JOINT NON-RIGID REGISTRATION AND RESTORATION OF RECTO-VERSO ANCIENT MANUSCRIPTS

P. Savino, L. Bedini, A. Tonazzini

Istituto di Scienza e Tecnologie dell'Informazione
Consiglio Nazionale delle Ricerche
Via G. Moruzzi, 1, 56124 Pisa, Italy

ABSTRACT

Ancient manuscripts written on both pages of the sheet are frequently affected by ink bleeding from the reverse side, which produces a significant degradation of the text. Effective digital image restoration techniques may require the use of the content of both document sides, thus needing their perfect alignment. Usually, recto and verso are not aligned either for rigid misalignments occurring during acquisition, or for non-rigid deformations of the sheet. In this paper we propose a novel method to jointly register and restore color recto-verso manuscript images in a piecewise manner, by subdividing the images into sub-images that exhibit apparent, different deformations of one with respect to the other. For each pair of corresponding sub-images, a specific projective transformation is computed, the two sub-images are registered, and then restored with a pixel-by-pixel algorithm that returns free of interferences versions of the images in their original acquisition layout. The projective transformation is estimated exploiting the precise computation of the shifts of a large number of small corresponding recto and verso patches, via correlation of their gradients. The experiments show that this combined procedure of local registration plus restoration can provide an excellent removal of bleed-through, while leaving unaltered the salient features of the original manuscripts.

Index Terms— Document image enhancement, Image registration, Bleed-through removal.

1. INTRODUCTION

The seeping of ink from the reverse side (bleed-through) is a severe deterioration affecting most ancient archival documents. This degradation, which cannot be physically removed, greatly impairs the legibility and fruition of the manuscript contents, besides being highly unpleasant.

Image processing techniques applied to the digital acquisitions of such manuscripts have proven to be often effective, especially those exploiting information of both recto and verso [1, 2, 3, 4, 5, 6, 7, 8, 9]. Using recto-verso images entails however their preliminary, very accurate alignment. Although nowadays many libraries and archives have specialized digitization devices, e.g. high resolution CCD cameras, mounted on mechanical equipments that guarantee a stable setup, misalignments frequently occur, due to the human intervention for repositioning the sheet to acquire the verso image, or to accidental movements of the camera. Furthermore, the sheet might be intrinsically deformed in a non-rigid way.

Registration of recto-verso images is a challenging task, due to sparsity of bleed-through, different color of the same stroke in the two sides, and different degree of curvature of recto and verso in books. Most methods consider a global affine transformation, estimated by minimizing the intensity differences between the two sides [10], or by using the Fourier-Mellin transform [11], or by a block-by-block strategy, where for each block in the recto the corresponding verso block is searched for in a larger window by spatial cross-correlation [12]. Projective transformations have also been proposed, e.g. in [13], where a feature-based method employs the corners detected from extracted character contours, and in [14], where the features are corresponding points located from the relative shifts of pairs of small recto-verso patches, computed via cross power spectrum. Non-rigid registration is proposed in [15], combining a global affine transformation with a free-form, hierarchical transformation based on B-splines and detected recto-verso corresponding points. A thin plate spline smoothness constraint is then applied for minimizing the residual complexity between the two sides, as done in [16].

In this paper we assume that the deformation between the two sides, though not describable by a unique geometrical transformation, is locally rigid. The images are subdivided into sub-images where we assume a homogeneous deformation. Each pair of sub-images are first registered, and then restored individually with the algorithm in [9]. Since the two processes act pixel-by-pixel, our final results are the free of interferences, un-registered images of the two sides, without any need of mosaicking the registered sub-images.

