Image Registration of Ancient Documents

V. V. V. A. O. Martinez G. L. de Campos
Departamento de Engenharia de Computação e Sistemas Digitais
Escola Politécnica — Universidade de São Paulo
São Paulo, SP, Brazil.

Abstract *This work describes a technique to align images of ancient documents. Ancient documents present a significant level of bleed-through. An important technique for bleed-through removal requires four images of each leaf: the two sides with a white and a black background. To align the images of the document pages, an image registration approach is required. The images have acquisition differences due to translation, rotation and geometric distortions. The adopted image registration approach is divided into three stages. The first is the control point selection: images of the pages are segmented into text lines and then each line is segmented into words. Therefore, each selected control point corresponds to the beginning of a text word. In the second stage, a matching is performed between the extracted control points of both images. Finally, the target image is mapped using thin plate spline functions to coincide it with the reference image.*

Keywords: align, image registration, thin plate spline, image of ancient documents.

1 Introduction

Ancient books constitute a legacy that should be preserved. With digitizing techniques it is possible to have these old books in digital format and able to be reached by a wider public. One of the problems encountered for good book reproduction is bleed-through.

Because most types of paper are partly translucent, a digital image of one side of a document is often contaminated by images on the other side. This phenomenon is known as *bleed-through*. In old books, the paper is more transparent than modern paper and the problem of transparency is more evident.

The bleed-through is clearly an unwanted phenomenon that compromises the quality of the reproduced image. To this end, a method for removing bleed-through was proposed by Stolfi [10]. The method requires that each side of the document be imaged twice, once against a black background and another against a white background. Intuitively, with a black background the images are not affected by the marks on the other side, so they are not affected by bleed-through. With white background, the bleed-through is more apparent.

Scanning one of the sides of the document twice may cause differences in the images due to translation, rotation and deformation. The method above requires that the images of pages with black and white background be aligned before the removal of bleed-through. To align the two images of the same page, an image registration approach is used. There aren't many works related with alignment of two pages of the same document. The works that deal with this, like [12] and [6], don't deal with deformation. This paper presents an image registration technique developed for this purpose.

There are several techniques for each one of the steps of image registration and there are several combinations of these techniques. In this work, the page segmentation in words is used as control point selection. A novel control point matching, based in line by line point extraction, is presented. Finally, a TPS function is applied as a mapping function.

In Section 2, the image registration process is introduced. Section 3 describes image registration with thin plate splines. The control point selection is presented in Section 4 and the control point correspondence is described in Section 5. Section 6 describes the results. Finally, the last section of the paper presents the conclusion.

2 Image Registration

Image registration is the process of overlaying two or more images of the same scene which present differences due to the acquisition process, like temporal or viewpoint differences. A number of image registration methods have been presented in the literature, for example Brown [2] and den Elsen et al. [3].

Image registration consists of three main steps. The first step is the selection of control point in the reference and target images. The second is the determination of the correspondence between the control points. The last is the determination of one mapping function between both images so that the differences are minimized.

The mapping function between the images are f_x and f_y:

$$u = f_x(x, y) \quad \text{and} \tag{1}$$

$$v = f_y(x, y). \tag{2}$$

These functions relate the reference image coordinates (x, y) to the target image coordinates (u, v).

3 Thin Plate Spline Mapping Functions

When the differences between images include geometric distortions, it is necessary to use elastic models to register images. A method to register images with these distortions is Thin Plate Spline – TPS proposed by Goshtasby [5]. In this work, the object of interest is paper, which may be modeled as an elastic body and TPS was chosen to be the mapping functions.

The problem of finding the mapping functions f_x and f_y with TPS can be formulated as an interpolation problem:

$$S = \{(x_i, y_i, u_i): \quad \text{for } i = 1, \ldots, n\} \tag{3}$$

and

$$Q = \{(x_i, y_i, v_i): \quad \text{for } i = 1, \ldots, n\}. \tag{4}$$

One would like to find smooth surfaces $f_x(x, y)$ and $f_y(x, y)$. The surface $f_x(x, y)$ must pass through all points in S and the surface $f_y(x, y)$ must pass through all points in Q.

