Robust Multiscale Affine 2D-Image Registration through Evolutionary Strategies

Héctor Fernando Gómez García[1], Arturo González Vega[1], Arturo Hernández Aguirre[1], José Luis Marroquín Zaleta[1], and Carlos Coello Coello[2]

[1] Center for Research in Mathematics, Department of Computer Science, Guanajuato, Gto. 36240, MEXICO
hector,gonzart,artha,jlm@cimat.mx
[2] CINVESTAV-IPN Sección de Computación, México, D.F. 07300, MEXICO
ccoello@cs.cinvestav.mx

Abstract. We propose a robust methodology based on multiscale analysis, affine transforms, and evolutionary strategies for solving the image registration problem. The approach is found robust for the affine registration of medical images.

1 Introduction

The goal of image registration is to find the best correspondence between images of the same scene. The intuitive approach to this problem is to find "relevant features" on the images that can be used to bring them into correspondence. As noted in [5], finding relevant features is one of three main components of any image registration problem; the second is similarity metric, and the third is search space and strategy. Our approach to image registration can be described around these components as follows: 1) feature space is not created, nor induced or searched; images are sampled and a few points are used for matching. 2) It uses a metric based on pixel intensity to measure image correspondence, and 3) the search space is kept small by subsampling the images whereas the optimization mechanism is implemented through evolutionary strategies.

2 The image registration problem

For the sake of sufficiency of the article we describe the registration problem (defined elsewhere in the literature). Assume two images $I_1(x, y)$ and $I_2(x, y)$ are available from the same object but the object changes position from one to other. The images are 2-dimensional arrays with some intensity value at every pixel position (x, y). The image registration problem is *to find the mapping between two images* I_1 and I_2 that gives the best correspondence. Equation 1

$$I_2(x, y) = I_1(f(x, y)) \tag{1}$$

is the formal registration model where function $f : I_1(x, y) \rightarrow I_2(x, y)$ performs the mapping between images. Approximations to f can be constructed by some transformations: affine and projective amongst several. No transform applies to all problems thus in choosing a suitable transform it is advisable to consider the sources of misregistration. They are generally due to sensor noise, sensor type, and changes in scene conditions [8]. This paper describes an image-to-image registration technique without use of any knowledge about the sensors. Images are taken with the same instrument but (simulated) from different positions, thus the only source of misregistration is related to changes in scene conditions. Note that the image registration problem is clearly a function approximation one. That is, f is unknown but it will be approximated through affine transformations.

2.1 Affine transforms

An affine transform is a linear transform composed of the following geometric transformations: translation, rotation, scaling, stretching, and shearing [7]. As noted affine transforms are sound basis for our mapping function approximation problem since the source of misregistration can be tackled as follows: distortions due to different sensor orientation are corrected by translation and rotation, changes in altitude are corrected by scaling, and stretching and shearing correct distortions due to changes in the viewing angle [5]. A useful subset of affine transform that combines rotation and translation is called *rigid-body transform*. In Equation 2, (x_2, y_2) is the transformed coordinate (x_1, y_1) after translations (t_x, t_y), rotation by angle θ, and scaling by factor s (an affine transform is only linear when translation is zero [5]).

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + s \begin{bmatrix} cos\ \theta & -sin\ \theta \\ sin\ \theta & cos\ \theta \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \tag{2}$$

The general $2D$ affine transformation (the basis of our approach) is expressed as shown in Equation 3. The matrix $\begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$ accounts for rotation, scaling, stretching, and shearing.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{1,3} \\ a_{2,3} \end{bmatrix} + \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \tag{3}$$

Hence we should understand the affine registration problem as the problem of finding the parameter set $\{a_{1,1}, a_{1,2}, a_{1,3}, a_{2,1}, a_{2,2}, a_{2,3}\}$ for the affine transform that best mimics the function $f : I_1(x, y) \to I_2(x, y)$. Provided function approximation is sound we still need to address the question of how well two images match. The similarity measure used affects the matching quality. One popular similarity measures is normalized cross-correlation [14]; since its computation is expensive in this paper we use a simpler and reliable well known approach: the sum of absolute difference of pixel intensities.

