Proc. of SPIE Vol. 2034, Wavelet Applications in Signal and Image Processing, ed. A F Laine (Nov 1993) Copyright SPIE

A multiresolution image registration procedure using spline pyramids

Michael Unser and Akram Aldroubi

Biomedical Engineering and Instrumentation Program, Bldg 13, Room 3W13
National Center for Research Resources, National Institutes of Health
Bethesda, Maryland 20892 USA

Charles R. Gerfen

Laboratory of Cell Biology, Bldg 36, Room 2D-10
National Institute of Mental Health, National Institutes of Health
Bethesda, MD 20892 USA

ABSTRACT

We present an iterative multiresolution algorithm for the translational and rotational alignment of digital images. An image is represented by an interpolating spline. Coarser versions of this continuous image model are obtained by using spline approximations at various scales (polynomial spline pyramid). We use a coarse-to-fine updating strategy to compute the alignment parameters iteratively, using a variation of the Levenberg-Marquardt non-linear least-squares optimization method. This approach yields very precise image registration with subpixel accuracy. It is also much faster and more robust than a comparable single-scale implementation, because the resolution of the underlying image model is adapted to the step size of the algorithm.

1. INTRODUCTION

Image registration techniques play a crucial role in applications in which images with different translational and rotational parameters need to be compared or combined for further processing[2]. Examples in biomedical imaging include averaging techniques for noise reduction in high resolution electron-micrographs[4, 9], statistical analyses of series of PET images, multimodality imaging[15], and the alignment of autoradiographic slices for three-dimensional volume reconstruction. Most available algorithms for translational and rotational alignment use iterative correlation techniques[2, 5, 9]. As a result, they tend to be computationally quite intensive; their accuracy is also usually limited due to angular and translational sampling.

In this paper, we present an alternative technique that uses a continuous polynomial spline image model and takes advantage of the multiresolution structure of the underlying function spaces[14]. Instead of performing a systematic search over the parameter space as is done with most correlation techniques, the present algorithm uses a gradient-based least-squares optimization procedure. This approach is used in combination with a coarse-to-fine iteration strategy which substantially improves the overall performance of the algorithm. First, the computational load is reduced significantly since most iterations are performed at the coarser levels of the pyramid. Second, the procedure converges more rapidly since the spatial resolution of the underlying image model is adapted to error on the current parameter estimates. Finally, the algorithm is less likely to get trapped in a local optimum because the initial search is performed on a very coarse grid.

A key feature of this algorithm is its ability to provide image registration with subpixel accuracy; this property will be illustrated with some experimental examples. Other issues that will be considered are the behavior of the algorithm in the presence of noise, as well as the comparison with other techniques.

There are several reasons for using polynomial splines in this particular application. First, they provide a natural framework for formulating the image registration problem in the continuous domain. Second, polynomial splines

have a simple explicit form that makes them easy to manipulate. Finally, they are ideally suited for multi-scale processing because of their multiresolution properties.

2. STATEMENT OF THE PROBLEM

Image registration, as it is defined here, means finding a spatial rigid transformation (translation + rotation) that optimally maps an object image s onto a reference map r (or vice versa). In the following discussion, both images s and r will be represented by bidimensional spline functions of the continuous spatial variables x and y. The optimal mapping from one image onto another is defined as the one that minimizes the L_2-norm of the difference.

2.1 Image function spaces

$L_2(R^2)$ (which will be abbreviated to L_2) is the vector space of measurable, square-integrable bidimensional functions $s(\boldsymbol{x})$, $\boldsymbol{x}=(x,y)\in R^2$. L_2 is a Hilbert space whose metric $\|\cdot\|$ (the L_2-norm) is derived from the inner product

$$\langle r,s\rangle = \langle r(\boldsymbol{x}),s(\boldsymbol{x})\rangle = \iint_{(x,y)\in R^2} r(x,y)s(x,y)dxdy, \tag{1}$$

$$\|s\| = \sqrt{\langle s,s\rangle} = \sqrt{\iint_{(x,y)\in R^2} s(x,y)^2 dxdy}. \tag{2}$$