The required surfaces may be found through the expressions

$$\begin{aligned} f_x(x, y) = & a_0 + a_1 x + a_2 y + \\ & \sum_{i=1}^{n} F_i r_i^2 \log r_i^2 \end{aligned} \tag{5}$$

and

$$\begin{aligned} f_y(x, y) = & b_0 + b_1 x + b_2 y + \\ & \sum_{i=1}^{n} G_i r_i^2 \log r_i^2, \end{aligned} \tag{6}$$

where $r_i^2 = (x - x_i)^2 + (y - y_i)^2$.

The coefficients a_0, a_1, a_2 and F_i, for $i = 1, \dots, n$, are determined by substituting the n control points in the Equation (5) and by solving the system of $(n+3)$ linear equations

$$f(x_i, y_i) = u_i, \quad \text{for } i = 1, \dots, n, \qquad (7)$$

$$\sum_{i=1}^{n} F_i = 0, \quad \sum_{i=1}^{n} F_i x_i = 0, \quad \text{and} \qquad (8)$$

$$\sum_{i=1}^{n} F_i y_i = 0. \qquad (9)$$

The surface $f_x(x, y)$ represents the first component of the mapping function and the surface $f_y(x, y)$ representing the second component is determined similarly. After the coefficients are determined and the mapping functions f_x e f_y are obtained, it is possible to register the images.

One comment about the cost for calculating the coefficients is necessary. This cost may be very high and is dependent on the number of control points. The works [4] and [1] deal with this subject. For this reason, one may wish to obtain a reduced control points set.

4 Control Point Selection

The control point selection is based on document processing techniques. In the literature, there are several references to document processing like in [9] and [11].

The control point selection was made with the constraint that the images are composed only of text. The control point selection is based on page segmentation. The page is segmented into lines and then into words. Each word of the text corresponds to one control point. This segmentation process is similar that used by Muge et al. [8]. Segmentation into lines is performed with the computation of the projection profile across horizontal lines of the image. Before the page is segmented in lines it is necessary to correct the inclination of the page. There are many techniques for skew correction and a list of them can be found in [11].

The word segmentation relies on a length assumption: the space between two characters is much smaller then the space between two words. To solve this problem, it is possible to construct a histogram of the lengths of white spaces that are in the line from a vertical projection profile. This word segmentation policy may fail if the spaces between words and characters are irregular. In this case, some words can be split apart and others can be aggregated.

5 Control Point Matching

Control point matching is complicated by the fact that the above procedure may result in a different number of points in each image. There are many works in literature that present techniques for control point correspondence. If n is the number of control points, the best techniques have complexity $O(n^2 \log n)$.

A novel control point matching algorithm was developed, relying on the segmentation performed line by line and on the fact that the control points are ordered by the column of the beginning of each word. The idea of the algorithm is, given two sets of control points, one for each line, the sets are visited comparing pairs of points. If the distance between one pair of points is in a determined threshold, then this points are matched and the algorithm continues with the next points. In the other case, the algorithm searches sets for a point with the smallest column coordinate and the next comparison will be between the point of the set with the largest column coordinate and the next point of the set with the smallest one. The control point matching algorithm complexity is linear with the number of control points.

The fact that the control point selection is performed line by line may be applied for any approach of control point matching, reducing the computational complexity of this work.

6 Results

To test the image registration techniques described images obtained from the book *Auto da Fé em Coimbra* [7] were used, and they were acquired in gray level. In Figures 1 and 2 is possible to see the reference image and the target image. In Figure 6 is the sum of the two images. The control points extracted are presented in Figures 3 and 4 represented by white squares. Finally, Figure 5 illustrates the image registration using TPS, and Figure 7 shows the sum of image registration and reference image.

The comparison between Figure 6 and Figure 7 shows that the differences between images were significantly reduced and the result is considered good. In a careful analysis of the Figure 7, it is possible to note that in the region where there are words in italic, the alignment is not good. This occurs because the spaces between words in italic are smaller, i. e. different, than those between other words, and it is an expected result.

It is important to note that the deformation of those images was intentionally much larger then the deformations that occur in real situations, in order to stress the process to its limits.

