A NEW EFFICIENT CROSS-CORRELATION BASED IMAGE REGISTRATION TECHNIQUE WITH IMPROVED PERFORMANCE

Kostas Berberidis and *Irene Karybali*
Dept. of Computer Engineering and Informatics
and Computer Technology Institute
University of Patras, School of Engineering
26500 Rio - Patras, Greece
Tel: +30 61997708; fax: +30 61991909
e-mails: { berberid, karybali}@ceid.upatras.gr

ABSTRACT

In many applications the image registration task turns out to be a fundamental prerequisite for any further processing and analysis. In the proposed paper a new efficient technique is presented, appropriate for registering images which are translated and rotated versions of a reference image. The technique consists of two main parts: an efficiently implemented prewhitening part, and an iterative part which yields the unknown displacements after a few iterations. The cross-correlation operations involved in the iterative part are performed efficiently via a new scheme based on a proper partitioning of the images and the use of the FFT. Note that this efficient scheme is also applicable to the conventional spatial cross-correlation method. The new registration technique exhibits a superior performance as compared to the conventional cross-correlation method.

1 INTRODUCTION

In many image processing applications it is necessary to form a pixel-by-pixel comparison and find the underlying correspondence between two or more images. The different images are either images of the same object taken from different sensors or images of the same object taken at different times. Thus, between these images there may be translational shifts, rotational differences, scale and perspective view differences [1]. Remote Sensing Systems, Synthetic Aperture Radar Imaging, Computed Tomography using Magnetic Resonance Imaging, are some of the many applications in which image registration is required. In some applications the shifts and differences are detected and corrected off-line, while in others the acquired sequence of images has to be aligned in real time.

Over the last two decades a broad range of registration techniques has been developed for various types of data and problems [2], [3]. The existing literature is divided into two major classes, namely the techniques based on area correlation either in spatial or frequency domain, and the techniques based on matching properly chosen features or models of the images. The several techniques are compared according to criteria as accuracy, applicability and restrictions, level of automation and computational complexity.

Phase correlation and spatial cross-correlation are two well-known techniques for image registration. In this paper we concentrate on spatial cross-correlation for the following main reason. If the images have white noise, i.e. noise which is spread across all frequencies, the location of the peak will be inaccurate since the phase difference at each frequency is corrupted. In this case, methods which find the peak of the spatial cross-correlation sequence are preferable [2].

The proposed technique restores both translational and rotational differences between two images. The technique comprises two main parts. In the first part an appropriate pre-filtering is applied increasing the discrimination capability of the whole technique. More specifically, a two-dimensional prediction error filter corresponding to the reference image is computed, using an efficient algorithm, and then applied to both the reference and the displaced image. The second part is a cross-correlation based iterative procedure which having as inputs the pre-whitened images yields after a few iterations the displacement parameters for both translation and rotation. The implementation of the most costly steps of the iterative part (i.e. cross-correlations) is done efficiently resulting in a considerable reduction of the overall computational cost of the technique. It should be noted that the scheme developed for the cross-correlations' computation is applicable to other spatial cross-correlations methods as well. The proposed registration technique has been compared with the conventional cross-correlation method, implemented according to the suggested iterative algorithm. As shown via extensive simulations (and justified theoretically) the performance of the proposed technique is superior to the conventional method and in most cases the improvement is considerable.

This work was supported in part by the General Secretariat of Research & Technology of Greece under grant $\Pi ABET$-2000 no. 00BE363, and in part by the Computer Technology Institute (CTI) of Patras.

The basic steps of the proposed technique are discussed in section 2. In section 3 efficiency issues concerning the implementation of the several steps of the technique are presented. Finally, in section 4, results of extensive simulations are provided.

2 THE PROPOSED TECHNIQUE

Let us denote as A the so-called reference image and as B the second image which is a translated and rotated version of the first one. The aim is to correct translation and rotation displacements between A and B. The basic steps of the proposed technique are described in the following two subsections.