In our formulation, images are represented by polynomial splines at various scales. These functions are piecewise polynomials of degree n with the additional smoothness constraint that the polynomial segments are connected in a way that insures the continuity of the function and its derivatives up to order n-1. At the finer scale, there is exactly one knot per pixel and the image is represented by a spline that provides an exact interpolation of the initial gray level values[10]. The corresponding fine-grid spline function space $V_0 \in L_2$ can be defined as follows:

$$V_0 = \left\{ s_0(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in Z^2} c(\boldsymbol{k})\beta^n(\boldsymbol{x}-\boldsymbol{k}) : c(\boldsymbol{k}) \in l_2(Z^2) \right\}, \tag{3}$$

where the 2D generating function $\beta^n(\boldsymbol{x}) = \beta^n(x)\cdot\beta^n(y)$ is a tensor-product B-spline, and where $\{c(\boldsymbol{k})\}_{\boldsymbol{k}\in Z^2}$ is a 2D array of B-spline coefficients. The function $\beta^n(x)$ is the central B-spline[8] of degree n. It is generated by repeated convolution of a B-spline of degree 0

$$\beta^n(x) = \beta^0 * \beta^{n-1}(x) \tag{4}$$

with

$$\beta^0(x) = \begin{cases} 1, & x\in\left[-\frac{1}{2},\frac{1}{2}\right) \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

To specify the corresponding multiresolution image approximations, we consider the sequence of dyadic dilations of our basic spline space V_0. Specifically, V_i - the spline space at scale i - is defined as

$$V_i = \left\{ s_i(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in Z^2} c_i(\boldsymbol{k})\beta^n(\boldsymbol{x}/2^i - \boldsymbol{k}) : c_i(\boldsymbol{k}) \in l_2(Z^2) \right\}. \tag{6}$$

One scale increment corresponds to an enlargement of the basis functions by a factor of two; the spacing between the knots is also increased in the same proportion. Thus, the variable i may be interpreted as a coarseness index. For n odd, the following sequence of spline subspaces is nested

$$L_2 \supset \cdots \quad V_0 \supset V_1 \cdots \supset V_i \cdots \supset V_I \quad \cdots \supset \{0\}, \tag{7}$$

and this structure generates a multiresolution analysis of L_2 in the sense defined by Mallat[6]. The polynomial spline pyramid representation of a signal $s\in L_2$ corresponds to a sequence of fine-to-coarse minimum error approximations[14]

$$\{s_0 = P_0 s \in V_0, \cdots, s_i = P_i s \in V_i, \cdots, s_I = P_I s \in V_I\}, \tag{8}$$

where the operator P_i represents the orthogonal projection on V_i, and where $(I+1)$ is the depth of the pyramid. The various levels in this pyramid are entirely specified by their B-spline coefficients $c_i(\boldsymbol{k})$ as in (6), or, equivalently, by their sample values at the knots $s_i(2^i\boldsymbol{k})$ (cardinal representation). The former characterization is well suited for the evaluation of derivatives[12] and geometrical transformations, while the latter is more appropriate for the computation of error estimates. In either case, the pyramid coefficients are evaluated by repeated application of a digital lowpass filter followed by a decimation by a factor of two[14].

2.2 Image transformations

The translation operator $T_{\boldsymbol{a}} : L_2 \rightarrow L_2$ with displacement vector $\boldsymbol{a} = (a_x, a_y)$ is defined as

$$T_{\boldsymbol{a}}\, s(\boldsymbol{x}) = s(\boldsymbol{x} - \boldsymbol{a}). \tag{9}$$

Rotations are all specified with respect to a central pivot point (x_0,y_0); typically, the center of the image. The rotation operator $R_\theta : L_2 \rightarrow L_2$ with angular parameter θ is defined as

$$R_\theta\, s(\boldsymbol{x}) = s(\boldsymbol{x}_0 + \boldsymbol{R}(\theta)\cdot(\boldsymbol{x} - \boldsymbol{x}_0)) \tag{10}$$

where $\mathbf{R}(\theta)$ is the 2×2 rotation matrix

$$\boldsymbol{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}. \tag{11}$$

In general, translations and rotations do not commute. However, it is still possible to switch the order of application by using the identity

$$T_{\boldsymbol{a}} R_\theta s = R_\theta T_{\boldsymbol{b}} s \qquad \text{where} \quad \boldsymbol{a} = \boldsymbol{R}(\theta)\cdot\boldsymbol{b} \;\Leftrightarrow\; \boldsymbol{b} = \boldsymbol{R}(-\theta)\cdot\boldsymbol{a}. \tag{12}$$

Both $T_{\boldsymbol{a}}$ and R_θ are isometries from L_2 into L_2 in the sense that they preserve the L_2-norm

$$\forall s \in L_2, \quad \|s\| = \|T_{\boldsymbol{a}} s\| = \|R_\theta s\|. \tag{13}$$

We can use these last two properties to derive the identity

$$\left\|T_{\Delta\boldsymbol{a}} R_{\Delta\theta} s - R_{-\theta_1} T_{-\boldsymbol{a}_1} r\right\| = \left\|T_{\boldsymbol{a}_2} R_{\theta_2} s - r\right\| = \left\|s - R_{-\theta_2} T_{-\boldsymbol{a}_2} r\right\|, \tag{14}$$

with

$$\begin{cases} \boldsymbol{a}_2 = \boldsymbol{a}_1 + \boldsymbol{R}(\theta_1)\cdot\Delta\boldsymbol{a} \\ \theta_2 = \theta_1 + \Delta\theta. \end{cases} \tag{15}$$

This equation constitutes the basis for our iterative updating scheme. It will allow us to improve the current alignment parameters $\boldsymbol{a}$ and θ by considering small perturbations on the signal s.

