

Image Registration Using Markov Random Coefficient Fields

Edgar Román Arce-Santana and Alfonso Alba

Facultad de Ciencias
Universidad Autonoma de San Luis Potosi
San Luis Potosi, Mexico
{arce, fac}@fciencias.uaslp.mx

Abstract. Image Registration is central to different applications such as medical analysis, biomedical systems, image guidance, etc. In this paper we propose a new algorithm for multi-modal image registration. A Bayesian formulation is presented in which a likelihood term is defined using an observation model based on linear intensity transformation functions. The coefficients of these transformations are represented as prior information by means of Markov random fields. This probabilistic approach allows one to find optimal estimators by minimizing an energy function in terms of both the parameters that control the affine transformation of one of the images and the coefficient fields of the intensity transformations for each pixel.

Keywords: Image Registration, Markov Random Fields, Bayesian Estimation, Intensity Transformation Function.

1 Introduction

Image registration is the alignment of images that may come from the same or different source. This task is very important to many applications involving image processing or analysis such as medical analysis, biomedical systems, image guidance, depth estimation, and optical flow. A special kind of registrations is called Multimodal Image Registration, in which two or more images coming from different sources are aligned; this process is very useful in computer aided visualization in the medical field.

In the literature, there are basically two classes of methods to register multimodal images: those based on features such as edge locations, landmarks or surfaces [6][7][11], and those based on intensity [1][19][4][16]. Within the intensity methods there are two popular ones. Partitioned Intensity Uniformly (PIU) [19][5], proposed by Woods et al, is one of them. In this method it is assumed that uniform regions in one of the images correspond to regions, also uniform, in the other one. To achieve the registration, a corresponding measure is established based on the statistical characteristics of both images. The goal of this method is to use this measure to minimize the variance of intensity ratios. The other method that has shown good results is the registration based on mutual

information (MI), proposed by Viola et al [18]. In this method, statistical dependencies between images are compared, establishing a metric based on the entropy of each image and the join entropy. Even though the method is theoretically robust, it is complicated to implement and requires vast computational resources. Another drawback of MI is that it completely ignores spatial information such as edges or homogenous regions.

A method related to the work proposed in this paper is presented in [10]. It focuses only on elastic registration of multimodal images; it uses an iterative scheme that iterates between finding the coefficients of polynomial intensity transformations and registration using the demons method [17]. This method makes the assumption that there are at most two functional dependencies between intensities. This restriction limits its applications since there are cases, as those found in medical imaging, where inhomogeneity and noise are presented in both images to register.

In this work, we present a more general registration method, in which a probabilistic model permits the characterization of the image registration by means of linear intensity transformation functions. Rigorously based on Bayesian estimation, the main goal of this method is to establish the parameters of the affine transformation, and at the same time, determine in a probabilistic framework the coefficient values of these linear functions for each pixel to achieve the image registration. These transformations have the purpose to estimate the adequate intensity changes that match the intensity values between the images. In this approach, the coefficients of the linear intensity transformations (labeled MRFCF) are represented as Markov Random Fields (MRF)[2], giving in this way the prior information about the homogeneity of the intensity changes.

The paper is organized as follows: in Section 2, we give an introduction to MRF and present the Bayesian framework of image registration using affine transformation and MRFCF; in Section 3, we describe some experiments and results; finally in Section 4, some conclusions are presented.

2 Bayesian Framework for Multimodal Image Registration

2.1 Markov Random Fields

In this subsection, we present the basic definition of Markov Random Fields, for more detail refer to [2][12][8]. Let L be the discrete pixel lattice where 2D images of size $n \times m$ are observed:

$$L = \{(i, j) | 1 \leq i \leq n, 1 \leq j \leq m\}. \quad (1)$$

To simplify the notation, the pixels in a $n \times m$ image can be conveniently re-index by a number r taking values in $\{1, 2, \dots, n \times m\}$. The sites in L are related to one another via a neighborhood system. A neighborhood system for L is given by

$$N = \{N_r | \forall r \in L\}, \quad (2)$$

where N_r are the sites neighboring r . The neighborhood relationship has the following properties:

- a site is not neighboring to itself: $r \notin N_r$;
- the neighboring relationship is mutual: $r \in N_{r'} \iff r' \in N_r$.