3 Related work

Image registration has been approached from a large variety of techniques, we should only mention in this section those articles closely related to this article. The work of Fitzpatrick and Grefenstette [6] is one of the first works on registration of medical images based on Genetic Algorithms. Brown [5] noted that probabilistic methods are more suitable for registration and segmentation of medical images, thus less than 7% of the methods (95) accounted by Maintz [11] use a form of evolutionary algorithm as optimization tool. Ankenbrandt and Buckles [1] use genetic algorithms for scene recognition, Bhanu and Sungkee [4] describe several methods based on evolutionary techniques for image segmentation. Roberts and Howard [12] use genetic programming for orientation detection, Ross et. al. [13] also use genetic programming but for edge detection; Bhandarkar [3] recently compared several techniques for image segmentation using evolutionary computation. Louchet [10] applied evolutionary strategies to stereovision.

4 Multiscale representation

A multiscale representation, also called Gaussian pyramid, is a set of images generated by the successive application of smoothing and subsampling operations over a source image of dimensionality d. At each step the new image contains only $\frac{1}{2^d}$ pixels of the previous image [9]. A typical multiscale pyramid is shown in Figure 1. The successive application of smoothing and subsampling operations

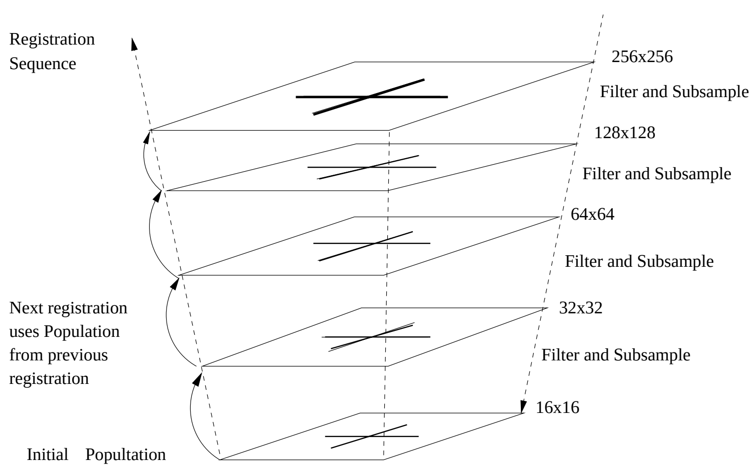

Fig. 1. Multiscale Gaussian pyramid for Image registration

helps to eliminate unnecessary details while keeping important features. These features are very important for the first iteration of our method at the lowest level of resolution (bottom of pyramid). Inherent to multiscaling is the reduction in the size of the image, a property that reduced processing time without altering precision.

5 The multiscale affine image registration method

Our approach to image registration is based on a multiscale representation. The sequential application of smoothing and subsampling operations is performed in a top-down fashion. Then registration is performed bottom-up. The whole registration process is as follows and it is shown in Figure 1

- Top-Down step. Apply a smoothing and subsampling procedure K times to both images I_1 and I_2. A set of K subsampled images is computed and stored (K of each image). The smoothing procedure (to prevent alising) is a low-pass filter implemented by image convolution with a Gaussian kernel ($\sigma = 0.5$).
- Bottom-Up step. Register the images at the bottom of the pyramid (lowest resolution). Initial population is seeded with individuals mutated out of the identity transform, that is, individuals are mutations of the identity matrix, and the zero translation vector. Once the $(\mu + \lambda)$-ES algorithm reaches a nominal fitness value or number of iterations, the registration process is repeated but this time using the images at the immediate level above. For seeding the next initial population the object variables of the best individual of the previous registration step are mutated. Control variables are generated anew. At any registration level no more than 256 sampled pixels are used to compute the fitness value. Notice in Figure 1 the image at the bottom has only 256 pixels, but the immediate above must be sampled because it has 1024 pixels (although the image itself was derived by the sampling procedure of the top-down step, it is again sampled to compute the fitness function).

A $(\mu + \lambda)$ Evolutionary Strategy is used for searching and optimization of the six real variables that control the general affine transform. Crossover operation for control and object variables is generalized intermediate, mutation is uncorrelated (no rotation angles), in accordance with that version of the algorithm as described by Bäck [2]. Population size does not change during the process. No knowledge (landmarks) from the image has to be derived or located during the process, our method uses only 256 sample points equally spaced and distributed over the image. Thus, about 0.4 of image pixels are used for registration in our experiments.