The geometrical transformation between the two sub-images is estimated by least mean square error (LMSE) from

978-1-5090-5524-1/16/$31.00 ©2016 IEEE

a large number of corresponding points located exploiting the relative shifts of pairs of small recto-verso patches at the same position. However, rather than using the shift property of the Fourier Transform of the intensities, here the shifts are computed by cross-correlation of the gradients of the patches. Indeed, though geometrically identical, recto-verso images have different intensities. Specifically, dark strokes in one side (foreground text) are lighter in the other (bleed-through pattern), and viceversa. As for images of a same scene taken with different sensors (e.g., RGB color channels), we thus inferred that recto and verso mainly correlate in correspondence of the object borders or textures, and that a measure of correlation is more reliable when performed on the gradients of the patches. For each patch in a side, the matrix of the correlations of its gradient with the gradients of all the patches of same size in the other side is computed, and the maximum of this matrix defines the relative shift we are looking for.

We compare the results of global and local registration in terms of normalized mutual information (NMI) of the registered images and in terms of quality of the restoration. In both cases, we obtain significant improvements.

The paper is organized as follows. In Section 2, the registration method is described and experimented both in its global and local modality, with qualitative and quantitative comparisons of the results. Section 3 summarizes the model-based algorithm for bleed-through removal, and provides examples of the final results of the entire computational chain. Finally, Section 4 concludes the paper with the description of possible improvements and extensions.

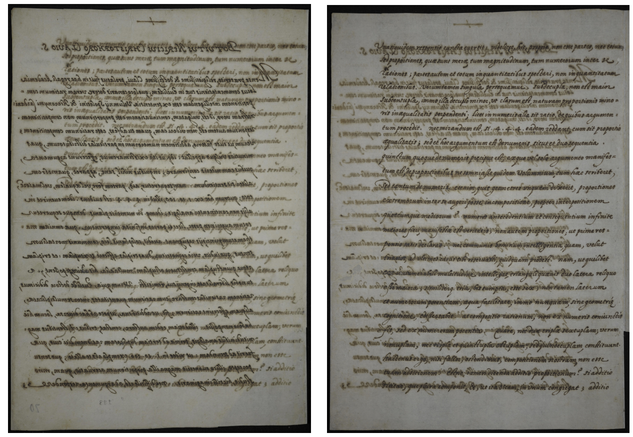
2. THE REGISTRATION ALGORITHM

The proposed registration method has been experimented on several color manuscripts. In the following discussion we refer to the results obtained on the color recto-verso pair of size 3114×4404, shown in Figure 1.

No matter if we are considering the entire manuscript or one of it sub-images, the related recto and verso images, of same size $M \times N$, are indicated by $r_c(i,j)$ and $v_c(i,j)$, where $c = \{R, G, B\}$ and $i = 1, ..., M$, $j = 1, ..., N$. Their registered versions are indicated as r_c^{REG} and v_c^{REG}. We assume a rigid, projective transformation from the recto to the verso, and a perfect alignment of the three color planes in the two sides, which entails a unique transformation for the three pairs of homologous channels. Thus, we can compute this transformation on a single pair of channels only, and discard the color subscript in the variables. Using homogeneous coordinates, a projective transformation is given by the following matrix:

$$P = \begin{bmatrix} s_x \cdot cos(\theta) + p_x b_x & -s_y \cdot sin(\theta) + p_y b_x & b_x \\ s_x \cdot sin(\theta) + p_x b_y & s_y \cdot cos(\theta) + p_y b_y & b_y \\ p_x & p_y & 1 \end{bmatrix} \quad (1)$$

where b_x and b_y, s_x and s_y, p_x and p_y represent translations, scale factors and projective deformations, respectively, along



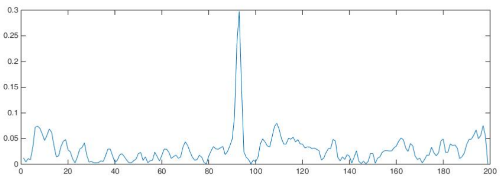
(a) (b)

Fig. 1. Original recto-verso manuscript: (a) recto (b) verso. Reproduction by courtesy of The Historical Archive of the Pontificia Università Gregoriana, APUG 529, c. 138r/v.