Proof: We start with the left hand side of (14) and apply the norm-preserving inverse transformation $T_{\boldsymbol{a}_1} R_{\theta_1}$

$$\left\|T_{\Delta\boldsymbol{a}} R_{\Delta\theta} s - R_{-\theta_1} T_{-\boldsymbol{a}_1} r\right\| = \left\|T_{\boldsymbol{a}_1} R_{\theta_1} T_{\Delta\boldsymbol{a}} R_{\Delta\theta} s - r\right\|.$$

We then use (12) to reverse the order of the central translation and rotation

$$\left\|T_{\boldsymbol{a}_1} R_{\theta_1} T_{\Delta\boldsymbol{a}} R_{\Delta\theta} s - r\right\| = \left\|T_{\boldsymbol{a}_1} T_{\boldsymbol{R}(\theta_1)\cdot\Delta\boldsymbol{a}} R_{\theta_1} R_{\Delta\theta} s - r\right\| = \left\|T_{\boldsymbol{a}_1 + \boldsymbol{R}(\theta_1)\cdot\Delta\boldsymbol{a}} R_{\theta_1+\Delta\theta} s - r\right\|,$$

which proves the desired result. □

2.3 The multiresolution alignment problem

The registration problem can now be stated as follows. Let s and r be the continuous representations of the object and reference images, respectively. Then, the goal is to find the optimal translation and rotation parameters (a_0, θ_0) that minimize the error measured in the L_2-norm

$$(a_0, \vartheta_0) = \arg \min_{(a,\theta)} \|T_a R_\theta s - r\|^2, \tag{16}$$

where T_a and R_θ are the operators defined by (9) and (10), respectively.

Our approach is to cast this problem in a multiresolution framework and to consider the signal and reference approximations $s_i = P_i s$ and $r_i = P_i r$ in a spline pyramid with levels i=0,...,I. Starting at the coarser level I, we successively solve the sequence of problems for i=I down to 0

$$(a_i, \theta_i) = \arg \min_{(a,\theta)} \|T_a R_\theta s_i - r_i\|^2. \tag{17}$$

The optimum alignment parameters at scale i are determined iteratively by considering a perturbation $(\Delta a, \Delta\theta)$ on the previous solution. For this purpose, we use the equivalent form of the criterion given by the left hand side of (14), and update the new parameter values according to Eq. (15). This procedure ultimately yields the solution to our initial problem when i=0. The approximation problem at a given scale is solved by using a variation of the Marquardt-Levenberg least-squares optimization algorithm, which is iterative and requires the evaluation of the partial derivatives of s_i with respect to Δa, and $\Delta\theta$.

3. ALGORITHM DISCRETIZATION AND IMPLEMENTATION

3.1 Approximating the L_2-norm

The formal description of the algorithm in Section 2.3 uses L_2-norms which are difficult to compute in practice. To facilitate the implementation, we have chosen instead to use a sum approximation of the integral using the discrete samples of the underlying spline functions. This leads to the following discrete approximation formula of the quadratic error between the signals s and r at scale i

$$\|s_i - r_i\|^2 \cong 2^{2i} \sum_{(k,l)\in Z^2} \left[s_i^c(k,l) - r_i^c(k,l)\right]^2 = 2^{2i} \|s_i^c - r_i^c\|_{l_2}^2, \tag{18}$$