Fitness function is based on the similarity measure "absolute difference of intensities", as follows:

$$fitness = \frac{1}{1 + \frac{1}{N}\sum_{\{x,y\}\in\Omega} |I_1(f(x,y)) - I_2(x,y)|} \tag{4}$$

Where Ω is the set of sample points, N is the cardinality of Ω (256), and $f(x,y)$ the required transformation. The fitness function takes values in $[0,1]$ to represent $[nomatch \rightarrow perfectmatch]$. Another strategy to create the set Ω is to take random samples with uniform probability distribution. The set can be generated anew between 3 and 5 times per level, improving the registration ability of the algorithm over noisy images. Since an affine transformation AT maps pixel's integer coordinates of image I_2 into real coordinates on image $I1$, a cubic spline

interpolation procedure IP is used to calculate the proper intensity value of each transformed coordinate.
An important issue related to the fitness function is the quality of the registration implied by the fitness value (Equation 4). In Table 1 we show the value of the summation term for several fitness function values that are used in our experiments. It is clear that a change of .01 in the fitness value implies a change of 2.58 in the summation term. Since the summation term accounts for 256 pixels and each pixel intensity lies in the interval $[0, 255]$, a high value of fitness implies high image matching. Yet, moderate values of fitness also imply good matching.

Fitness Value	Σ term (256 pixels)
1.0	0.0
.99	2.58
.98	5.22
.97	7.91
.96	10.6
.95	13.4
.94	16.3

Table 1. Relation fitness function - summation of intensity error

6 Experiments

After some trials we found that a (100 + 150)-ES finds solutions with average approximation error no greater than 10^{-4} (on each objective variable), in average time of 200 seconds. A $(250 + 50)$-ES finds solutions with average approximation error no greater than 10^{-2}, in average time of 30 seconds. For all experiments reported in the following sections, the population parameters are: $\mu = 250$, and $\lambda = 50$. Since to each objective variable corresponds one control parameter, an individual is composed by six control variables (Equation 3), and six objective variables. Control parameters (variance) for matrix coefficients are $\sigma_{1,1} = \sigma_{1,2} = \sigma_{2,1} = \sigma_{2,2} \geq 0.1$, and for translation coefficients $\sigma_{1,3} = \sigma_{2,3} \geq 1.0$ Every $\sigma < 5.0$ We used a PC computer with Xeon processor running at 1.7 Ghertz with 1 Gbyte of memory; all algorithms are programmed in C++. In the sequel we describe three experiments.
The goal of the first experiment is to verify the ability of the method to consistently reach the optimum and to measure the error. In the second experiment we contrast overall convergence. Two of the three test images are known to be hard to register: a slice of MRI of the human brain, and an angiogram.

6.1 Robustness and consistency of the method

This experiment is designed to measure the approximation error on each of the six parameters. Therefore, a fixed set of affine transformation coefficients (see Equation 3) is randomly generated and used in all 70 runs. The matrix coefficients (that imply rotation, scaling and shearing) are: 0.819684, 0.089627, 0.16434, 0.40437 The coefficients denoting translations over the axes are: 25.532335, and 39.6115 For these experiments we used a MRI image of the brain, as shown in Figure 2