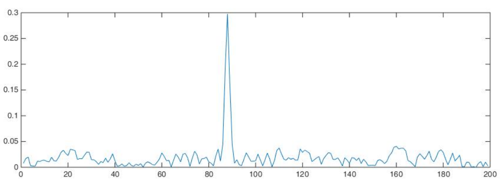
the x and y axes, and θ is the angle of rotation. Thus, given the image v of the verso, its versions registered on the recto v^{REG} will be built by applying to each pixel (i,j) of a ordered grid the transformation above, and then computing its value by bicubic interpolation of the values of those verso pixels that surround the possible non-integer location so found.

To determine the seven parameters of the transformation in eq.(1) by using LMSE, a minimum of four pairs of corresponding locations are necessary. However, to compensate for possible small mismatches of the corresponding locations, a much larger number of pairs should be prudently used. Here we assume that, whatever the transformation is, at the very local level it can be always approximated with a translation. Thus, we propose to select pairs of corresponding points by determining the displacements, in the verso side, of a large number of small patches picked-up in the recto.

Provided that these patches are not pure background and contain some text or bleed-through, each displacement can be computed by looking for that verso patch that maximally correlates with the recto patch at hand. If the sub-image deformation is not too strong, the verso region to be explored can be much smaller than the entire sub-image, thus saving computational time. More in detail, given a patch of size $m \times n$ centered in $(\hat{i}, \hat{j})$, we compute the correlation between this patch and every verso patch of same size spanning, pixel-by-pixel, a verso region of size $2m \times 2n$ and centered in $(\hat{i}, \hat{j})$. The novel aspect of this technique is that, rather than measuring the correlation of the intensities, we measure the correlation of the gradients of the patches. The rationale for this is that images of a same scene, but with different colors or intensity, have higher correlation just in correspondence of the object borders, that is at the higher intensity derivatives. The maximizer of the $2m \times 2n$ correlation matrix identifies the location of the verso patch exactly matching with the recto patch at hand. Figure 2 shows the plots of the correlation values along the x and y directions passing for the matrix maximizer. Note how the two peaks are well defined.



(a)



(b)

Fig. 2. Plots of the correlation values along the x and y directions passing for the maximizer of the correlation matrix.

To register, either globally or locally, the images in Figure 1 we used patches of size 101×101. Figures 3 (a) and (b) show the superposition in transparency of a detail of recto and verso before and after global registration. It is apparent that even global registration give satisfactory results. However, there are characters that still appear "split", clearly indicating that, for these images, the deformation is not homogeneous. As per the local registration of these images, we empirically chose the number and size of sub-images by examining the superposition of the original entire recto and verso. A finer criterium would consist of first performing a global registration and then choosing the sub-images looking at the areas where this global registration performed poorly. Using 12 sub-images, we obtained the excellent result shown in Figure 3 (c) for the detail selected for comparison. This quality is representative of all the 12 sub-images. The entire registered image cannot be shown, since, as said, we restore them individually, without mosaicking the registered sub-images.

We measured also quantitatively the performance of the two registration modalities by computing the Normalized Mutual Information (NMI) between the original recto, r, and the registered verso, v^{REG}:

$$NMI(r, v^{REG}) = \frac{MI(r, v^{REG})}{\sqrt{H(r) \times H(v^{REG})}} \quad (2)$$

with

$$MI(r, v^{REG}) = H(r) + H(v^{REG}) - H(r, v^{REG}) \quad (3)$$

being $MI(r, v^{REG})$ the Mutual Information between r and v^{REG}, and H the entropy. For the entire images of Figure 1, the NMI was 0.084 when un-registered, 0.207 after global registration, and 0.215 after local registration. For the sub-images containing the detail shown in Figure 3 the NMI was 0.059 when un-registered, 0.180 after global registration of the entire images, and 0.223 after local registration.

