Image Registration Using Markov Random Coefficient and Geometric Transformation Fields

E. R. Arce-Santana, A. Alba
Facultad de Ciencias
Universidad Autonoma de San Luis Potosi
San Luis Potosi, Mexico
Email: {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 multimodal image registration. A Bayesian formulation is presented in which a likelihood term is defined using an observation model based on coefficient and geometric fields. These coefficients, that represent the local intensity polynomial transformations, as the local geometric transformations, are modelled 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 fields, making the registration between the images possible.

I. INTRODUCTION

Image registration is the alignment of images, this task is very important to many applications 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.

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 [1][2][3], and those based on intensity [4][5]. Within the intensity methods there are two popular ones: Partitioned Intensity Uniformly (PIU) [6] and those based on mutual information (MI), proposed by Viola et al [7]. A drawback of these methods is that they ignores completely spatial information such as edges or homogenous regions.

A method related to the work proposed in this paper is presented in [8]. It focuses only on the elastic registration of multimodal images. This method makes the assumption that there are at most two functional dependencies between intensities. In [9], the authors propose a method that uses finite element for modelling both the geometric transformations and the intensity corrections. Other related work is found in [10], where an affine transformation between images is modelled locally using a linear approximation (first order Taylor’s expansion). To enforce global consistency it is necessary to impose smoothness constrains in both the geometric and the intensity transformation parameters.

In this work, we present a more general registration method, Rigourously based on Bayesian estimation, the main goal of this method is to determine at each pixel, the parameters of the local affine transformation, and at the same time, the coefficient values of the polynomial intensity transfer functions to that achieve the image registration. In this approach, the coefficients of the intensity transformations (labeled MRCF) and the parameters of the geometric transformations (labeled MRGTF) are represented as Markov Random Fields (MRF)[11], giving in this way the prior information about the homogeneity of the intensity and geometric changes. This method, called MR-CGTF, gives the possibility to estimate complex geometric transformations, at global or local scales, by means of a flexible spline-based model that take into account spatial variations.

The paper is organized as follows: in Section II, we present the Bayesian framework of image registration using MRCGTF; Section III shows some experiments and results; and finally, in Section IV, some conclusions are presented.

II. BAYESIAN ESTIMATION

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

\[ I_2(T(r)) = g(I_1(r)) + \eta(r), \]  

(1)

where \( I_1, I_2 \) are the images to register; \( T \) is the local global 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. In particular, we model this transformation with a local polynomial function given by

\[ g(I_1(r)) = K_n(r)I_1(r)^n + K_{n-1}(r)I_1(r)^{n-1} + ... + K_0(r), \]

(2)

where the \( K \)'s are Markov random coefficient fields (MRCF) that describe the intensity transformation at each pixel \( r \); the degree of the polynomial depends on the complexity of the required transformation. Given the observation model (1) and the polynomial functions (2), one can estimate their parameters using Bayesian estimation theory.

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

\[ p(I_1, I_2|K_n, ..., K_0, T) = \frac{1}{Z_L} \exp\left[-\sum_{r \in L} V_T(r)\right], \]

(3)
where
\[ V_T(r) = \frac{(I_2(T(r)) - K_n(r)I_1(r))^2}{2\sigma^2}. \] (4)

In this model, the \( K \)'s, and \( T \) are assumed independent; hence, one can express \( p(K_n, \ldots, K_0, T) \) as a product of independent probabilities.

In the general case, when the geometric transformation may be very complex, such as an elastic transformation, one can define a locally-affine transformation for each pixel as:
\[ T(r) = f(r, \phi_1(r), \ldots, \phi_\tau(r)), \] (5)
where \( f(\cdot) \) is the geometric affine transformation, and \( \phi_i(r), i = 1, \ldots, 7 \) are the affine transformation parameters given by a spline-based models, which are defined as linear combinations of basis functions \( N_j, j = 1, \ldots, J \) each one of which has a local finite support:
\[ \phi_i(r) = \sum_{j=1}^{J} \omega_{ij} N_j(r), \] (6)
where the \( \omega \)'s are MRF (MRGTF) used to estimate this transformation at each pixel that determine the weights of each basis function for each parameter of the geometric transformation.