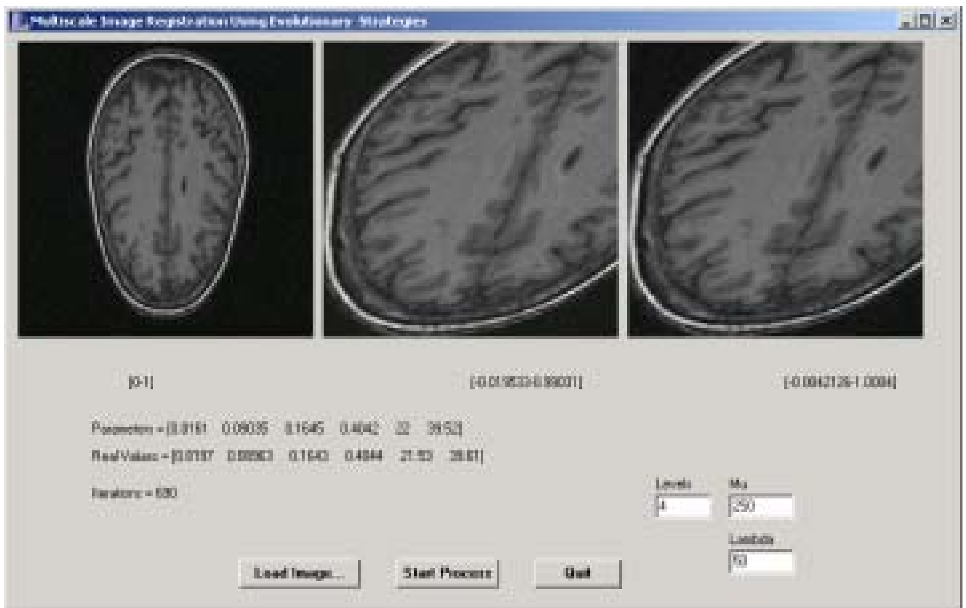

Fig. 2. MRI of brain used in experiment 1

In Table 2 we resume the robustness and consistency ability of the method to reach the highest fitness value (1.0). Out of 70 runs:

- 24 runs reached $fitness \geq 0.99$
- 29 runs reached $fitness \geq 0.98$
- 44 runs reached $fitness \geq 0.97$
- 47 runs reached $fitness \geq 0.96$
- 52 runs reached $fitness \geq 0.95$
- 63 runs reached $fitness \geq 0.94$

The rest of the experiments (7) had $fitness < 0.94$ Information in Table 2 shows that for fitness value of 0.94 the coefficients begin to deteriorate. To the human eye this level of error is not apparent, yet the data for that row indicates that the translation coefficients differ in six pixels (average). Thus, for the experiments of the next section we take a fitness value of up to 0.95 as the minimum required to indicate “good registration”.

Interval	Measure	**0.8196**	**0.0896**	**0.1643**	**0.4043**	**21.5323**	**39.6115**
0.99-1.0	Avg.	0.8194	0.0904	0.1649	0.4049	21.4508	39.4623
0.99-1.0	Std.Dev.	0.0031	0.0041	0.0025	0.0010	0.9496	0.3939
0.98-1.0	Avg.	0.8190	0.0884	0.1660	0.4050	21.7839	36.6115
0.98-1.0	Std.Dev.	.0031	.0064	.0038	.0010	1.2821	0.5188
0.97-1.0	Avg.	0.8171	0.0830	0.1694	0.4052	22.9012	38.9512
0.97-1.0	Std.Dev.	0.0045	0.0113	0.0072	0.0010	2.3640	0.8696
0.96-1.0	Avg.	0.8164	0.0825	0.1701	0.4050	23.1073	38.8960
0.96-1.0	Std.Dev.	0.0056	0.0126	0.0079	0.0022	2.7110	0.9956
0.95-1.0	Avg.	0.8154	0.0869	0.1703	0.4038	22.4673	39.1022
0.95-1.0	Std.Dev.	0.0063	0.0261	0.0093	0.0057	4.3408	1.5174
0.94-1.0	Avg.	0.7792	0.0803	0.1886	0.4032	27.9592	37.1533
0.94-1.0	Std.Dev.	0.1547	0.0495	0.0491	0.0167	15.1011	4.6558

Table 2. Error in transformation parameters by interval

6.2 Overall convergence experiments

In this set of experiments we ran 20 random registrations with each one of the test images: MRI-brain (Figure 2), Diana (Figure 4), and angiogram (Figure 3), and checked for convergence. As explained before, if fitness reaches 0.95 the registration is counted as good. The same problems were also registered using the Gauss-Newton optimization algorithm. In this gradient descent technique we also used the same procedure and fitness function. Table 3 shows the number of successful runs (convergence) of each algorithm for every image.

Test image	Multiscale Evolutionary	Gauss-Newton
MRI-Brain	19/20	11/20
Angiogram	20/20	6/20
Diana	17/20	1/20

Table 3. Convergence in 20 random experiments for two techniques

7 Discussion and conclusions

From the first set of experiments we have shown our method consistently finds the solutions with good accuracy. The second set of experiments is a clear proof of the robustness of the method. The Gauss-Newton based method, as any other based on gradient descent, is prone to fall in local minima. The evolutionary approach is a strategy that shares global knowledge among the individuals of the population, thus convergence to the solution occurs with high probability.