3. THE RESTORATION ALGORITHM

In both registration modalities, global and local, the recto side is considered as our reference image, so that it remains geometrically unaltered. The restoration algorithm thus applies to r_c and v_c^{REG}, either at the level of the entire image or at the level of sub-images, independently for each pair of color channels. Dropping again the subscript c, this algorithm, detailed in [9], is based on the following data model:

$$\begin{aligned} D_r(i,j) &= \hat{D}_r(i,j) + q_v(i,j)\left[h_v(i,j) \otimes \hat{D}_v^{REG}(i,j)\right] \\ D_v^{REG}(i,j) &= \hat{D}_v^{REG}(i,j) + q_r(i,j)\left[h_r(i,j) \otimes \hat{D}_r(i,j)\right] \end{aligned} \quad (4)$$

with

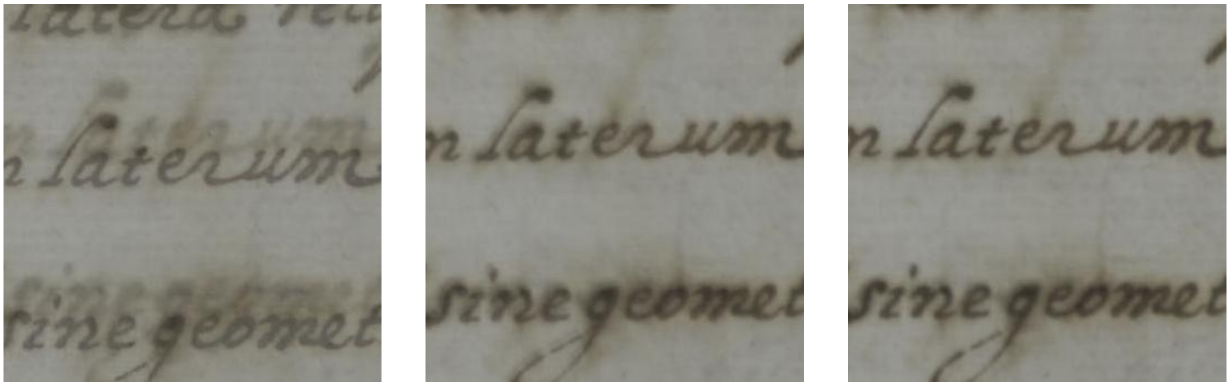
$$D_r(i,j) = -log\left(\frac{r(i,j)}{R_r}\right), \quad D_v(i,j) = -log\left(\frac{v^{REG}(i,j)}{R_v^{REG}}\right) \quad (5)$$

where $D_r(i,j)$ and $D_v^{REG}(i,j)$ are the recto and verso observed optical densities, and $\hat{D}_r(i,j)$ and $\hat{D}_v^{REG}(i,j)$ are the ideal ones, at pixel (i,j). Constants R_r and R_v^{REG} represent the predominant values of the two backgrounds, and h_r and h_v^{REG}, are unit volume Point Spread Functions (PSF), describing the smearing of ink penetrating the paper. These PSFs are stationary, empirically chosen as Gaussian functions, but characterized by pixel-dependent gains $q_r(i,j)$ and $q_v^{REG}(i,j)$, having the physical meaning of interference levels from one side to the other.

To invert the system of eq. (4) the model parameters $q_r(i,j)$ and $q_v^{REG}(i,j)$, $\forall(i,j)$, must be known. Ideally, they should be both zero in background and occlusion pixels, thus permitting to leave them unchanged, and mutually exclusive, i.e. one positive and the other zero, for pixels of foreground in one side and see-through in the other. Since we do not know a priori the nature of the various pixels, we estimate the interference levels at every pixel (i,j) from the data as:

$$\begin{aligned} q_r(i,j) &= \frac{D_v^{REG}(i,j)}{h_r(i,j) \otimes D_r(i,j) + \epsilon} \\ q_v^{REG}(i,j) &= \frac{D_r(t)}{h_v^{REG}(i,j) \otimes D_v^{REG}(i,j) + \epsilon} \end{aligned} \quad (6)$$