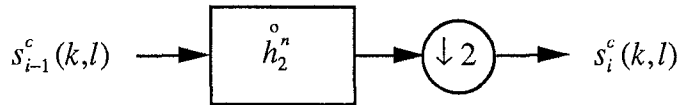
where the superscript "c" denotes the signal coefficients in the cardinal representation. These coefficients are obtained by sampling the underlying spline functions at the nodes of the pyramid

$$s_i^c(k,l) := s_i(x,y)\big|_{x=2^i k, y=2^i l}. \tag{19}$$

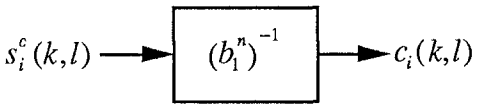
Note that the factor 2^{2i} in (18) provides the appropriate scale normalization. The use of this discrete approximation formula can be justified theoretically by the fact that there exist two positive constants A and B (that are close to one) such that the following inequality always holds

$$\forall r, s \in L_2, \quad 2^{2i} A \|s_i^c - r_i^c\|_{l_2}^2 \le \|s_i - r_i\|^2 \le 2^{2i} B \|s_i^c - r_i^c\|_{l_2}^2. \tag{20}$$

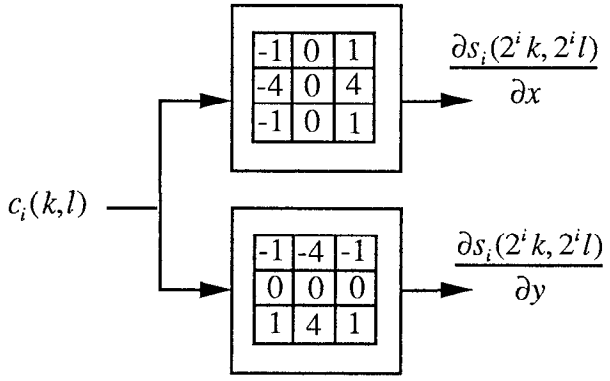
These constants may be determined from the extrema of the frequency response of the digital filter that performs the conversion between the cardinal and orthogonal spline representations[14]. For n=3, the value of these bounds is A=0.486 and B=1. The lower bound of this inequality represents the worst possible case; usually, the quality of the approximation (18) will be much better than that. Moreover, it can be shown that the distinction between the L_2 and l_2-norms in (18) progressively vanishes as n increases. This result comes as a consequence of the convergence properties of cardinal and orthogonal spline basis functions[1, 11].



(a) Reduction procedure for the spline pyramid



(b) Determination of the B-spline coefficients



(c) Determination of the image gradient

Fig 1 : Block diagram representation of various B-spline processing steps required for registration. The convolution masks in (c) are those associated with a cubic spline image model.

The approximation formula (18) is also valid when the two functions are translated and rotated with respect to the initial grid, although this situation is more difficult to analyze mathematically. Here too, the quality improves withincreasing n. Asymptotically, the approximation formula is exact because the corresponding spline approximation is bandlimited and because the underlying sinc basis functions are orthogonal[11].

3.2 Polynomial spline pyramid and B-spline processing

The polynomial spline pyramid representations of the signal s and r are obtained by filtering and decimation[14]. Let $s(k,l)$ denote the initial input image array. Then, the samples of the successive levels of the polynomial spline pyramid are computed iteratively as

$$\begin{cases} s^c_0(k,l) = s(k,l) \\ s^c_i(k,l) := s_i(x,y)\big|_{x=2^i k, y=2^i k} = [\overset{\circ}{h}^n_2 * s^c_{i-1}(k,l)]_{\downarrow(2,2)}, \end{cases} \tag{21}$$

where $[\cdot]_{\downarrow(2,2)}$ denotes a sub-sampling by a factor of two in the x and y directions, and where $\overset{\circ}{h}^n_2$ is the optimal prefilter for the cardinal representation. For more details on the specification of the optimal decimation filter, we refer to our previous work[14] (cf. Table III, p. 368). The sample values of the spline approximation at scale i will be used to specify the error criterion to be optimized.

Also required is the B-spline representation of these signals in order to implement rotations and translations of the reference image r, and to evaluate the partial derivatives of s for the optimization algorithm. Each level of the corresponding B-spline pyramid is obtained by digital filtering of the corresponding sampled representation:

$$c_i(k,l) = \left(b^n_1\right)^{-1} * s^c_i(k,l), \tag{22}$$

where $(b_1^n)^{-1}$ denotes the impulse response of the so-called direct B-spline filter. The most efficient way to perform this filtering is to use a recursive algorithm[10]. Finally, the components of the corresponding image gradient are evaluated by convolution with the appropriate masks[13]. A schematic representation of these various B-spline image processing operations is given in Fig. 1.