We can define a graph (L, N) , where L contains the nodes and N determines the link between the nodes according to a neighborhood system. A clique C for (L, N) is a subset of sites in L such that for all $r, s \in C$ such that $r \neq s$, we have that $r \in N_s$ and $s \in N_r$. In a first order neighborhood system (the four nearest sites to r), cliques may be composed of either single sites $c = \{r\}$, or a pair of neighboring sites $c = \{r, r'\}$, thus the collections of single-cliques C_1 and pair-cliques C_2 are defined as

$$C_1 = \{\{r\} | r \in L\},$$

$$C_2 = \{\{r, s\} | r \in N_s, s \in N_r\}.$$

Let $F = \{F_1, \dots, F_{n \times m}\}$ be a family of random variables defined on L , where a realization of F_r can take a value f_r in Ω ; we denote a realization of a joint event as $F = f$. F is said to be a Markov random field on L with respect to a neighborhood system N if the following conditions are satisfied:

- $p(f) > 0, \forall f \in \mathbf{F}$,
- $p(f_r | f_{L-\{r\}}) = p(f_r | f_{N_r})$.

The Hammersley and Clifford theorem [2] establishes that an MRF has an equivalence with a Gibbs distribution, which has the following form

$$p(f) = \frac{1}{Z_f} \exp\left\{-\sum_C V_C(f)\right\}, \quad (3)$$

where Z_f is a normalizing constant, the sum in the exponential ranges over the cliques of the given neighborhood system on L , and $\{V_C\}$ are the potential functions, each one depending on the values of f at the sites that belong to the clique C . These potential functions, together with the neighborhood system, control the appearance of the sample field f .

2.2 Bayesian Estimation

To describe the probabilistic framework for multimodal image registration, we assume first that the observation model in each pixel is given by

$$I_2(T(r)) = g(I_1(r)) + \eta(r), \quad (4)$$

where I_1, I_2 are the images to register; T is the affine transformation that aligns the images I_1, I_2 ; and $\eta(r) \sim N(0, \sigma^2)$. $g(I_1(r))$ is the intensity transformation function which may be, in general, very complex such as logarithmic, gamma, contrast-stretching, inverse, polynomial, or thresholding transformations, (see

more details in [9]). In particular, we model this transformation with a locally linear function given by

$$g(I_1(r)) = K_1(r)I_1(r) + K_2(r), \tag{5}$$

where K_1 and K_2 are Markov random coefficient fields (MRCF) that describe the intensity transformation at each pixel r . Given the observation model (4) and the linear functions (5), one can estimate their parameters using Bayesian estimation theory, following the steps [13]:

1. Find the likelihood of the observation $p(I_1, I_2|K_1, K_2, T)$.
2. Using the prior distributions $p(K_1, K_2, T)$, find the posterior distribution $p(K_1, K_2, T|I_1, I_2)$.
3. Define an appropriate cost function $C(\widehat{K}_1, \widehat{K}_2, \widehat{T}, K_1, K_2, T)$, that assigns a cost to estimators $\widehat{K}_1, \widehat{K}_2, \widehat{T}$, given that the true values are K_1, K_2, T .
4. Find the optimal estimators K_1^*, K_2^*, T^* by minimizing

$$Q(\widehat{K}_1, \widehat{K}_2, \widehat{T}) = E[C(\widehat{K}_1, \widehat{K}_2, \widehat{T}, K_1, K_2, T)|I_1, I_2]. \tag{6}$$

Now, we proceed to analyze each step in detail.

Assuming that $\eta(r)$ (normal) is known and *iid*, the likelihood function can be written as

$$p(I_1, I_2|K_1, K_2, T) = \frac{1}{Z_L} \exp[-\sum_{r \in L} V_T(r)], \tag{7}$$

where

$$V_T(r) = \frac{(I_2(T(r)) - K_1(r)I_1(r) - K_2(r))^2}{2\sigma^2}. \tag{8}$$

In this model, K_1, K_2 , and T are assumed independent; hence, one can express $p(K_1, K_2, T)$ as a product of independent probabilities. Now, the probability of T is considered constant, and K_1, K_2 are MRF, resulting in the prior distribution:

$$\begin{aligned} p(K_1, K_2, T) &= p(K_1)p(K_2)p(T) \\ &= \frac{1}{Z_P} \exp[-\sum_C V_C(K_1) - \sum_C V_C(K_2) + \log p(T)]. \end{aligned} \tag{9}$$

