MULTIRESOLUTION IMAGE REGISTRATION

M. Corvi and G. Nicchiotti

Elsag Bailey, Central Research Dept.,
16154 Genova, Italy
email nicchiot@elsag.it

ABSTRACT

The paper describes an automatic registration procedure based on a multiresolution analysis of images. The approach is quite general and can be applied to a large variety of images. Furthermore the algorithm is very robust and can effectively cope with a considerable range of transformations, since the registration is obtained iteratively at different multiresolution scales. The procedure is completely automatic, and relies on the grey level information content of the images. We have applied the algorithm to test images of banknotes, aerial stereo pairs, multispectral and SAR images. In all the cases we have obtained excellent results, outperforming the best algorithms available at present and used in industrial applications.

1. INTRODUCTION

Image registration aims to overlay two different images of the same scene or two images of similar objects. It is used in vision based systems to carry out comparisons and difference estimates and represents an important tool for multitemporal analysis of remotely sensed data and quality control purposes (banknotes, printed circuits, etc.)[1]-[2].

The registration process is usually carried out in three steps. The first step consists of the selection of reference points on the first image the Control Points (CPs). Next for each CP the best match on the second image is determined. Finally the parameters of the best transformation connecting the matched CPs are estimated, usually with a least mean square (LMS) algorithm [3].

In this work we describe a fully automatic registration procedure based on a multiresolution analysis of images which is obtained by means of a discrete wavelet transform (DWT) [4].

An image registration approach based on multiresolution analysis has been already proposed by [5], but our algorithm results more robust, faster and requires less computer memory. Indeed in our approach there are two main differences with respect to [5].

We use the residual images of the DWT and a clustering technique as basis for the initial estimate of the registration transformation, while [5] uses only the DWT coefficients. We have found that this strategy increases the robustness of the whole procedure providing a reliable initial guess for the registration transformation, even when the images are strongly rotated and/or shiftec with respect to each other. This reliable initial guess allows us to use the wavelet decomposition as proposed by Mallat in [6] instead of the *à − trous* algorithm used in [5]. This is the reason for a considerable memory saving. Another minor difference is that while [5] appears to use only maxima of the DWT coefficients as matching points, we use both maxima and minima. Hence we have more points for the matching procedure and the LSM estimation.

In the next section we sketch our approach and give some details on the multiresolution analysis. In the third section we describe the strategy we use to estimate the initial guess of the transformation. Section four deals with the refinement of the initial guess. Section five presents some experimental results and the final section is devoted to the conclusions and further discussions.

2. MULTIRESOLUTION ANALYSIS

The registration algorithm consists of three blocks: the multiresolution analysis, the initial guess of registration transformation and the refinement of this transformation.

In the first block the multiresolution analysis is obtained with a DWT according to Mallat pyramidal scheme [6]. The DWT of the two images I and $\tilde{I}$ yields two series of images of decreasing size, the DWT coefficients,

$$I_1, I_2, ..., I_L \text{ and } \tilde{I}_1, \tilde{I}_2, ..., \tilde{I}_L,$$

and the two images of the residues of the DWT:

We want to acknowledge E. Ottaviani for helpul discussions

0-8186-7310-9/95 $4.00 © 1995 IEEE

$$I^* = DWT(I_L) = (DWT)^L(I)$$
$$\tilde{I}^* = DWT(\tilde{I}_L) = (DWT)^L(\tilde{I})$$

The transformed images are not overdetermined as in the *à − trous* algorithm,where no decimation occurs. The main advantage of using Mallat scheme is a considerable computational and memory saving which could result essential in remote sensing application where large images are used.