7 Conclusion

In this paper, the alignment was performed considering that images have acquisition differences due to translation, rotation and geometric distortions. To minimize these differences TPS was used as a mapping function. An algorithm for control point selection was developed, constrained to pages with text only, with good results.

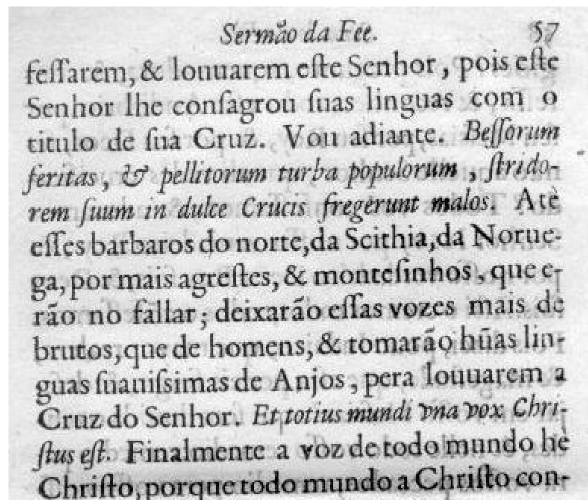

Sermão da Fee. 57
feſſarem, & louuarem eſte Senhor, pois eſte
Senhor lhe conſagrou ſuas linguas com o
titulo de ſua Cruz. Vou adiante. *Beſſorum
feritas, & pellitorum turba populorum, ſtrido-
rem ſuum in dulce Crucis fregerunt malos.* Atè
eſſes barbaros do norte, da Scithia, da Norue-
ga, por mais agreſtes, & monteſinhos, que e-
rão no fallar; deixarão eſſas vozes mais de
brutos, que de homens, & tomarão hũas lin-
guas ſuauiſsimas de Anjos, pera louuarem a
Cruz do Senhor. *Et totius mundi vna vox Chri-
ſtus eſt.* Finalmente a voz de todo mundo he
Chriſto, porque todo mundo a Chriſto con-

Figure 1: Reference image scanned with white background.

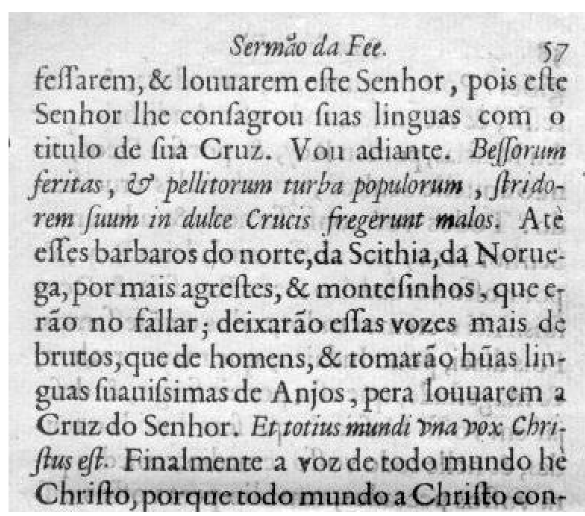

Sermão da Fee. 57
feſſarem, & louuarem eſte Senhor, pois eſte
Senhor lhe conſagrou ſuas linguas com o
titulo de ſua Cruz. Vou adiante. *Beſſorum
feritas, & pellitorum turba populorum, ſtrido-
rem ſuum in dulce Crucis fregerunt malos.* Atè
eſſes barbaros do norte, da Scithia, da Norue-
ga, por mais agreſtes, & monteſinhos, que e-
rão no fallar; deixarão eſſas vozes mais de
brutos, que de homens, & tomarão hũas lin-
guas ſuauiſsimas de Anjos, pera louuarem a
Cruz do Senhor. *Et totius mundi vna vox Chri-
ſtus eſt.* Finalmente a voz de todo mundo he
Chriſto, porque todo mundo a Chriſto con-

Figure 2: Target image scanned with black background.