Using (8), (9), and the Bayes rule, one finds the posterior distribution as:

$$p(K_1, K_2, T|I_1, I_2) = \frac{1}{Z} \exp[-U(K_1, K_2, T)], \tag{10}$$

where Z is a normalizing constant composed by $1/Z_L$ and $1/Z_P$, and

$$U(K_1, K_2, T) = \sum_{r \in L} V_T(r) + \sum_C V_C(K_1) + \sum_C V_C(K_2) - \kappa, \tag{11}$$

where V_T is given by (8), and κ is a noninformative constant; the potential function V_C considers only cliques of size 2, that is, nearest-pair sites $\langle r, s \rangle$ which are one unit apart:

$$V_C(K(r), K(s)) = \lambda_{r,s}(K(r) - K(s))^2, \tag{12}$$

where $\lambda_{r,s}$ is a positive regularization parameter that may depend on the sites $\langle r, s \rangle$; however, we used the same λ for all $\langle r, s \rangle$ in our implementation.

Let $\theta = [K_1, K_2, T]$ denote the estimator vector, and define the cost function $(1 - \delta(x))$, where

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

To find the optimal estimator θ^* , using this cost function, we see that

$$\begin{aligned} Q(\hat{\theta}) &= \int_{\theta \in \Theta} (1 - \delta(\hat{\theta} - \theta)) p(\theta | I_1, I_2) d\theta \\ &= \int_{\theta \in \Theta} p(\theta | I_1, I_2) d\theta - \int_{\theta \in \Theta} \delta(\hat{\theta} - \theta) p(\theta | I_1, I_2) d\theta \\ &= 1 - \int_{\theta \in \Theta} \delta(\hat{\theta} - \theta) p(\theta | I_1, I_2) d\theta. \end{aligned} \tag{14}$$

Therefore, to minimize (14), we need to find $\hat{\theta}$ that maximizes $p(\hat{\theta} | I_1, I_2)$, which is equivalent to finding

$$K_1^*, K_2^*, T^* = \arg \min_{K_1, K_2, T} U(K_1, K_2, T), \tag{15}$$

which is called maximum a posteriori (MAP) estimator.

2.3 Minimization Algorithm

The minimization of (15) may be achieved using different unconstrained optimization algorithms (see [15]). However, in this paper, we have used an efficient Newtonian gradient descent algorithm (NGD) [14]. This method is based on the idea of moving, in each iteration, in a direction d such that $\nabla U \cdot d < 0$ (i.e., a descent direction). The convergence may be accelerated if one considers each element of $K_1(r), K_2(r)$, and each element of the affine transformation T as the position of a particle of unit mass, subject to a force equal to $-\partial U / \partial K_1(r)$ (respectively, $-\partial U / \partial K_2(r), -\partial U / \partial T$). The equation of motion of these particles may be obtained from Newton’s second law. The discretization of these equations gives way to an iterative gradient descent algorithm with inertia:

$$\begin{aligned} K_1^{(t+h)}(r) &= \frac{2}{\alpha h + 1} K_1^{(t)}(r) + \frac{\alpha h - 1}{\alpha h + 1} K_1^{(t-h)}(r) \\ &\quad - \frac{h^2}{\alpha h + 1} \nabla_{K_1(r)} U^{(t)} \end{aligned} \tag{16}$$

$$\begin{aligned} K_2^{(t+h)}(r) &= \frac{2}{\alpha h + 1} K_2^{(t)}(r) + \frac{\alpha h - 1}{\alpha h + 1} K_2^{(t-h)}(r) \\ &\quad - \frac{h^2}{\alpha h + 1} \nabla_{K_2(r)} U^{(t)} \end{aligned} \tag{17}$$

$$\begin{aligned} T^{(t+h)} &= \frac{2}{\alpha h + 1} T^{(t)} + \frac{\alpha h - 1}{\alpha h + 1} T^{(t-h)} \\ &\quad - \frac{h^2}{\alpha h + 1} \nabla_T U^{(t)}, \end{aligned} \tag{18}$$

where α is a friction coefficient, $U^{(t)} = U(K_1^{(t)}(r), K_2^{(t)}(r), T^{(t)})$, and h is the step size. This method differs from the typical gradient descent in that the friction coefficient α allows the algorithm to avoid, in many cases, becoming trapped in local minima. Notice that if $\alpha = 1/h$ the NGD is a typical gradient descent method.