3. INITIAL GUESS ESTIMATION

Since our first experiments of registration with a multiresolution approach, we realized the success of the procedure depends strongly on the first step which grossly evaluates the parameters of the transformation. A wrong start can lead the search of the registration transformation completely off on a direction that may not be recovered at later steps when the parameters receive only minor adjustments. Particularly the algorithm described in [5] does not seem enough robust to face with large rotations between the two images . We overcame the problem by initially estimating the registration parameters with a clustering algorithm that uses as CPs the maxima and the minima of the residues of the DWT, I^* and $\tilde{I}^*$.
In this step it is assumed that the transformation between the images is roughly approximated by a rototranslation: $(\Delta x, \Delta y)$ represents the translation vector, c,s cosine and sine of the rotation angle. A complex bidimensional accumulator array $\mathbf{B}$ is used to determine a discrete approximation for the translation vector. The local minima and maxima over the residue images of the DWT are selected. For *every* two couples of maximum-maximum and/or minimum-minimum points $(P, \tilde{P})$ and $(Q, \tilde{Q})$, where P and Q refer to the image I^* and $\tilde{P}, \tilde{Q}$ to $\tilde{I}^*$, the parameters of the transformation which better approximates the rototranslation that links the segment $(\overline{PQ})$ and the segment $(\overline{\tilde{P}\tilde{Q}})$ are calculated,

$$c = ((\tilde{P}_x - \tilde{Q}_x) * (P_x - Q_x) + (\tilde{P}_y - \tilde{Q}_y) * (P_y - Q_y))/d$$
$$s = ((\tilde{P}_x - \tilde{Q}_x) * (P_y - Q_y) + (\tilde{P}_y - \tilde{Q}_y) * (P_x - Q_x))/d$$
$$\Delta x = 0.5(\tilde{P}_x + \tilde{Q}_x) - c * (P_x + Q_x) - s * (P_y + Q_y)$$
$$\Delta y = 0.5(\tilde{P}_y + \tilde{Q}_y) + s * (P_x + Q_x) - c * (P_y + Q_y)$$

where $d = (P_x - Q_x)^2 + (P_y - Q_y)^2$ and the pedices x, y indicate the x, y coordinates of the pixels.
The bin $\mathbf{B}(\Delta x, \Delta y)$ is incremented if the lenght of the two segment is grossly similar that is

$$|||P - Q|| - ||\tilde{P} - \tilde{Q}||| < D_0$$

where D_0 is a fraction (usually 1/10) of the average distance D between CPs within the images. Hence only the couples such that the lenghts of the vectors $\overline{PQ}$ and $\overline{\tilde{P}\tilde{Q}}$ are similar contribute to the evaluation of the initial guess of the registration transformation. Furthermore short vectors should account less then long vectors because the uncertaintites in the transformation parameters due to errors in the coordinates of the CPs are larger in the former case. Hence the bin increment should be proportional to the distance between the CPs if this is small and saturate when this reaches the average distance between CPs.
Finally we want to enforce "coherent" transformations (i.e. characterized by the same rotation angle).To achieve this goal, according to the scheme proposed in [7], we choose to increment the bin by a complex number with phase given by the rotation angle $\theta = arctan(s/c)$ and modulus equal to

$$min(D, 0.5 * (||P - Q|| + ||\tilde{P} - \tilde{Q}||))$$

When all the couples of segments have been examined, the bin with highest modulus is chosen to determine the translation vector of the initial guess of the registration transformation f. Its phase describes the rotation angle.
The initial guess supplied by this strategy is so reliable that allows us not only to make more robust the registration algorithm as proposed in [5] but also there is no need for the redundancy of the overdetermined *à − trous* to estimate the transformation and we could implement the pyramidal approach with all its related advantages.

4. TRANSFORMATION REFINEMENT

In the third, final, block the constraint of rototranslation is eliminated and the transformation f is refined trough a series of steps, at each of which a matching procedure produces the optimal transformation f_0 relating $f(I)$ to $\tilde{I}$, and f is replaced by $f_0 \circ f$. This refinement of the transformation f is carried out in a loop over the multiresolution levels l, from L down to L_{min}. At each step, the algorithm finds an affine transformation f_0 that relates the multiresolution level $f(I_l)$ with $\tilde{I}_l$.
The comparison of $f(I_l)$ and $\tilde{I}_l$ consists of estabilishing a correspondence between CPs represented by local maxima and minima on the first image and local maxima and minima on the second image. This correspondence is performed by matching extremal points according to their spatial positions in the images. The matching strategy is very simple and it is performed as

follows.First, the neighbouring size S is obtained as the least distance between two maxima or two minima in the same image. Next, two maxima (or minima), P and $\tilde{P}$, in $f(I_l)$ and $\tilde{I}_l$, respectively, are matched if P and $\tilde{P}$ lie within a distance S. The matched CPs represent the input data for LSM estimation of the parameters of the affine transformation f_0 obtained by minimizing

$$\sum \|f_0(P) - \tilde{P}\|^2$$

where the sum runs over the pairs $(P\tilde{P})$ of matched points. The refinement is repeated at each level of resolution as long as the L^2-distance between the grey levels of the two images $\tilde{I}$ and $f(I)$ decreases.