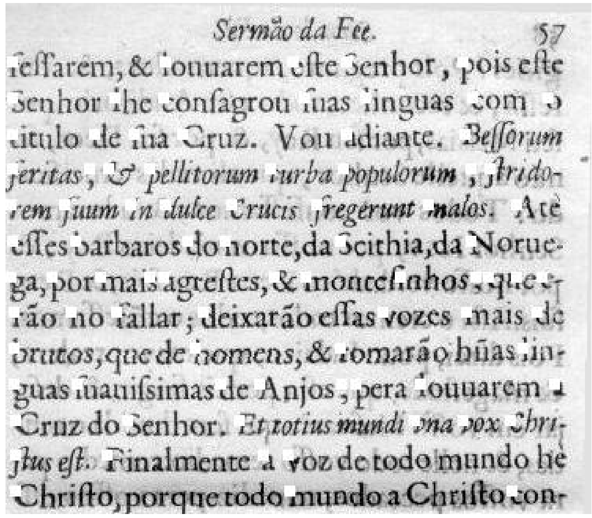


Figure 3: Control points of reference image.

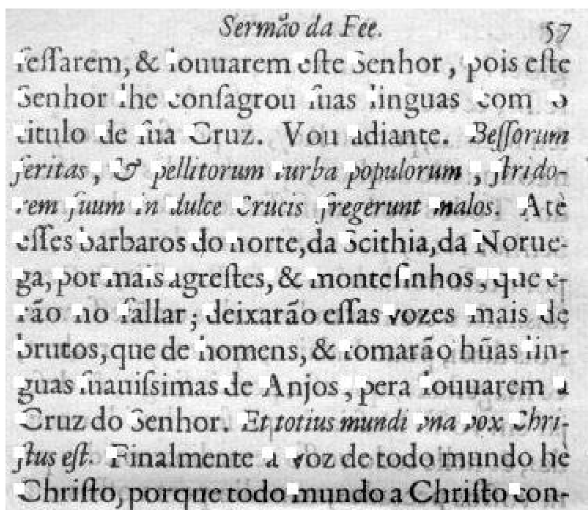


Figure 4: Control points of target image.

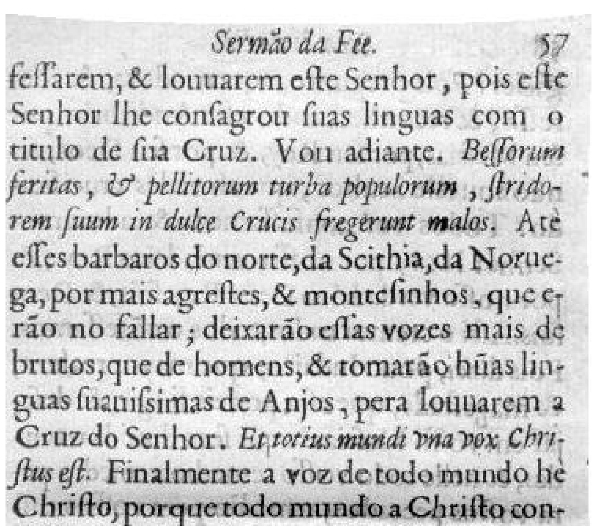
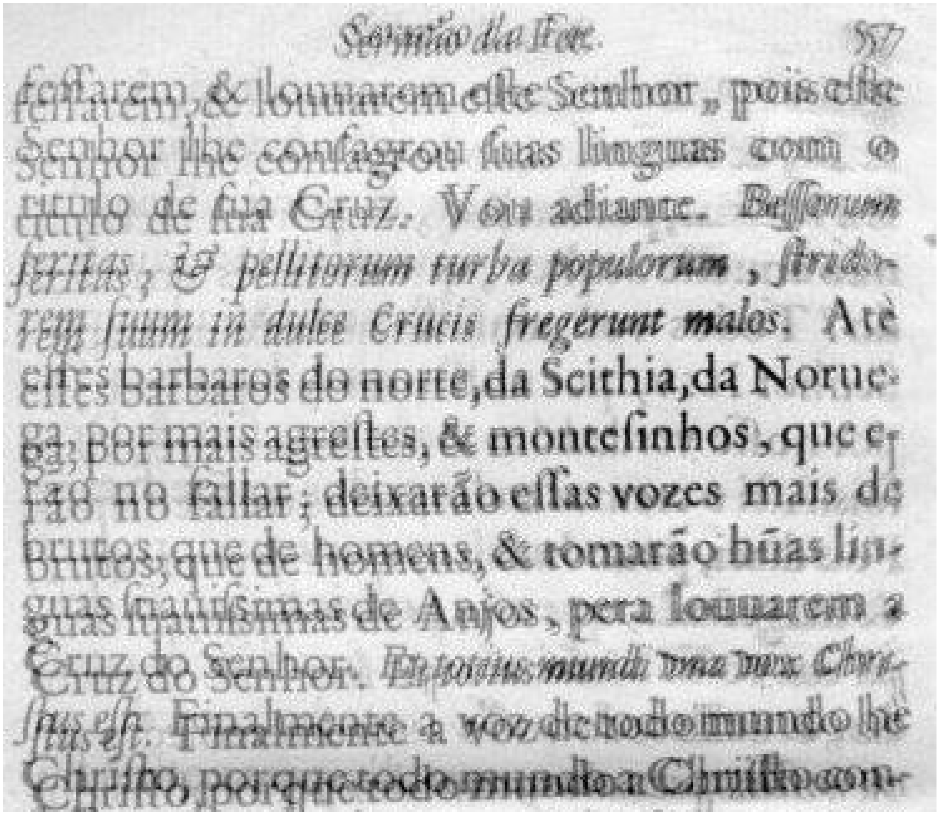