3 Results and Discussion

In the following section, we present some experiments involving different kind of images to test the performance of the algorithm. First, we show the ability of the MRCF to compute the intensity changes required to achieve the image registration between multimodal images and its robustness to noise and inhomogeneities. Second, we compare our algorithm against the method proposed by Viola et al. [18]. All these experiments were performed on a PC-based workstation running at 3.0 GHz.

3.1 Experiments

In order to test the ability of the proposed algorithm to find the coefficients of the linear intensity transformation functions, we built a one-dimensional signal of 126 samples. Fig. 1a shows the signal I_1 (thicker line), which is the negative of I_2 (thinner line), and shifted five samples to the left of I_2 . In the plot in Fig. 1b, we can see the thicker line composed by the intensity transformation $K_1(r)I_1(r) + K_2(r)$, and $I_2(r - d)$, where d is the displacement found by the proposed algorithm; to appreciate the matching between the signals, the thicker line is plotted few values below I_2 . In fact, the signal I_1 was built by setting $I_1(r) = -I_2(r + d)$; one can observe in Figures 1c and 1d how the MRCF K_1, K_2 approach this transformation.

In order to test the robustness of the algorithm to noise, the following experiment consisted in the registration of the images in Fig. 2a and Fig. 2b; this last one was built artificially. We added normal random values to the image in Fig. 2b. The true relative mean error (TRME) between the true parameter vector $\theta^* = [0.2094, 2.0, -36, -18]$ (corresponding to the angle, scale, and displacements in (x, y)) and the vector values θ_i obtained by the algorithm for different noise standard deviations $\sigma = \{0, 2, 4, 6, 8, 10, 12, 14, 16\}$ is plotted in Fig. 3. This error has the advantage of independently taking into account the unit scales of the quantities to evaluate, and it is computed for each θ_i as follows

$$TRME_i = \frac{\sum_{k=1}^4 \left| \frac{\theta_i^*(k) - \theta_i(k)}{\theta_i^*(k)} \right|}{4}. \quad (19)$$

In all the experiments, we used the same set of values for the parameters of the algorithm; in all tests, the error was less than 3.0%.

We also applied this registration approach to different kinds of brain images coming from different sources or processes. The first experiment consisted in registering a Magnetic Resonance (MR) image in Fig. 4a and the Computed Tomography

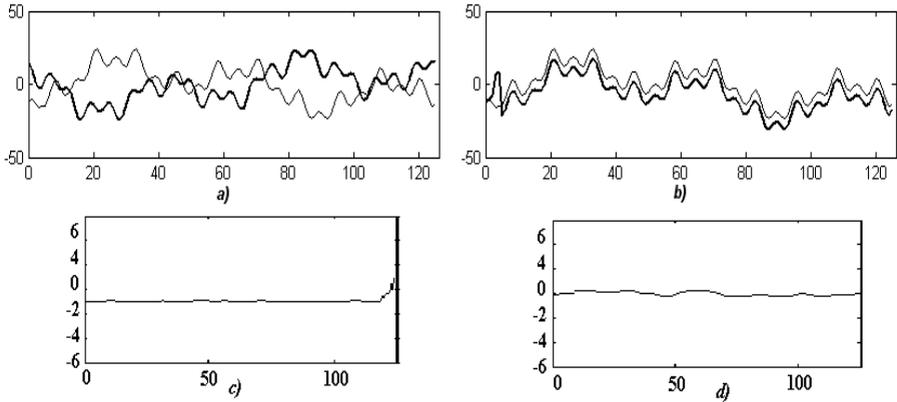


Fig. 1. a) Original signals I_1, I_2 ; b) aligned signals; c) K_1 field; d) K_2 field

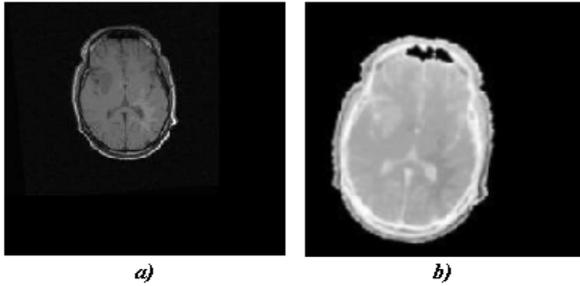


Fig. 2. a), b) images I_1, I_2 to align; c) transformed image I_1 ; d) difference between image transformed I_1 and registered I_2

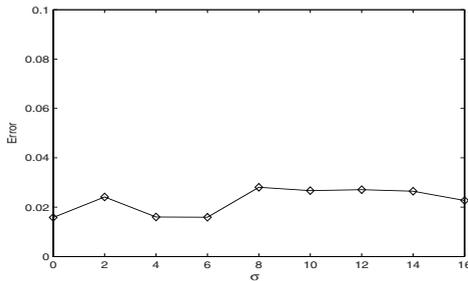


Fig. 3. True relative mean error

(CT) image in Fig. 4b. We can see in Fig. 4c the transformation of the MR-image using the MRCF to match the image in Fig. 4b that together with the estimated affine-parameters produce the superimposed registration shown in Fig. 4d.