5. EXPERIMENTAL RESULTS

We use 512×512 eight-bit images. The classes of images used for the testing includes banknotes, mechanical pieces, aerial stereo pairs, multispectral and SAR images. In the experiment it is also possible to numerically rotate shift and deformed one of the images before the registration. Therefore we can compare an image with its rototranslated copy; in this case the registration algorithm should recover the rotation angle and the shift vector. This provides an indication of the quality of our registration algorithm.

The pyramidal transform (DWT) is carried out up to the fourth level. The initial estimate of the transformation using the complex clustering usually employs 20 to 30 points on each image and achieves a RMS between 10 and 20 grey levels, depending on the images. Nevertheless, the images are roughly well overlayed and this provides a suitable starting point for the refinement procedure based on DWT coefficients. The refinement of the transformation f is carried up from the scale level $L = 4$ down to the level $L_{min} = 3$. The details at finer scales turned out not to improve significantly the quality of the results. Indeed the information at these details turns out to be too much corrupted by the noise present in the images. At level four there are about one hundred CPs and at level three there are several hundreds. Of these about half are matched in pairs between the two images.

As a measure of the performance, we used the *RMS Grey Level Error (RMS-GLE)* that is the mean of the absolute value of the grey level difference between $f(I)$ and $\tilde{I}$. A high value indicates a poor registration. Another parameter is the average distance between CPs *RMS Distance Error (RMSDE)*,

$$\delta = \sqrt{\frac{1}{N} \sum \|f(P) - \tilde{P}\|^2},$$

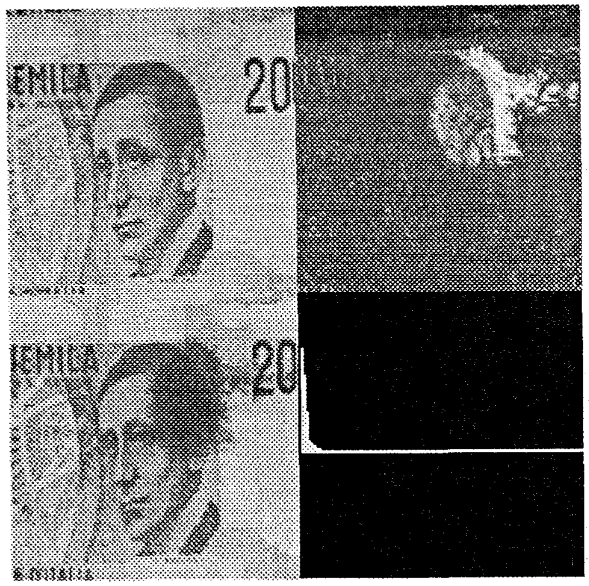

Figure 1: Results of the algorithm applied to the quality control of banknotes

where the sum runs over the CP pairs $(P, \tilde{P})$ and N is their number. Usually we employ *RMS-GLE* to evaluate the performances over two similar images as defect free notes, while *RMSDE* is more meaningful in the case of multispectral images. We obtained *RMS-GLE* below three grey levels for banknote images and more generally *RMSDE* resulted about one pixel.

Figure 1 and 2 show some results of our algorithm. In each figure the left quadrants contain the original images. The bottom (top in fig. 2) right quadrant shows the absolute difference between the images magnified by a factor five (ten in fig. 2) . The white curve in the top (bottom in fig. 2) right quadrant represents the histogram of absolute values of the grey level difference. referred to their overlapping region.

To verify the capability of the algorithm in dealing with large rotations, we tried to register a couple of image taken in two different spectral bands one of which had been numerically rotated and shifted. The rotation angle range from zero to 90 degrees and the shift vector was ten pixels in both directions. The results are reported in table 1 and show that the registration algorithm can recover the rotation angle within 0.05 degrees and the shift vector within 0.4 pixels.

