the efficacy of the segmentation algorithm by applying it both on the original images and on the pre-processed ones. To carry out this test, the verso image must be horizontally flipped and registered with respect to the recto channel. Fig. 3 illustrates the registration process. In the top row, the original recto and verso scans are shown. In the middle row, the flipped verso overlaps the recto in transparency. The pair is depicted by two different colors (blue for the recto and red for the verso) to highlight the differences between the two images. It can be noticed that the two overlapped images look translated horizontally. Finally, at the bottom of Fig. 3, the overlapped registered pair is shown. It is apparent that the registration produced good results. After registration, a symmetric orthogonalization can be applied, which reduces the degradations and already provides some improvement in readability (Fig. 4).

The output of the orthogonalization is then input to the Sauvola algorithm in order to segment the main text; the results are compared with the output of the Sauvola algorithm as applied directly to the original images (Table II). It is apparent that the segmentation applied to the orthogonalized images and to the original ones produces similar results, since in this case the contrast between foreground and bleed-through are already good in the original, and the two patterns do not interfere so much.

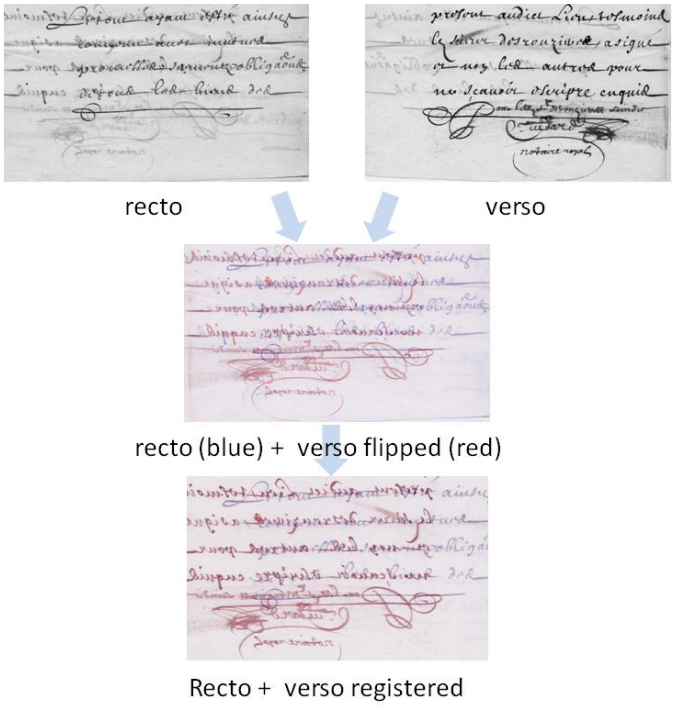


Figure 3. Flipping and registering the recto-verso pair

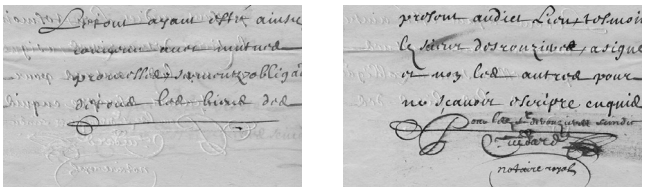
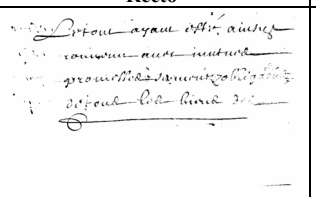
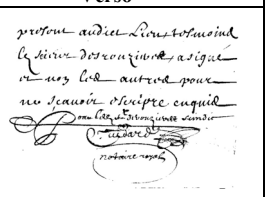
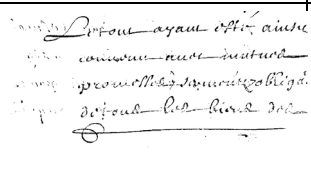
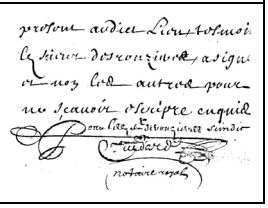


Figure 4. Output of the decorrelation algorithm

TABLE II. RESULT OF THE SEGMENTATION PERFORMED ON THE ORIGINAL AND ORTHOGONALIZED IMAGES OF THE FIRST EXPERIMENT

	Recto	Verso
Originals		
Orthogonalized		

The second case study refers to a recto-verso document affected by a stronger bleed-through (Fig. 5). The segmentation algorithm in this case does not allow for a good separation of the foreground pattern (Fig. 6): indeed, even varying the parameters of the Sauvola algorithm, it is not possible to remove the bleed-through without eliminating also the foreground.