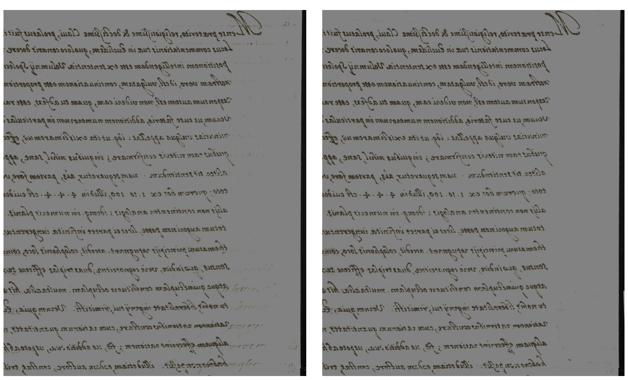
where ϵ is a small positive number to avoid indeterminacies or infinity. Eqs. (6) derive from the model of eq. (4) assuming that the ideal densities are both zero, which, at least in principle, holds true only for background pixels. For pixels of foreground in one side and bleed-through in the other only one equation makes sense, whereas in the occlusion pixels none of the two equations holds true. Assuming bleed-through lighter than foreground, for all pixels we maintain



(a) (b) (c)

Fig. 3. Example of recto-verso superposition before and after registration: (a) superposition of a detail the original images in Figure 1; (b) superposition of the same detail after global registration; (c) superposition of the detail after local registration.

the smallest between the two computed interference levels, and set to zero the other. This allows to correctly discriminate the two instances of foreground in one side and see-through in the other, but leaves a wrong positive interference level in one side in correspondence of some background pixels and in all the occlusion pixels. In [9] we commented on the consequences of this fact and proposed some partial remedies.



(a) (b)

Fig. 4. Restoration of a portion of the recto manuscript in Figure 1 (a) after: (a) global registration; (b) local registration.

On the basis of the available q_r and q_v^{REG} maps, $\hat{D}_r$ and $\hat{D}_v^{REG}$ are computed through the following single step:

$$\hat{D}_r(i,j) = D_r(i,j) - q_v(i,j)\left[h_v^{REG}(i,j) \otimes D_v^{REG}(i,j)\right]$$
$$\hat{D}_v^{REG}(i,j) = D_v^{REG}(i,j) - q_r(i,j)\left[h_r(i,j) \otimes \hat{D}_r(i,j)\right] \quad (7)$$

In case of global registration, this pixel-by-pixel restoration algorithm returns directly the two restored sides. In the local case, registration and restoration are performed individually on each pair or sub-images, thus obtaining a set of restored recto and verso sub-images, denoted as $\hat{r}_k$ and $\hat{v}_k^{REG}$, $k = 1, ..., K$, where K is the total number of sub-images. Since the recto is not affected by the registration process, we again obtain the restored version of the entire image straightforwardly, by recomposing the K sub-images. In order to obtain the restored verso, one way could be to repeat the process by inverting recto and verso. In our experiments we adopted a more direct approach, which does not require to register the images again. This consists in restoring each pixel (i, j) in each v_k^{REG}, and then, given the transformation matrices P_k, associating that restored value to the original pixel in v_k.

A comparison between the restoration obtained with the two registration modalities is shown in Figures 4 (a) and (b), for a large portion of the recto in Figure 1. It is apparent that local registration improves the restoration result.

4. CONCLUSION AND FUTURE WORK

We proposed the joint registration and restoration of recto-verso misaligned manuscripts affected by bleed-through, assuming the relative deformation to be locally rigid. For each pair of sub-images, the verso one is registered on the recto one, a pixel-by-pixel restoration is performed, and then the verso restored pixels are re-allocated in their original position, thus preserving the original appearance of the acquired images. A major novelty is the estimation of the various geometrical transformations based on the cross-correlation of image gradients, rather than image intensities. The results show a significant quantitative improvement of the registration performance, with a consequent much more effective removal of the bleed-through. Once the sub-regions are determined, all the procedure is fully automatic.

Future improvements could regard the automatic determination of the sub-regions as well, based, for instance, on the measures of NMI. Another issue is a more effective way to handle the occlusions in the restoration algorithm.