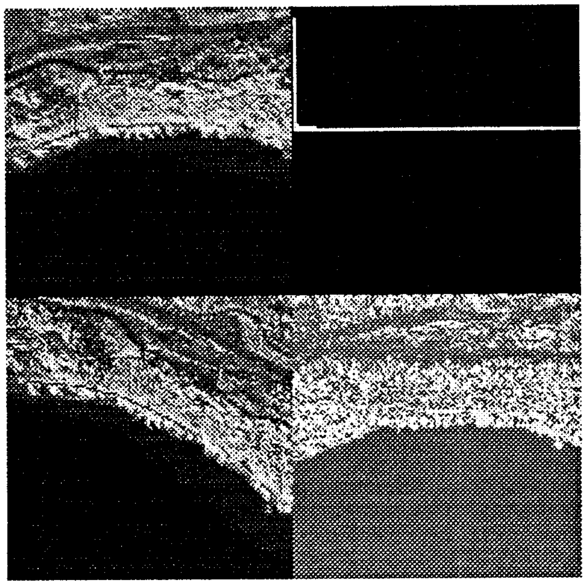

Figure 2: Results of the algorithm in the remote sensing domain

$\theta[deg]$	$\Delta\theta[deg]$	Δ Shift [pix]
0	0.02	0.00
1	0.03	0.28
2	0.01	0.29
4	0.02	0.33
6	0.01	0.28
10	0.02	0.10
15	0.02	0.16
20	0.00	0.34
30	0.00	0.10
45	0.03	0.20
60	0.02	0.06
80	0.01	0.38
90	0.04	0.15

Table 1: Results obtained with different rotation angles θ. In the second and third columns are presented the errors $\Delta\theta$, Δ Shift reported in the recovery of the rotation angle and of the shift vector.

6. CONCLUSIONS

We have presented a novel completely automatic image pair registration algorithm. The algorithm is based on a multiresolution approach and achieves very good performances. It represents a general tool suitable for multitemporal analysis of remotely sensed data and quality control purposes Our approach is close to that of [5] altough the two algorithms differ in a few respects. The main difference is that we use the residual images of the DWT and a clustering technique as basis for the initial estimate of the registration transformation. This strategy increases the robustness of the procedure providing a reliable initial guess even when the images are strongly rotated. Moreover it allows us to implement a pyramidal approach for the DWT instead of the *à − trous* scheme. By this means we achieve a cosiderable speed up in execution time and reduce the memory storage needs. The algorithm proved to be robust with respect to both correlated (defects on one of the notes) and uncorrelated noise (speckles on SAR images) and to sensor or illumination changes (multispectral images). The approach have been proved to be able to treat large rototranslations and it relies only on the hypothesis of small scale deformations (less than 10%).

Further improvements are conceivable by localizing the transformation f. This can be obtained by subdividing the image domain into patches and estimating the best transformation parameters in the individual patches. The parameters of the transformation f are then determined by interpolating among the parameters of the patches. A final enhancement can be obtained by carrying out the search of f in more complex groups, e.g. higher order polynomial transformations.

References

1. R.Kwok et al. *An automated system for mosaicking spaceborne SAR imagery*, Intern. J. Remote Sensing **5** (1990) 507.
2. A.A. Rodriguez and J.R. Mandeville *Image registration for the automatic inspection of printed circuits patterns using CAD reference data*, Mach. Vis. App. **6** (1993) 233.
3. A.Goshtasby, *Image registration by local approximation methods*, Image Vis. Comp. **6** (1988) 255.
4. I.Daubechies, "Ten Lectures on Wavelets", SIAM, Philadelphia PA, 1992.
5. J.P.Djamdji et al. *Geometrical registration of images: the multiresolution approach*, PE&RS **59** (1993) 645.
6. S. Mallat, *A theory for multiresolution signal decomposition*, IEEE PAMI. **11** (1989) 676.
7. T.J.Atherton, and D.J.Kerbyson, *The coherent circle Hough transform*, Proc. Brit. Mach. Vision Conf. 1993, p.269.