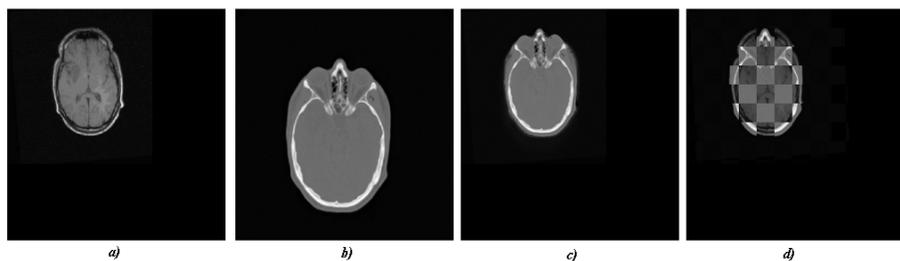


Fig. 4. *a)* MR-image, *b)* CT-image, *c)* transformed MR-image, *d)* superimposed registration

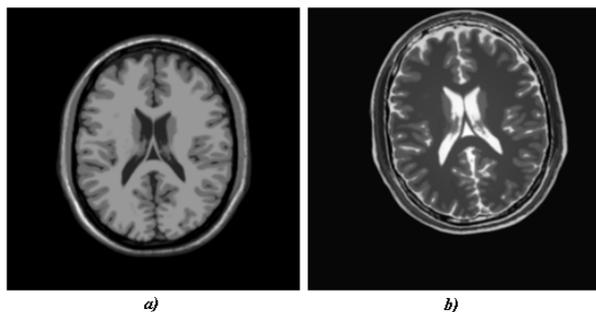


Fig. 5. *a)* T1-image, *b)* T2-image

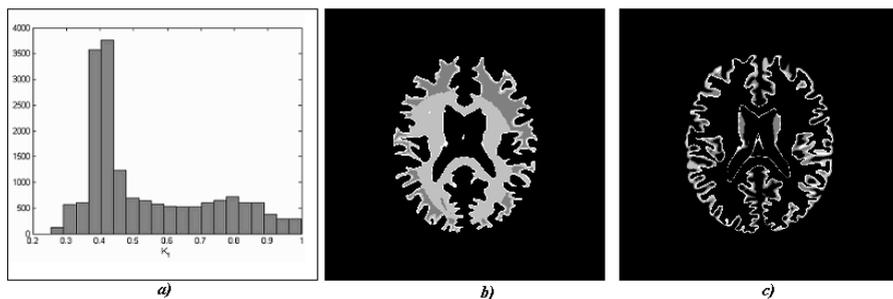


Fig. 6. *a)* Histogram of $0 < K_1 < 1$; quantized values of white and gray matter: *b)* $0.2 \leq K_1 < 0.6$, *c)* $0.6 \leq K_1 < 1$

In the next experiment, we examine the coefficient values of the fields K_1 and K_2 obtained by registering a synthetic magnetic resonance image spin-lattice relaxation time (T1), and spin-spin relaxation time (T2) obtained from the Brainweb Database [3]; these are shown in Fig. 5. The T1-image was produced with 0% of Gaussian noise and 0% intensity shading (inhomogeneity), while the T2-image with 0% of noise and 40% of inhomogeneity. A histogram of the values of

$0 < K_1(r) < 1$ is plotted in Fig. 6a. These values correspond to a region of the T1-image where it is necessary to reduce the intensity levels in order to match the intensity values of the same regions in the T2-image. We can see that there are two modes localized approximately at 0.4 and 0.8, corresponding respectively to the white and gray matter. These distributions show that it is necessary to have a set of coefficient values (i.e., different intensity transformation functions) to adjust the intensities of T1 to approximate those of T2 in these regions, mainly due to their inhomogeneity. This is more evident in Figures 6b and 6c where we separated the white and gray matter using the K_1 -interval values $[0.2, 0.6)$ and $[0.6, 1)$, and thresholded in intervals of 0.1.