2.1 Cross-correlation of pre-whitened images

The conventional cross-correlation technique has the following disadvantages. At first, the possible presence of a secondary maximum makes it difficult to select the actual maximum. Moreover, the cross correlation peak may be rather broad, so that the accuracy in determining its position may be poor.

In order to sharpen the cross-correlation peak we suggest whitening the input images. The whitening method adopted in this paper is to prefilter the images with a two dimensional prediction error filter corresponding to a properly selected area of the reference image. Thus an amount of redundant information is taken away from A and B and the resulting images are less correlated. In this manner an enhancement of the discrimination capability of the cross-correlation method can be achieved. In subsection 3.1 an efficient scheme to apply this procedure is proposed.

Let us now recall that the predicted value of a pixel $x(m,n)$ is given by [4]

$$\widetilde{x}(m,n) = \sum_i \sum_j a(i,j)x(m-i,m-j) \qquad (1)$$

where $a(i,j)$ is a 2-D linear prediction coefficient. If an $M \times M$ causal region of support for the linear prediction array is selected, then the 2-D linear prediction coefficients minimizing the cost function

$$J_{LP} = E\{|x(m,n) - \widetilde{x}(m,n)|^2\} \qquad (2)$$

are computed via the following set of normal equations

$$\sum_{i,j\epsilon\Omega} \sum a(i,j) r_{xx}(k-i,l-j) = \begin{cases} 1, (k,l) = (0,0) \\ 0, \quad otherwise \end{cases} \qquad (3)$$

where r_{xx} is the auto-correlation coefficients of the given image. The prediction error filter is an $M \times M$ array having 1 in the bottom right corner of the QP region of support and the coefficients -a(i,j) in the other places.

If the displacement where either translation or rotation only, then the procedure would be simply the following. The prediction error filter h_A, for image A, is first computed and then is convolved with both A and B to yield E_A and E_B, respectively. Subsequently the conventional spatial cross-correlation technique is applied to E_A and E_B and the unknown displacement is the one corresponding to the best match, which occurs when the cross-correlation measure takes its maximum value. However, if both translational and rotational displacements are present then the iterative algorithm of next subsection is suggested. A similar iterative procedure based on the so called AMDF method has been proposed in [5].

2.2 Summary of the proposed technique

The proposed technique consists of the three steps described below. The first two steps take place once and constitute the non-iterative part while the third step is the iterative part of the technique.

1. Compute the prediction error filter, h_A, for image A.

2. Compute the convolution E_A of h_A and A, and the convolution E_B of h_A and B.
   Set $i = 1$, $E_{A_0} = E_A$, $E_{B_0} = E_B$

3. Iterative step
   3.1 Translation

   - Cross-correlate $E_{A_{i-1}}$ and $E_{B_{i-1}}$ for all possible translations. Find the position (x_i, y_i) of the best match.

   - Translate the one image with respect to the other according to the translational displacement (x_i, y_i) and extract the overlapping subimages $\widetilde{E}_{A_i}$,$\widetilde{E}_{B_i}$.

   3.2 Rotation

   - Rotate $\widetilde{E}_{A_i}$ by -θ, -θ + $\Delta\theta$, . . . , θ and cross-correlate the common areas of the rotated $\widetilde{E}_{A_i}$ and $\widetilde{E}_{B_i}$. Find the position θ_i of the best match.

   - Rotate the one image in terms of the other according to the rotational displacement θ_i and extract the overlapping subimages E_{A_i}, E_{B_i}.

   3.3 If x_i, y_i, $\theta_i \neq 0$, then set $i = i+1$ and repeat step 3.

The algorithm terminates when no translational and rotational displacements from one step to another can be detected. The overall translational and rotational displacements can be found by summing the respective displacements obtained at each iteration. These displacements represent the overall transformation of one image into the other.

3 EFFICIENCY ISSUES

In this section efficient implementation schemes are proposed for the most computationally thirsty steps of the technique.