Now, to impose global consistency, we consider the \( K \)'s and the \( \omega \)'s as MRF, resulting in the prior distribution:
\[ p(K_n, \ldots, K_0, T) = p(K_n) \cdots p(K_0) p(\omega_1) \cdots p(\omega_\tau) \]
\[ \frac{1}{Z_P} \exp[-\sum_{i=0}^{n} \sum_C V_C(K_i) - \sum_{i=1}^{7} \sum_C V_C(\omega_i)] \] (7)
where \( \frac{1}{Z_P} \) is a normalizing constant and \( V_C \) are potential functions, each one of which depends only on the values at the sites that belong to the clique \( C \) (see [12] for details).

Using (4), (7), and the Bayes rule, one can estimate the posterior distribution as:
\[ p(K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau | I_1, I_2) = \frac{Z}{Z_P} \exp[-U(K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau)], \] (8)
where \( Z \) is a normalizing constant composed by \( 1/Z_L \) and \( 1/Z_P \), and
\[ U(K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau) = \sum_{r \in L} V_f(r, \phi(1)(r), \ldots, \phi(\tau)(r)) + \sum_{i=0}^{n} \sum_C V_C(K_i) + \sum_{i=1}^{7} \sum_C V_C(\omega_i), \] (9)
where \( V_f \) is the likelihood function depending on the affine transformation and the intensity polynomial function:
\[ \frac{V_f(r, \phi_1(r), \ldots, \phi_\tau(r))}{Z_P} = \frac{(I_2(f(r, \phi_1(r), \ldots, \phi_\tau(r))) - K_n(r)I_1(r))^2}{2\sigma^2}. \] (10)
The potential functions \( V_C \) consider only cliques of size 2, that is, the nearest-pair sites \( < r, s > \) in the case of the intensity-coefficients \( K \)'s, and \( < j, k > \) in the spline subgrids \( \omega \)'s, which are one unit apart:
\[ V_C(K_i(r), K_i(s)) = \lambda_{r,s}^{(i)} (K_i(r) - K_i(s))^2, \] (11)
\[ V_C(\omega_{ij}, \omega_{ik}) = \lambda_{j,k}^{(i)} (\omega_{ij} - \omega_{ik})^2. \] (12)
where \( \lambda_{r,s}^{(i)} \) and \( \lambda_{j,k}^{(i)} \) are positive parameters that do not necessarily depend on the sites \( < r, s > \) or \( < j, k > \).

Let \( \theta = [K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau] \) denote the estimator vector, and define the cost function \( (1 - \delta(x)) \), where \( \delta(x) \) is the Kronecker’s delta function. To find the optimal estimator \( \theta^* \), using this cost function, we see that
\[ Q(\hat{\theta}) = 1 - \int_{\theta \in \Theta} \delta(\hat{\theta} - \theta)p(\theta | I_1, I_2) d\theta. \] (13)
Therefore, to minimize (13), we need to find \( \hat{\theta} \) that maximizes \( p(\theta | I_1, I_2) \), which is equivalent to finding
\[ K_n^*, \ldots, K_0^*, \omega_1^*, \ldots, \omega_\tau^* = \arg\min_{K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau} U(K_n, \ldots, K_0, \omega_1, \ldots, \omega_\tau), \] (14)
which is called maximum a posteriori (MAP) estimator.

The minimization of (14) may be achieved using different unconstrained optimization algorithms. However, in this paper, we have used an efficient Newton gradient descent algorithm (NGD) [13]. A very important aspect in the minimization algorithm is to ensure that it does not converge towards regions where the differences between the images is explained, for the most part, by the intensity transformation, and thus the geometric transformation is never found. This can be achieved coordinating the geometric and the intensity transformations, as soon as choosing the adequate parameters \( \lambda_{r,s} \) and \( \lambda_{j,k} \). One way to obtain this stability is by using a multiscale strategy in the number of control nodes at the spline subgrid (the number \( J \) for the \( \omega_{ij} \)). Initially, few nodes are set allowing the splines to model global transformations (see an example in Fig. 3). At finer scales, more nodes are added in order to model any local details in the transformation fields that may be required for a faithful registration. The values at the finer subgrids can be obtained by bilinear interpolation between nodes at the previous coarse subgrid.