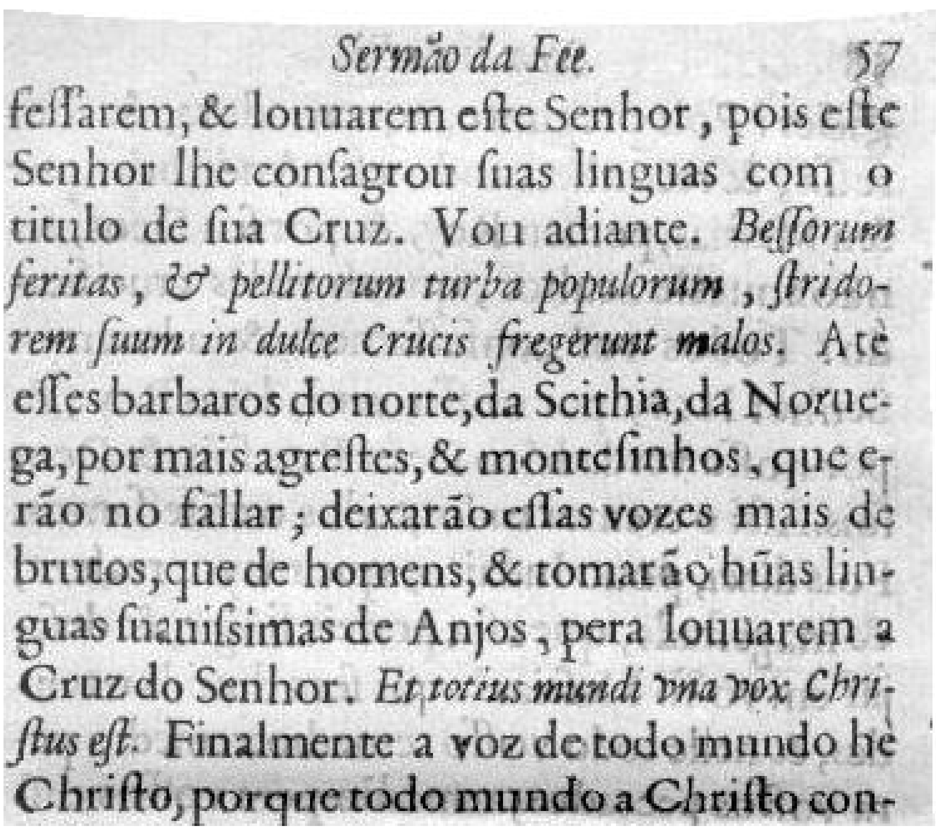
Figure 5: Registration of target image.