3.3 Non-linear least squares optimization

One iteration of the optimization algorithm can be described as an attempt to find the best parameter increment (Δa, $\Delta\theta$) such that

$$(\Delta a, \Delta\theta) = \arg \min_{(\Delta a, \Delta\theta)} \left\| T_{\Delta a} R_{\Delta\theta} s_i - r_i^* \right\|^2, \tag{23}$$

where $r_i^* = R_{-\theta} T_{-a} r_i$ represents a rotated and translated version of the reference according to the current parameter estimates. At scale i where the step size is 2^i, the discretized form of the error criterion to be minimized is

$$\varepsilon_i^2(\Delta a_x, \Delta a_y, \Delta\theta) = \sum_{k \in A_i} \left((T_{\Delta a} R_{\Delta\theta} s_i)(2^i k) - r_i^*(2^i k) \right)^2, \tag{24}$$

where A_i represents the region of interest in the image. Our optimization procedure uses a slight modification of the standard Levenberg-Marquardt method for non-linear least squares curve fitting. This technique requires the explicit knowledge of the partial derivatives of the function to be optimized. The derivatives that are needed here can be determined as follows

$$\left. \frac{\partial T_{\Delta a} R_{\Delta\theta} s_i(x,y)}{\partial \Delta a_x} \right|_{\Delta a = 0, \Delta\theta = 0} = -\frac{\partial s_i(x,y)}{\partial x}, \tag{25}$$

$$\left. \frac{\partial T_{\Delta a} R_{\Delta\theta} s_i(x,y)}{\partial \Delta a_y} \right|_{\Delta a = 0, \Delta\theta = 0} = -\frac{\partial s_i(x,y)}{\partial y}, \tag{26}$$

$$\left. \frac{\partial T_{\Delta a} R_{\Delta\theta} s_i(x,y)}{\partial \Delta\theta} \right|_{\Delta a = 0, \Delta\theta = 0} = (y - y_0)\frac{\partial s_i(x,y)}{\partial x} - (x - x_0)\frac{\partial s_i(x,y)}{\partial y}. \tag{27}$$

Since we know the values of $\partial s_i/\partial x$ and $\partial s_i/\partial y$ at the grid points (c.f. Section 3.2), these formulas can be used directly to calculate the gradient of the criterion to be optimized, as well as a first order approximation of the Hessian matrix. The best current parameter increments (Δa_x, Δa_y, $\Delta\theta$) are then computed using the recommended Marquardt rule[7]. The alignment parameters are finally updated as in (15), and the reference image is translated and rotated accordingly by resampling its B-spline representation on the appropriate grid. This procedure is iterated until the minimum is reached.

Because of the resampling step, the optimization is always performed around the point (Δa_x=0, Δa_y=0, $\Delta\theta$=0); this constitutes a variation of the standard Marquardt-Levenberg method. The benefit of this approach is a significant reduction in the number of computations because the partial derivatives of the function on the grid points, as well as the Hessian matrix, need only be computed once per scale.

Starting at the coarser scale I, the algorithm first mimics a standard steepest descent procedure, and progressively switches to an inverse-Hessian method as the minimum is approached. The transition to the next finer scale (i-1) occurs once convergence has been reached at resolution i. The method then operates in a quasi-Newton mode until it finally reaches the finer level of the pyramid. The advantage of this strategy is that the largest number of iterations is spent during initialization at the coarser scale; all subsequent finer scale updates usually require no more than one or two iterations since the previous solution already provides a very close estimate. The complexity per iteration is $O(N/4^i)$ where N denotes the total number of pixels in the image.

Initialization:
Compute the polynomial spline pyramid for the object s and reference r (cf. Fig. 1a)
Initialize $(\boldsymbol{a}, \theta)$

For $i=I$ *down to* 0
    Compute the B-spline coefficients for r_i and s_i (cf. Fig. 1b)
    Compute $\partial s_i/\partial x$ and $\partial s_i/\partial y$ at the grid point (cf. Fig. 1c)
    Compute the 3×3 Hessian matrix
    Repeat
        Translate and rotate the reference r_i according to current estimate $(\boldsymbol{a}, \theta)$
        Compute the gradient of the error
        Compute the current error $\varepsilon_i^2(\boldsymbol{a},\theta)$
        Compute the Marquardt update for $(\Delta\boldsymbol{a}, \Delta\theta)$
        Update $(\boldsymbol{a}, \theta)$
    Until convergence
End

Fig 2 : Summary of the iterative multiresolution registration procedure.

4. RESULTS

4.1 Cubic spline implementation

We used a cubic spline image model (n=3) to implement the algorithm. All the image processing operators described in Section 3.2 are separable; they were implemented by successive one-dimensional processing along the rows and columns.