3.1 Computation of the Prewhitened Images

As already mentioned, before performing the image prewhitening we need to compute the 2-D prediction error filter. Note that the normal equations (3) are identical to the 2-D Yule-Walker equations for a causal 2-D AR process. In the given data case this system of equations can be solved with $O(M^4)$ complexity by using a Levinson-type algorithm which exploits the rich structure of the involved autocorrelation matrix [4].

Subsequently, as dictated by step 2 of the proposed technique, the 2-D prediction error filter must be convolved with both A and B. Let us assume, without loss of generality, that the sub-area of image A chosen for the further matching procedure has dimensions $N \times N$. Thus the $M \times M$ prediction error filter must be convolved with the $N \times N$ area of image A and with an $(N + D) \times (N + D)$ area of image B (if we assume that the displaced image is within an $(N+D)\times(N+D)$ area). The required convolutions can be done efficiently using FFT and the 2-D version of the overlap-save method [6]. The computational burden of this step is approximately $O(N^2 logM)$.

3.2 Computation of the Iterative Part Correlations

Sub-step 3.1 is the heaviest one of the iterative part in terms of complexity. Recall that in this sub-step, at each iteration, the optimum translational shift is seeked. The procedure of sub-step 3.1 can be considered as a two-dimensional search as opposed to the one of 3.2 which can be viewed as an one-dimensional search. More specifically, if we assume that the maximum translational shift at each dimension is L, then the $N \times N$ area of image A has to be correlated with a corresponding $(N + L) \times (N + L)$ area of image B. Moreover, the cross-correlation output has to be computed for $L^2 + L$ possible shifts. Commonly this procedure is done efficiently in the frequency domain using FFT and the required complexity is $O(N^2 logN)$.

It should be noticed however, that in most cases in practice, it can be reasonably assumed that $L << N$, meaning that the translational shifts are much less than the image size. In such cases, the above correlation is of a special type, since only a small number of correlation outputs relative to the image size is required. This observation was the motivation to derive an efficient scheme of complexity $O(N^2 logL)$. Unfortunately due to space limitations we shall only very briefly describe the scheme.

First, both images are properly partitioned into $2L \times 2L$ blocks. Specifically, image A is partitioned into non-overlapped and consecutive $2L \times 2L$ blocks. The upper left quarter of each block is the corresponding $L \times L$ part of A and the other three quarters have zero elements. On the other hand, image B is partitioned into properly overlapped $2L \times 2L$ blocks. Then, the corresponding blocks of the two images are cyclically convolved, using FFT, and the resulting arrays are summed in the frequency domain to form a $2L \times 2L$ array. Finally, we take the IFFT of this array and from the resulting one we keep only an appropriate part which corresponds to the desired correlation outputs. The main differences with the usual 2-D overlap-save method is that both 2-D sequences are partitioned (in different ways), and the summation of the blocks is done in the frequency domain, thus only one IFFT is needed at the end.

4 SIMULATION RESULTS

The performance of the proposed algorithm was tested on two different images, the image of Lena (Fig.1), of size 256×256, and an aerial image of size 290×445 (Fig.2). For each one, we considered a pair of scenes (sub-images) of size 200×200, with the second scene being a translated and rotated version of the first. Also, for each of the two image cases we have conducted 20 different experiments. The proposed method has been compared with the conventional cross-correlation method, implemented according to the iterative part of the proposed technique (i.e. step 3) so as to tackle both translation and rotation.

In order to evaluate the performance of our method we used two criteria, the number of completely successful registration (number of hits) and the total matching error defined as

$$D = \frac{1}{N_1 \cdot N_2} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} |S_A(i,j) - S_B(i,j)| \qquad (4)$$

where S_A and S_B are the subimages that resulted after the registration of images A and B and $N_1 \times N_2$ is the number of pixels in either image.