Sermão da Fee. 57

feſſarem, & louuarem eſte Senhor, pois eſte
Senhor lhe conſagrou ſuas linguas com o
titulo de ſua Cruz. Vou adiante. *Beſſorum*
feritas, & pellitorum turba populorum, ſtrido-
rem ſuum in dulce Crucis fregerunt malos. Atè
eſſes barbaros do norte, da Scithia, da Norue-
ga, por mais agreſtes, & monteſinhos, que e-
rão no fallar; deixarão eſſas vozes mais de
brutos, que de homens, & tomarão hũas lin-
guas ſuauiſsimas de Anjos, pera louuarem a
Cruz do Senhor. *Et totius mundi vna vox Chri-*
ſtus eſt. Finalmente a voz de todo mundo he
Chriſto, porque todo mundo a Chriſto con-

Figure 6: Sum of reference image and target image.



Sermão da Fee. 57

feſſarem, & louuarem eſte Senhor, pois eſte
Senhor lhe conſagrou ſuas linguas com o
titulo de ſua Cruz. Vou adiante. *Beſſorum*
feritas, & pellitorum turba populorum, ſtrido-
rem ſuum in dulce Crucis fregerunt malos. Atè
eſſes barbaros do norte, da Scithia, da Norue-
ga, por mais agreſtes, & monteſinhos, que e-
rão no fallar; deixarão eſſas vozes mais de
brutos, que de homens, & tomarão hũas lin-
guas ſuauiſsimas de Anjos, pera louuarem a
Cruz do Senhor. *Et totius mundi vna vox Chri-*
ſtus eſt. Finalmente a voz de todo mundo he
Chriſto, porque todo mundo a Chriſto con-

Figure 7: Sum of image registration and reference image.

Acknowledgments

We wish to thank Dr. Anna H. R. Costa (EPUSP) for her kind help.

References

[1] I. Barrodale, D. Skea, M. Berkley, R. Kuwahara, and R. Poeckert, *Warping digital images using thin plate splines*, Pattern Recognition, 26 (1992), pp. 375–376.

[2] L. G. Brown, *A survey of image registration techniques*, ACM Computing Surveys, 24 (1992), pp. 325–376.

[3] P. A. V. den Elsen, E. D. Pol, and M. A. Viergever, *Medical image matching – a review with classification*, IEEE Engineering in Medicine and Biology, 12 (1993), pp. 26–39.

[4] J. Flusser, *An adaptative method for image registration*, Pattern Recognition, 25 (1992), pp. 45–54.

[5] A. Goshtasby, *Registration of images with geometric distortions*, IEEE Transactions on Geoscience and Remote Sensing, 25 (1988), pp. 60–64.

[6] C. H. Leung, W. C. Tam, and Y. S. Cheung, *Recognition of handwritten chinese characters by elastic matching*, Image and Vision Computing, 16 (1998), pp. 979–988.

[7] F. Mendonça, *Sermão no Auto da Fé em Coimbra*, Na oficina de Diogo Gomez de Loureyro, 1619.

[8] F. Muge, I. Granado, M. Mengucci, P. Pina, V. Ramos, N. Sirakov, J. R. C. Pinto, A. Marcolino, M. Ramalho, P. Vieira, and A. M. Amaral, *Automatic feature extraction and recognition for digital access of books of the renaissance*, in Research and Advanced Technology for Digital Libraries, J. Borbinha and T. Baker, eds., vol. 1923 of Lecture Notes in Computer Science, Springer-Verlag, 2000, pp. 24–34.

[9] L. O'Gorman and R. Kasturi, eds., *Document Image Analysis*, IEEE Computer Society Press, 1995.

[10] J. Stolfi, *Removing optical bleed-through from imaged documents.* To appear.

[11] Y. Y. Tang, S. Lee, and C. Y. Suen, *Automatic document processing: a survey*, Pattern Recognition, 29 (1996), pp. 1931–1952.

[12] A. Zappalá, A. Gee, and M. Taylor, *Document mosaicing*, Image and Vision Computing, 17 (1999), pp. 589–595.