The image translations and rotations were computed by resampling the B-spline expansion (6) at the new pixel locations. Since the basis functions are compactly supported, the image value at a particular location (x, y) is evaluated by partial summation of the B-spline contributions within a local neighborhood. The explicit cubic B-spline formula that was used for this evaluation is

$$\beta^3(x) = \begin{cases} 2/3 - x^2 + |x|^3/2, & 0 \le |x| < 1^2 \\ 1/3 - |x|^3/6, & 1 \le |x| < 2 \\ 0, & 2 \le |x|. \end{cases} \tag{28}$$

An outline of the complete registration algorithm, which was partially described in Sections 3.2 and 3.3, is given in Fig. 2. In this formulation, it is the reference r that is brought into registration with the object image s. If instead it is the object that needs to be aligned as in the experiments below, then the role of the images s and r should simply be interchanged .

4.2 Experiments

Several test images were generated by translating and rotating a standard reference using cubic spline interpolation. These transformed images were then brought back into registration using our iterative procedure. The behavior of the algorithm is best analyzed by monitoring the evolution of the normalized quadratic error criterion (24) as a function of the number of iterations. Two examples of such curves are shown in Fig. 3. The displacement parameters for both images (Lena, and an MRI scan) were $\mathbf{a}$=(15,15) and θ=15°, and the processing was performed using a five level

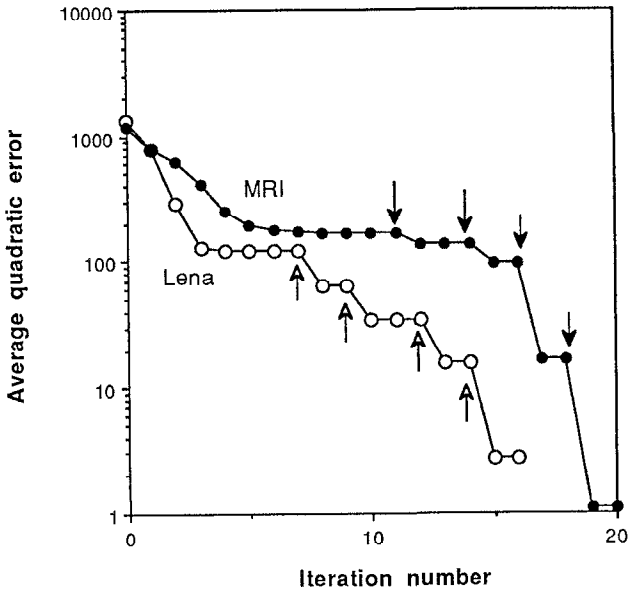


Fig 3 : Evolution of the quadratic error per pixel as a function of the number of iterations during the registration of the two test images "Lena" and "MRI".

pyramid (I=4). All parameters were initially set to zero. The arrows in Fig. 3 indicate the places of transition to the next scale. It can be seen that most iterations are spent at the coarsest level of the pyramid and that the initial improvement is quite progressive. Once the algorithm switches to the next finer scale, the convergence is almost instantaneous and this optimal behavior is maintained until the end. Note that a minimum of two iterations per scale was imposed in order to check convergence. In all the noise-free examples that we considered, the parameters were always estimated with a surprisingly good accuracy: Δa_x, $\Delta a_y < 10^{-3}$ pixels, and $\Delta\theta < 10^{-3}$ degrees. For comparison, we also tried a single scale optimization but this procedure did not converge unless the initial guess was within 2 pixels of the true solution. Even in this very favorable case, the single-scale approach took much longer than the multiresolution implementation.

We also tried to determine how good the initial guess needs to be so that the algorithm converges to the correct solution. Although these ranges are somewhat data dependent, we were able to identify some general guidelines. We found that the algorithm could generally tolerate an initial angular error of 20 degrees, and some times even more (30-35 degrees), depending on the quality of the initial translational estimates. The algorithm is more robust to translational errors because of its multiscale structure. We found that the least squares procedure could usually tolerate errors of the order of 1 to 2 pixels at the coarser scale. Therefore, our rule of thumb for a safe behavior is that the initial error should not exceed the following bounds: $\|\Delta a\|_\infty < 1.5 \cdot 2^I$, $|\Delta\theta| < 20°$, where (I+1) is the depth of the pyramid.

Next, we investigated the performance of the algorithm in the presence of noise. For this purpose, we considered the same experiment as before using the "Lena" picture while adding gaussian white noise with a standard deviation σ to both the reference and transformed object images. The results are summarized in Table I. The first observation is that the parameter estimates become less accurate for increasing levels of noise, which is to be expected. However, even under the most adverse conditions (SNR=0dB), the quality of the estimation is still very good (Δa_x, $\Delta a_y < 0.2$ pixels, and $\Delta\theta < 0.1$ degrees). Moreover, the presence of noise has very little effect on the estimation of the parameters at the coarsest level (c.f. 2nd row of Table 1). The corresponding graphs of the error as a function of the iteration number are shown in Fig. 4.