### III. RESULTS AND DISCUSSION

#### A. Rigid Image Registration

In order to make quantitative comparisons the following experiment consisted on a synthetic experiment where the geometric transformation was spatially-constant and rigid. For this reason, instead of using the spline subgrid framework, we applied a constant rigid transformation to all pixels, \( T(r) = f(r, \phi_1, \ldots, \phi_4) \), where \( \phi \)'s correspond to the angle, scale, x-translation, and y-translation. In these experiments we used a linear polynomial as intensity transformation and due to the geometric transformation \( T(r) \) applied to the entire image \( I_2 \) is considered constant, we can rewrite the prior distribution (7) as:
\[ p(K_1, K_0, T) = \frac{\exp[-\sum_C V_C(K_1) - \sum_C V_C(K_0) + \log p(T)]}{Z_P}, \] (15)
and the energy function is thus
\[ U(K_1, K_0, T) = \sum_{r \in L} V_f(r) + \sum_C V_C(K_1) + \sum_C V_C(K_0) - \kappa, \] (16)
where $\kappa$ is a noninformative constant. For the experiment, we chose 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. 1. The TRME of MRCF was of 1.1865%, while for MI was 71.0916%, both computed in 600 seconds.

B. Elastic Image Registration

Here, we present some experiments in which the geometric transformations are very complex, and others where there are also considerable intensity changes in the images to register.

The first experiment consisted in the registration of two images where the geometric transformation is elastic. These 128 $\times$ 128 sagittal MR-images, which are shown in Fig. 2, come from two different patients. Figure 3a shows the results obtained by transforming the image in Fig. 2b using equation (14) and a spline subgrid with a distance between neighboring control nodes of 16 pixels in both directions; Fig. 3b shows the geometric transformation applied to a checker-board in order to appreciate details of the local affine estimated transformation. We repeated the experiment but now using the spline subgrid multiscale framework. The method starts initially with a grid with nodes placed each 64 pixels in both directions and nodes were subsequently added until we obtained 8-spaced node grid. The results are shown in Fig. 4, we can observe in the first two rows how the coarse grids estimate a global geometric transformation, whereas the fine grids adjust small details and improve the overall quality of the registration (third and forth rows). The second column in Fig. 4 shows the superimposed images $I_1$ and the transformed image $I_2$, increasing the contrast to appreciate well the registration; finally, last column shows the geometric transformation in each scale applied to a checker-board image. The elapsed time to get the final result (last row in Fig. 4) was 84.32 seconds.

In the next experiment, we applied artificially a histogram equalization, which is a complex non-linear transformation, to the image in Fig. 2b, shown in Fig. 5b; notice the intensity difference between the two images and the intensity contrast in $I_2$. Figures 5c and 5d show, respectively, the aligned image $I_2$, and the superposition of the registered images. The required time to register these 256 $\times$ 256-images was 183 seconds.

IV. CONCLUSIONS

This work presents an algorithm rigourously based on Bayesian estimation in which two different kinds of Markov Random Fields are defined, representing the coefficients of polynomial intensity transfer functions (MRCF’s), and the parameters of a local affine transformations (MRGTF’s). These MRF’s are included in a very simple observation model (1) that allows one to estimate with high precision the necessary intensity changes and the geometric transformations to align two images from different sensors. These MRF models also allow the inclusion of spatial coherence as prior knowledge. Although the resulting posteriori energy function (14) is highly non-linear with respect to the geometric transformation, and quadratic with respect to the MRCF’s, it was successfully minimized using an efficient, simple, and easy to implement Newtonian gradient descent algorithm. In order
to find complex geometric transformations, we also propose a Spline Subgrid Multiscale Framework. This strategy has the purpose to stabilize and estimate the transformations with a gradually-increasing detail, until a successful alignment is achieved. We have shown the performance and stability of the algorithm to get high precision registrations in cases in which radial intensity changes exist. We demonstrate the registration accuracy, and the robustness to noise and intensity inhomogeneities of the proposed method by comparing it with the MI-algorithm as it was described in [7]. Finally, we prove the experimental performance of the algorithm during elastic image registrations, which may be achieved in few seconds, even in cases where complex intensity and geometric transformations are involved.

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’s, and (3) the application of MR-CGTF 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, 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