3.2 Comparisons

Here we present some comparisons with one of the most popular and referenced algorithms in the literature; this method was presented in [18] and it is based on Mutual Information theory. To do this, we obtained T1 and T2 images from the Brainweb and made several experiments. The first one consists in registering a T1-image with 3% of noise and 20% of inhomogeneity versus a set of T2-images (similar to that shown in Fig. 5) having different level of noise and 40% of inhomogeneity; the set of images were previously transformed using a known affine transformation. In both algorithms, the transformation T was initialized with the identity. Due to the stochastic nature of the MI method, it required to let the program run for 300 seconds ten times for each image pair. However, since MRCF is deterministic, we let the program run 300 seconds once for each image pair. The results are plotted on Figure 7. Notice that MI does not always converge to an acceptable solution in most cases (large variances), while MRCF reached a TRME below 1% in all cases.

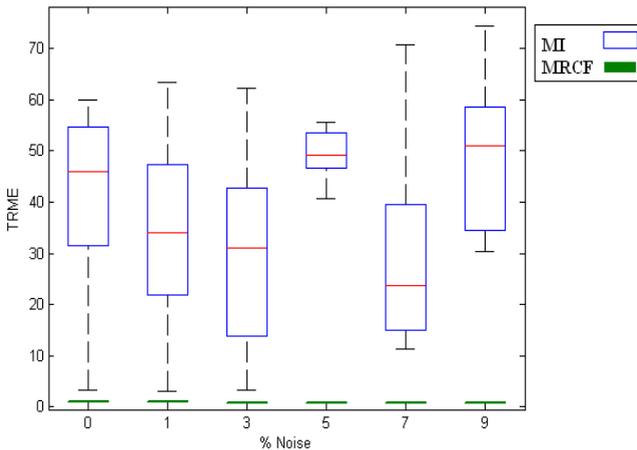


Fig. 7. Boxplot of results obtained by MI and MRCF

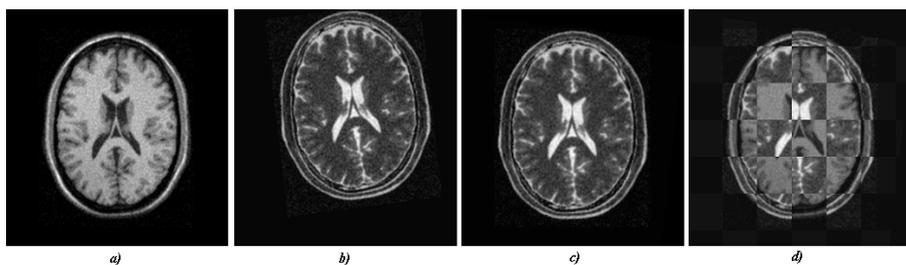


Fig. 8. *a)* T1-image, *b)* T2-image, *c)* transformed T1-image, *d)* superimposed registration

For the final comparison, we choose the hardest case in which a T1-image with 9% of noise and 40% of inhomogeneity was registered with a T2-images with also 9% of noise and 40% of inhomogeneity. The results obtained by MRCF are shown in Fig. 8. The TRME of MRCF was of 1.1865%, while for MI was 71.0916%, both computed in 600 seconds.

4 Conclusions

This work presents an algorithm rigorously based on Bayesian estimation in which two Markov Random Fields (K_1, K_2) represent the coefficients of linear intensity transfer functions applied to each pixel. These functions are included in a very simple observation model (4) that allows one to estimate with high precision the necessary intensity changes and the parameter values of the affine transformation to match the images to register. Another important characteristic of this energy function is that it includes spatial coherence as priori knowledge by means of the MRCF (see equations (11), (12)). Although the resulting posteriori energy function (15) is highly non-linear with respect to the affine transformation parameters, and quadratic with respect to the MRF's K_1, K_2 , it was successfully minimized using an efficient, simple, and easy to implement Newtonian gradient descent algorithm.

The paper also presents examples that illustrate the generality of the algorithm to estimate the coefficient values K_1 and K_2 of the local linear functions to approach the intensity transformation needed to achieve the image registration. We showed the performance and stability of the algorithm to get high precision registrations in cases in which radical intensity changes exist, as those shown in Figures 1, 4, 5 and 8. Preliminary results show that the fields K_1 and K_2 may also yield discriminatory information about the different regions in the images, which may be useful for a posterior segmentation process. Finally, we demonstrate the robustness of the proposed algorithm to noise and intensity inhomogeneities, outperforming the MI-algorithm as it was described in [18].