5. REFERENCES

[1] Grani Adiwena Hanasusanto, Zheng Wu, and Michael S. Brown, "Ink-bleed reduction using functional minimization," in *The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13-18 June 2010*, 2010, pp. 825–832.

[2] Yi Huang, Michael S. Brown, and Dong Xu, "User-assisted ink-bleed reduction," *IEEE Transactions on Image Processing*, vol. 19, no. 10, pp. 2646–2658, 2010.

[3] Róisín Rowley-Brooke, François Pitié, and Anil C. Kokaram, "A non-parametric framework for document bleed-through removal," in *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, 2013, pp. 2954–2960.

[4] Anna Tonazzini, Emanuele Salerno, and Luigi Bedini, "Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique," *IJDAR*, vol. 10, no. 1, pp. 17–25, 2007.

[5] Anna Tonazzini, Ivan Gerace, and Francesca Martinelli, "Multichannel blind separation and deconvolution of images for document analysis," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 912–925, 2010.

[6] Róisín Rowley-Brooke and Anil C. Kokaram, "Bleed-through removal in degraded documents," in *Document Recognition and Retrieval XIX - DRR 2012, 19th Document Recognition and Retrieval Conference, part of the IS&T-SPIE Electronic Imaging Symposium, Burlingame, California, USA, January 25, 2012, Proceedings*, 2012.

[7] Farnood Merrikh-Bayat, Massoud Babaie-Zadeh, and Christian Jutten, "Using non-negative matrix factorization for removing show-through," in *Latent Variable Analysis and Signal Separation - 9th International Conference, LVA/ICA 2010, St. Malo, France, September 27-30, 2010. Proceedings*, 2010, pp. 482–489.

[8] Francesca Martinelli, Emanuele Salerno, Ivan Gerace, and Anna Tonazzini, "Nonlinear model and constrained ML for removing back-to-front interferences from recto-verso documents," *Pattern Recognition*, vol. 45, no. 1, pp. 596–605, 2012.

[9] Anna Tonazzini, Pasquale Savino, and Emanuele Salerno, "A non-stationary density model to separate overlapped texts in degraded documents," *Signal, Image and Video Processing*, published online: 13 December 2014 2014.

[10] Eric Dubois and Anita Pathak, "Reduction of bleed-through in scanned manuscript documents," in *PICS 2001: Image Processing, Image Quality, Image Capture, Systems Conference, Montréal, Quebec, Canada, April 22-25, 2001*, 2001, pp. 177–180.

[11] Anna Tonazzini, Gianfranco Bianco, and Emanuele Salerno, "Registration and enhancement of double-sided degraded manuscripts acquired in multispectral modality," in *10th International Conference on Document Analysis and Recognition, ICDAR 2009, Barcelona, Spain, 26-29 July 2009*, 2009, pp. 546–550.

[12] Jie Wang, Michael S. Brown, and Chew Lim Tan, "Accurate alignment of double-sided manuscripts for bleed-through removal," in *The Eighth IAPR International Workshop on Document Analysis Systems, DAS 2008, September 16-19, 2008, Nara, Japan*, 2008, pp. 69–75.

[13] B. Li, W. Wang, and H. Ye, "Multi-sensor image registration based on algebraic projective invariants," *Optics express*, vol. 21, no. 8, pp. 9824–9838, April 2013.

[14] Pasquale Savino and Anna Tonazzini, "Digital restoration of ancient color manuscripts from geometrically misaligned recto-verso pairs," *Journal of Cultural Heritage*, vol. 19, pp. 511521, 2016.

[15] Jie Wang and Chew Lim Tan, "Non-rigid registration and restoration of double-sided historical manuscripts," in *2011 International Conference on Document Analysis and Recognition, ICDAR 2011, Beijing, China, September 18-21, 2011*, 2011, pp. 1374–1378.

[16] Andriy Myronenko and Xubo B. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Trans. Med. Imaging*, vol. 29, no. 11, pp. 1882–1891, 2010.