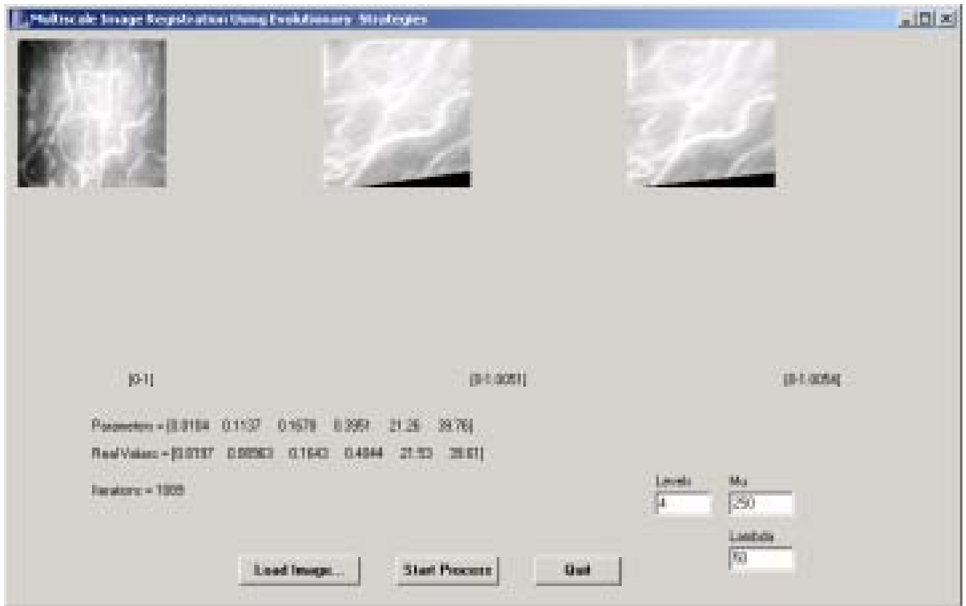

Fig. 3. Evolutionary multiscale registration of an angiogram

Fig. 4. Evolutionary multiscale registration of "Diana"

In general our approach contradicts several authors [5] who have found weak properties in evolutionary methods for image registration.

Future work A fitness function based on normalized cross-correlation (and other approaches) will be studied. Other problem worth of studying is the registration of images with noise. The combination of non-linear transforms with evolutionary techniques is a promising approach to image registration.

Acknowledgments The first two authors acknowledge support from CONACyT through a scholarship to complete the Master in Science program at CIMAT. The third author acknowledges partial support from CONCyTEG project No. 01-02-202-111 and CONACyT No. I-39324-A. The fourth author acknowledges partial support from CONACyT project No. 34575-A. The last author acknowledges support from CONACyT project No. NSF-CONACyT 32999-A.

References

1. C.A. Ankenbrandt, B.P. Buckles, and F.E. Petry. Scene recognition using genetic algorithms with semantic nets. *Pattern Recognition Letters*, 11:285–293, 1990.
2. Thomas Back. *Evolutionary Algorithms in Theory and Practice*. Oxford University Press, New York, 1996.
3. S.M. Bhandarkar and H. Zhang. Image segmentation using evolutionary computation. *IEEE Transactions on Evolutionary Computation*, 3(1):1–21, April 1999.
4. B. Bhanu and Sungkee Lee. *Genetic Learning for Adaptive Image Segmentation*. Kluwer Academic Publishers, Massachusetts, 1994.
5. Lisa G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, December 1992.
6. J.M. Fitzpatrick, J.J. Grefenstette, and D. Van Gucht. Image registration by genetic search. *Proceedings of IEEE Southeastern Conference*, pages 460–464, 1984.
7. B. Jahne. *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*. Springer Verlag, Berlin, 1997.
8. Shang-Hong Lai and B. C. Vemuri. Reliable and efficient computation of optical flow. *International Journal of Computer Vision*, 29(2):87–105, 1998.
9. T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, Netherlands, 1994.
10. J. Louchet. Using an individual evolution strategy for stereovision. *Genetic Programming and Evolvable Machines*, 2(2):101–109, March 2001.
11. J. Maintz and M. Viergever. A survey of medical image registration. *Journal of Medical Image Analysis*, 2(1):1–36, 1998.
12. S.C. Roberts and D. Howard. Genetic programming for image analysis: Orientation detection. *Proceedings of the GECCO Conference 2000*, pages 651–657, 2000.
13. B.J. Ross, F. Fueten, and D.Y. Yashkir. Edge detection of petrographic images using genetic programming. *Proceedings of the GECCO Conference 2000*, pages 658–665, 2000.
14. J. C. Tilton. Comparison of registration techniques for GOES visible imagery data. *Proceedings of IRW, NASA GSFC*, pages 133–136, 1997.