Binarization was then applied to the decorrelated recto-verso pair. At the bottom of Fig. 6 a comparison between the results in the two cases is shown. We overlapped the segmentation outputs from the original and the decorrelated images, respectively, and highlighted in green the pixels that are only present in the former image. This representation shows that there is more bleed-through in the first output than in the second. It is thus apparent that, in this case, the preliminary application of enhancement techniques produces some improvement.

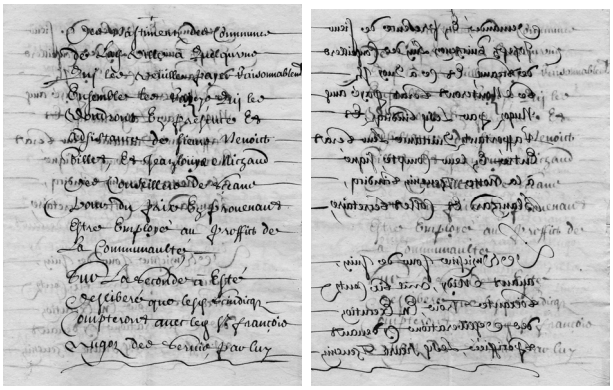


Figure 5. Original recto (left) and verso (right) images for the second test case

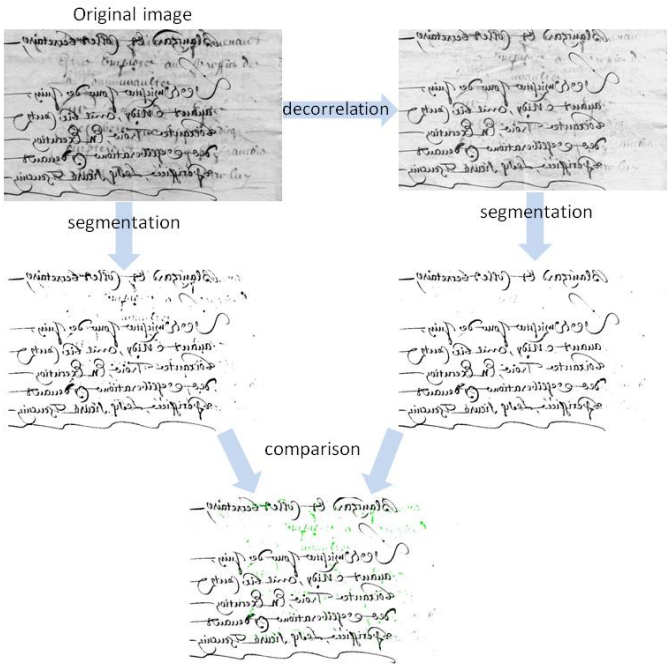


Figure 6. Comparison between the results of the second experiment

A third experiment (Fig. 7) has been conducted on another document with poor contrast between the writings and the interfering patterns. This last case study confirmed the effectiveness of the decorrelation technique to improve the results of the segmentation.

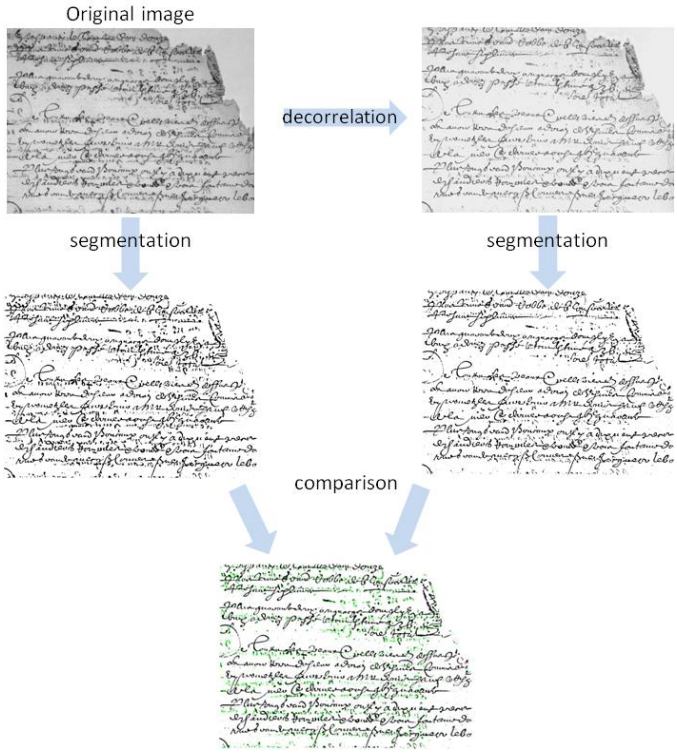


Figure 7. Comparison between the results of the third experiment

V. CONCLUSION

This paper presents a procedure to guide the user in the application of image processing techniques to improve the readability of ancient documents. This procedure includes registration, enhancement and segmentation, and is specifically oriented to the enhancement of recto-verso pairs affected by bleed-through.