The total matching error for the 20 experiments is plotted in Figure 3 for Lena image, and in Figure 4 for the aerial image. The dashed curve corresponds to the total matching error of the conventional cross-correlation technique , while the solid curve corresponds to the total matching error of the proposed technique. These simulation results show clearly that the proposed method performs better than the conventional cross-correlation method. In Table 1 and Table 2, the number of successful registration (hits), the mean value and the variance of the total matching error are shown, for Lena and aerial image, respectively. Our method has more hits and the mean values and variances of the total matching error are smaller. It must be noted that when the proposed technique fails to find the exact displacements, the estimation that makes is very close to the real ones, contrary to the corresponding estimates of conventional cross-correlation which in some cases may depart noticeably from the true ones.

We have conducted many other experiments, with different type of images, all leading to the same conclusion. Moreover, extensive simulations have shown that our technique performs also better than the conventional cross-correlation when either only translation or only rotation displacements are encountered. Thus, it was expected to perform better in the combined case as well, as shown above.

5 CONCLUSION

A new efficient technique has been developed, appropriate for registering images which are translated and rotated versions of a reference image. The technique consists of two main parts: an efficiently implemented prewhitening part, and an iterative part which yields the unknown displacements after a few iterations. It has been shown that the proposed technique exhibits better registration capabilities as compared to the conventional cross-correlation method and even when it fails to find the exact displacements it provides estimates very close to them. Moreover, a computationally efficiently scheme has been developed for computing the bulky computations involved in the correlation steps of the technique. This new scheme is applicable to any spatial cross-correlation based technique.

References

[1] W. P. Pratt, "Digital Image Processing", John Willey and Sons, 2001.

[2] L. G. Brown, "A Survey of Image Registration Techniques", ACM Computing Surveys, vol. 24, no. 1, pp. 326-376, Dec. 1992.

[3] J. Le Moigne et al, "First Evaluation of Automatic Image Registration Methods", Proc. IEEE IGARSS-98, Seattle, USA, July 1998.

[4] S. L. Marple Jr., "Digital Spectral Analysis", Prentice-Hall, 1987.

[5] R. A. Baggs, D. E. Tamir, T. Lam, "Image Registration Using the Length Code Algorithm", Proc. IEEE SHOUTHEASTCON-96, Tampa, Florida, April 1996.

[6] J.S. Lim, "Two-Dimensional Signal and Image Processing", Prentice-Hall, 1990.

Figure 1: The image of Lena

Figure 2: The aerial image

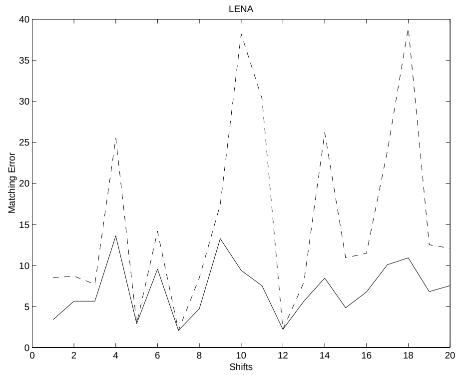

Figure 3: The matching error for Lena Image

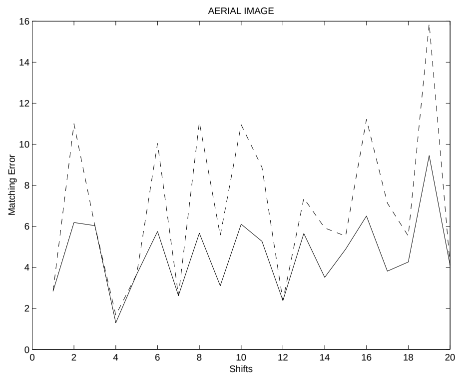

Figure 4: The matching error of Aerial Image

LENA	**Hits**	**Mean**	**Variance**
Cross-Correlation	3	15.5055	126.9687
Proposed method	8	7.0421	11.3229

Table 1: Results for Lena

AERIAL IMAGE	**Hits**	**Mean**	**Variance**
Cross-Correlation	6	6.9658	14.3385
Proposed method	11	4.6507	3.4834

Table 2: Results for the Aerial image