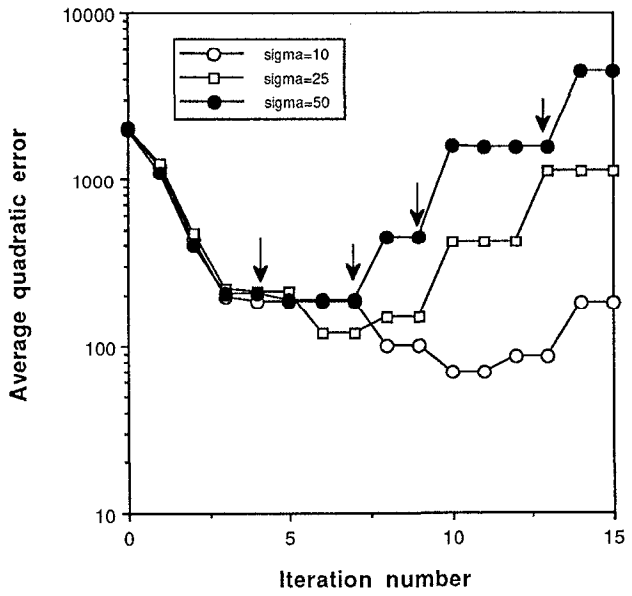


Fig 4 : Evolution of the quadratic error per pixel as a function of the number of iterations during the registration of the "Lena" under various noise conditions.

Signal-to-noise ratio	Coarse level estimate (I=4) $(a_{xI}, a_{yI}, \theta_I)$	Final estimate $(a_{x0}, a_{y0}, \theta_0)$	Final error ε_0^2
σ=0 (SNR→+∞)	(10.6333, 16.1656, 14.8806)	(14.9996, 15.0003, 14.9998)	1.075
σ=10 (SNR=7dB)	(10.5562, 16.1839, 14.9153)	(14.9944, 15.0072, 15.0027)	179.745
σ=25 (SNR=3dB)	(10.9498, 15.9976, 14.8332)	(14.9494, 15.0555, 15.0066)	1110.35
σ=50 (SNR=0dB)	(11.5896, 14.8966, 14.2377)	(14.8115, 15.1782, 15.0636)	4426.71

Table 1 : Results of the registration of the test image "Lena" with various levels of noise. The true parameter values are **a**=(15,15) and θ=15°.

Note that in the noisy case, the error ε_i^2 may increase as one jumps to the next finer scale, although the algorithm still converges to the correct solution. This behavior is not too surprising if one considers that the noise contribution at scale i is attenuated by a factor 2^i so that its effect is almost imperceptible at the coarser levels of the pyramid. This makes the algorithm very robust to measurement noise; its convergence properties are hardly affected at all.

4.3 Discussion

Our experimental results indicate that the algorithm is well behaved and that it is a capable of very precise image registration with subpixel accuracy. We have also shown that the multi-scale formulation is truly beneficial in the sense that it makes the algorithm much more robust. It also reduces the computational load substantially. For noise-free images, the algorithm usually provides an (almost) exact alignment with a very small residual error. The reason why the error is not exactly zero is that the underlying spline image model is not truly shift-invariant; in other words, re-sampling may result in a slight loss of information, although this effect tends to disappear with higher order splines[1, 11].

In the presence of noise, there is a certain uncertainty on the parameter estimates. Yaroslavky[16] has proposed a theoretical analysis for the translational case. In particular, he derived lower covariance bounds that take the form of a ratio between the noise variance and a shape factor that is image-dependent. In practice, similar estimates of the error covariance can be obtained by taking the product of the residual error and the inverse of the Hessian matrix that are both provided by the algorithm.

The present method has at least two advantages over standard correlation techniques[2]. First, it is extremely accurate because of the use of an underlying continuous image model. Correlation and Fourier techniques are unable to perform at the subpixel level because of obvious sampling limitations. Second, the present formulation offers some flexibility in the specification of the error criterion to be minimized which is defined over a region of interest A_0 (c.f. (24)), which can be user-specified. It is thereby possible to mask out certain structures in the image that are not well reproduced from one instance to another. Alternatively, one may introduce a spatial weighting function to emphasize certain important characteristics. Such modifications are not so easy to incorporate into Fourier-based correlation techniques[3, 9] which optimize a global criterion that is defined over the entire image.