Perspectives for future research include: (1) a generalization of the proposed methodology for the registration of 3D brain images, (2) the addition of a segmentation stage that takes advantage of the MRCF K_1 and K_2 , and (3)

the application of MRCF to other problems in computer vision and image processing.

Acknowledgments. For this research work, Edgar Arce-Santana was supported by Grant PROMEP/103.5/04/1387, C06-FAI-11-31.68, and Alfonso Alba by PROMEP/103.5/07/2416. The authors would like to thank Professor Jose Luis Marroquin Zaleta, CIMAT, Mexico, for sharing his insight.

References

1. Banerjee, P.K., Toga, A.W.: Image alignment by integrated rotational and translation transformation matrix. *Physics in Medical and Biology* 39, 1969–1988 (1994)
2. Besang, J.: Spatial interaction and statistical analysis of lattice systems. *J. Royal Statistical Soc. B* 36(1)(2), 192–236 (1974)
3. Cocosco, C.A., Kollokian, V., Kwan, R.K., Evans, A.C.: Brain web: Online interface to a 3DMRI simulated brain database. *NeuroImage* 5(2), Part 2/4, S425 (1997) (Proceedings of the 3rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997)
4. Ding, E., Kularatna, T., Satter, M.: Volumetric image registration by template matching. In: *Medical Imaging 2000*, pp. 1235–1246. SPIE, Bellingham, WA (2000)
5. Du, J., Tang, S., Jiang, T., Lu, Z.: Intensity based robust similarity for multimodal image registration. *International Journal of Computer Mathematics* 83, 49–57 (2006)
6. Fitzpatrick, J., West, J., Maurer, C.: Predicting error in rigid-body, point-based registration. *IEEE Transactions on Medical Imaging* 17, 694–702 (1998)
7. Frantz, S., Rohr, K., Stiehl, H.S., Kim, S.I., Weese, J.: Validation point-based MR/CT registration based on semi-automatic landmark extraction. In: *Proceeding of CARS 1999*, pp. 233–237. Elsevier, Amsterdam (1999)
8. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6(6), 721–741 (1984)
9. Gonzales, R.C., Woods, R.E., Eddins, S.L.: *Digital Image Processing Using Matlab*. Prentice-Hall, NJ (2004)
10. Guimond, A., Roche, A., Ayache, N., Meunier, J.: Three-Dimensional Multimodal Brain Wrapping Using the Demons Algorithm and Adaptive Intensity Corrections. *IEEE Transaction on Medical Imaging* 20(1), 58–69 (2001)
11. Hsu, L., Loew, M.H., Ostuni, L.J.: Automated registration of CT and MR brain images using hierarchical shape representation. *IEEE Engineering in Medicine and Biology Magazine* 18, 40–47 (1999)
12. Li, S.Z.: *Markov Random Field Modeling in Computer Vision*. Springer, Berlin (1995)
13. Marroquin, J.L., Mitter, S., Poggio, T.: Probabilistic Solution of Ill-Posed problems in Computational Vision. *J. Am. Statistical Assoc.* 82(397), 76–89 (1987)
14. Marroquin, J.L.: Deterministic Interactive Particle Models for Image Processing and Computer Graphics. *Graphical Models and Image Processing* 55(5), 408–417 (1996)
15. Nocedal, J., Wright, S.J.: *Numerical Optimization*. Springer, Heidelberg (1999)

16. Shekhar, R., Zagrodsky, V.: Mutual Information-based rigid and non-rigid registration of ultrasound volumes. *IEEE Transaction on Medical Imaging* 21, 9–22 (2002)
17. Thirion, J.-P.: Image matching as a diffusion process: An analogy with Maxwell's demons. *Med. Image Anal.* 2, 243–260 (1998)
18. Viola, P.A., Wells III, W.M., Atsumi, H., Nakajima, S., Kikinis, R.: Multi-modal Volumen Registration by Maximization of Mutual Infromation. *Medical Image Analysis* 1, 5–51 (1996)
19. Woods, R.P., Mazziotta, J.C., Cherry, S.R.: MRI-PET registration with automated algorithm. *Journal of Computer Asisted Tomography* 17, 536–546 (1993)