From the results presented, it is apparent that decorrelation is effective as a pre-processing technique to make segmentation more reliable, especially when there is a poor contrast between the writings and the interfering patterns.

This paper also presents an effective application of the Fourier-Mellin transform for the difficult problem of registering recto-verso scans of ancient documents. The resulting registration technique has been included in our overall procedure after some tests, partially reported in Section IV, that compare the performances of various general purpose registration algorithms on recto-verso images.

ACKNOWLEDGMENT

The authors are indebted to Prof. Massoud Babaie-Zadeh for sharing some of his test images.

REFERENCES

[1] H. Nishida and T. Suzuki, "A Multiscale Approach to Restoring Scanned Color Document Images with Show Through Effects", Proc. ICDAR 2003, Edinburgh, Scotland, 2003.

[2] Q. Wang, T. Xia, L. Li, and C.L. Tan, "Document Image Enhancement Using Directional Wavelet", Proc. IEEE Conf. Computer Vision Pattern Recognition, vol. 2, 2003, pp. 534-539.

[3] Y. Leydier, F. Le Bourgeois, and H. Emptoz, "Serialized unsupervised classifier for adaptive color image segmentation: application to digitized ancient manuscripts", Proc. Int. Conf. on Pattern Recognition, 2004, pp. 494-497.

[4] F. Drida, F. Le Bourgeois, and H. Emptoz, "Restoring ink bleed-through degraded document images using a recursive unsupervised classification technique", Proc. 7th Workshop on Document Analysis Systems, 2006, pp. 38-49.

[5] G. Sharma, "Show-through cancellation in scans of duplex printed documents", IEEE Trans. on Image Processing, vol. 10, pp. 736-754, 2001.

[6] E. Dubois and A. Pathak, "Reduction of bleed-through in scanned manuscript documents", Proc. IS&T Image Processing, Image Quality, Image Capture Systems Conference, 2001, pp. 177-180.

[7] B. Zitova and J. Flusser, "Image registration methods: A survey", Image and Vision Computing, vol. 21, pp. 997-1000, 2003.

[8] A. Cichocki and S. Amari, Adaptive Blind Signal and Image Processing. New York: Wiley, 2002.

[9] J. Sauvola and M. Pietaksinen, "Adaptive document image binarization", Pattern Recogn, vol. 33, pp. 225-236, 2000.

[10] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", Journal of Electronic Imaging, vol. 13, no. 1, pp. 146-165, January 2004.

[11] M. Lettner, M. Diem, R. Sablatnig, and H. Miklas, "Registration and Enhancing of Multispectral Manuscript Images", Proc. 16th European Signal Processing Conference (EUSIPCO08), Lausanne, Switzerland, 2008.

[12] J.P. Lewis, "Fast normalized cross-correlation", Proc. Vision Interface, Quebec City, Canada, 1995, pp. 120-123.

[13] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms", Optics Letters, vol. 33, No. 2, 2008.

[14] J.J.K. Ó Ruanaidh, T. Pun, "Rotation, Translation and Scale Invariant Digital Image Watermarking", Signal Processing, 66 303-317, 1998.

[15] G. Wolberg, and S. Zokai, "Robust image registration using log-polar transform", Proc. of IEEE Intl. Conf. on Image Processing, Sep. 2000.

[16] A. Tonazzini, L. Bedini, and E. Salerno, "Independent Component Analysis for document restoration," International Journal of Document Analysis and Recognition, vol. 7, pp. 17-27, 2004.

[17] A. Tonazzini, E. Salerno, and L. Bedini, "Fast correction of bleed-through distortion in grayscale documents by a Blind Source Separation technique", International Journal of Document Analysis and Recognition, vol. 10, pp. 17-25, 2007.

[18] H.S. Baird, "Document image defect models and their uses", Proc. ICDAR'92, 1992, pp. 62-67.

[19] E. Badekas and N. Papamarkos, "Automatic evaluation of document binarization results" Proc. 10th Iberoamerican Congress on Pattern Recognition, Havana, Cuba, 2005, p. 1005–1014.

[20] A. Tonazzini, G. Bianco, and E. Salerno, "Registration and enhancement of double-sided degraded manuscripts acquired," Submitted to ICDAR'09, 2009.

[21] J. R. Fienup, "Invariant error metrics for image reconstruction", Appl. Opt. vol. 36, Issue 32, pp. 8352-8357.