There are also ways of making the present algorithm more robust so that it can handle cases in which there is only a partial overlap. One approach is to subdivide the image into P regions of equal size. At the beginning of each iteration, a certain number of regions with potential outliers are identified based on the magnitude of their partial error contributions. These regions are then left out, and the estimation of the gradient and Hessian matrix is performed on the remainder of the set. This strategy may be appropriate for dealing with situations in which certain parts of the object (unknown a priori) may be missing, or when there is an incomplete match between the object and reference.

5. CONCLUSION

We have described a multiresolution registration procedure for the translational and rotational alignment of digital images. The main features of this approach can be summarized as follows:

- The algorithm provides extremely accurate estimates of the alignment parameters. In a noise-free environment, the alignment is essentially perfect. This is a consequence of a consistent design that uses a continuous polynomial spline image model.
- The multiresolution algorithm is superior to its single-scale version in a number of respects. First, it converges much more rapidly because the spatial resolution of the underlying image model is adapted to the step size of the algorithm. Second, it is much less likely to get trapped in a local minimum because of the smoothing effect of the pyramid. Finally, it is much faster because most iterations are performed at the coarsest resolution.
- The method is iterative and is therefore somewhat sensitive to the quality of the initial guess. However, we found that the algorithm could easily handle angular discrepancies as large as 20 degrees, and translational errors of the order of 2^I pixels, where $(I+1)$ is the depth of the pyramid.
- The algorithm is very robust to measurement noise. Even when the signal-to-noise ratio is as low as 0dB, it is still possible to obtain parameter estimates within a few tenths of a pixel, and less than one tenth of a degree. Noise has almost no effect on convergence because of the noise reduction properties of spline pyramids.

The present method should therefore provide an attractive alternative to more traditional correlation techniques, especially when subpixel accuracy is desirable, or when the registration should be performed over a region of interest of arbitrary shape.

6. REFERENCES

1. A. Aldroubi, M. Unser and M. Eden, "Cardinal spline filters : stability and convergence to the ideal sinc interpolator", *Signal Processing*, Vol. 28, No. 2, pp. 127-138, August 1992.
2. L.G. Brown, "A survey of image registration techniques", *ACM Computing Surveys*, Vol. 24, No. 4, pp. 325-376, 1992.
3. E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-9, No. 5, pp. 700-703, September 1987.
4. J. Frank, "Averaging of low-exposure electron micrographs of non-periodic objects", *Ultramicroscopy*, Vol. 1, pp. 159-162, 1975.
5. J. Frank, A. Verschoor and M. Boublik, "Computer Averaging of Electron Micrographs of 40S Ribosomal Subunits", *Science*, Vol. 214, pp. 1353-1355, 1981.
6. S.G. Mallat, "A theory of multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-11, No. 7, pp. 674-693, 1989.
7. W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes*, Cambridge University Press, Cambridge, GB, 1986.
8. I.J. Schoenberg, "Contribution to the problem of approximation of equidistant data by analytic functions", *Quart. Appl. Math.*, Vol. 4, pp. 45-99, 112-141, 1946.
9. B.L. Trus, M. Unser, T. Pun and A.C. Steven, "Digital image processing of electron micrographs: the PIC system II", *Scanning microscopy*, to appear.
10. M. Unser, A. Aldroubi and M. Eden, "Fast B-spline transforms for continuous image representation and interpolation", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No. 3, pp. 277-285, March 1991.
11. M. Unser, A. Aldroubi and M. Eden, "Polynomial spline signal approximations : filter design and asymptotic equivalence with Shannon's sampling theorem", *IEEE Trans. Information Theory*, Vol. 38, No. 1, pp. 95-103, January 1992.
12. M. Unser, A. Aldroubi and M. Eden, "B-spline signal processing. Part I : theory", *IEEE Trans. Signal Processing*, Vol. 41, No. 2, pp. 821-833, February 1993.
13. M. Unser, A. Aldroubi and M. Eden, "B-spline signal processing. Part II : efficient design and applications", *IEEE Trans. Signal Processing*, Vol. 41, No. 2, pp. 834-848, February 1993.
14. M. Unser, A. Aldroubi and M. Eden, "The L_2 polynomial spline pyramid", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 15, No. 4, pp. 364-379, April 1993.
15. P.A. van den Elsen, E.-J.D. Pol and M.A. Viergever, "Medical image matching: A review with classification", *IEEE Engineering in Medicine and Biology*, Vol. 12, No. 1, pp. 26-39, March 1993.
16. L.P. Yaroslavsky, "The theory of optimal methods for localization of objects in pictures", in: E. Wolf, ed., *Progress in Optics*, Elsevier, Amsterdam, to appear.