Medical Image Analysis 12 (2008) 731–741

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Dense image registration through MRFs and efficient linear programming ☆

Ben Glocker [a,b,*], Nikos Komodakis [a,c], Georgios Tziritas [c], Nassir Navab [b], Nikos Paragios [a,d]

[a] Laboratoire des, Mathématiques Appliquées aux Systèmes (MAS), Ecole Centrale de Paris, France
[b] Computer Aided Medical Procedures (CAMP), Technische Universität München, Germany
[c] Computer Science Department, University of Crete, Greece
[d] Equipe Galen, INRIA Saclay-Île-de-France, France

ARTICLE INFO

Article history:
Received 1 October 2007
Received in revised form 2 March 2008
Accepted 14 March 2008
Available online 7 April 2008

Keywords:
Discrete optimization
Deformable registration
Linear programming

ABSTRACT

In this paper, we introduce a novel and efficient approach to dense image registration, which does not require a derivative of the employed cost function. In such a context, the registration problem is formulated using a discrete Markov random field objective function. First, towards dimensionality reduction on the variables we assume that the dense deformation field can be expressed using a small number of control points (registration grid) and an interpolation strategy. Then, the registration cost is expressed using a discrete sum over image costs (using an arbitrary similarity measure) projected on the control points, and a smoothness term that penalizes local deviations on the deformation field according to a neighborhood system on the grid. Towards a discrete approach, the search space is quantized resulting in a fully discrete model. In order to account for large deformations and produce results on a high resolution level, a multi-scale incremental approach is considered where the optimal solution is iteratively updated. This is done through successive morphings of the source towards the target image. Efficient linear programming using the primal dual principles is considered to recover the lowest potential of the cost function. Very promising results using synthetic data with known deformations and real data demonstrate the potentials of our approach.

1. Introduction

Medical image analysis (Duncan and Ayache, 2000; Paragios et al., 2009) is an established domain in computational, mathematical and biological sciences. Recent advances on the acquisition side have made possible the visualization of human tissues as well as physiological and pathological indices related to them either occasionally or periodically. The ability to compare or fuse information across subjects with origins of different modalities is a critical and necessary component of computer aided diagnosis. The term used often to express this need is registration.

The registration problem often involves three aspects: (i) the transformation model, (ii) a similarity criterion and (iii) an optimization strategy.

Registration can be either global or local. Parametric models are often employed to address global registration with a small number of degrees of freedom, such as rigid or similarity. These models refer to a good compromise between performance and computational complexity. The registration problem in such a context is well posed since the number of variables to be determined is over-constrained from the number of observations. Dense image registration aims to go further and seeks individual correspondences between observations. The main goal is to determine relationships that locally express the correlation of the observations either for the same subject (acquisitions of different modalities or acquisitions of the same organ in time). Local alignment or dense/deformable registration is the term often considered to describe this task.

Deformable registration is one of the most challenging problems in medical imaging. The problem consists of recovering a local transformation that aligns two signals that have in general an unknown relationship both in the spatial domain and in the intensity domain. Several methods exist in the literature where specific measures are designed to account for this relationship and to optimize the transformation that brings these two signals together.

Local image alignment is often performed according to geometric or photometric criteria. Landmark-based methods (Hellier and Barillot, 2003; Rohr et al., 2003) are a classic example of geometric-driven registration. In such a setting, a number of anatomical key points (Pennec et al., 2000)/structures (segmented values) are identified both in the source and in the target image, and a transformation that aims to minimize the Euclidean distance

☆ This work is partially supported by Siemens Medical Solutions, Germany and from the French ANR-Blanc grant SURF (2005–2008).

* Corresponding author. Address: Computer Aided Medical Procedures (CAMP), Technische Universität München, Germany. Tel.: +49 8928917071; fax: +49 8928917059.

E-mail addresses: glocker@in.tum.de (B. Glocker), komod@csd.uoc.gr (N. Komodakis), tziritas@csd.uoc.gr (G. Tziritas), navab@in.tum.de (N. Navab), nikos.paragios@ecp.fr (N. Paragios).

1361-8415/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.media.2008.03.006

between these structures is to be recovered. The main limitation of these methods is related to the selection and extraction of landmarks, while their main strength is the simplicity of the optimization process.

Iconic registration methods (Cachier et al., 2003) seek for "visual" correspondences between the source and the target image. Such a problem is tractable when one seeks registration for images from the same modality due to an explicit photometric correspondence of the image intensities. Sum of squared differences, sum of absolute differences, cross correlation (Hajnal et al., 2001) or distances on subspaces that involve both appearance and geometry (intensities, curvature and higher order image moments) (Davatzikos et al., 1996) have been considered. On the other hand, it becomes more challenging when seeking transformations between different modalities with a non-linear or only statistical relation of intensities (Hermosillo et al., 2002). The measures that have often been used were normalized mutual information (Maes et al., 1997), Kullback-Liebler divergence (Zollei et al., 2005) and correlation ratio (Roche et al., 1998) to define similarity[1] between different modalities.

Once the similarity measure has been defined the next task consists of recovering the parameters that optimize the designed cost function. Parameters can be either searched or estimated. In the first case techniques like exhaustive search can be employed which are time consuming. On the other hand, one can use known optimization techniques, gradient-free or gradient-based to determine the optimal set of parameters starting from an initial guess (Klein et al., 2007). These methods require an important customization from one application to another since a correlation exists between the modalities/problem and the selection of the similarity measure. Furthermore, the optimization is often sub-optimal due the non-convexity of the designed cost functions. In particular when considering complex similarity functions defined on the continuous space, then the numerical approximation of the gradient in the discrete domain (image/volume plane) is very challenging leading to erroneous registration results.

The aim of our approach is to overcome both limitations present in all registration methods: dependency on the similarity measure selectionand on the initial conditions in a reasonable computation time.

In this article, we propose a novel technique that can be used either for inter or intramodal image registration. Towards satisfying the smoothness of the deformation field and reducing the dimensionality of the problem, we represent deformation through free form deformations (Sederberg and Parry, 1986). Our method reformulates registration as a Markov random field (MRF) optimization where a set of labels is associated with a set of deformations. Then we seek to attribute a label to each control point such that once the corresponding deformation has been applied, the similarity measure between the source and the target is optimal for all voxels. The optimization procedure is independent from the graph construction, and therefore any similarity measure can be used.

The remainder of this paper is organized as follows: In Section 2, we introduce the proposed registration framework, while in Section 3 we discuss the optimization aspects. Implementation details are given in Section 4 and experimental validation are part of Section 5. Section 6 concludes our paper.

2. Deformable registration

In order to introduce the concept of our approach (Glocker et al., 2007), we consider (without the loss of generality) the 2D image domain. Let us consider a source $f : \Omega = [1, N] \times [1, M] \rightarrow \mathscr{R}$ and a target image g. In general, these images are related with a non-linear transformation as well as a non-linear relation between intensities, that is

$$\forall \mathbf{x} \in \Omega \quad g(\mathbf{x}) = h \circ f(\mathscr{T}(\mathbf{x})), \tag{1}$$

where $\mathscr{T}(\mathbf{x})$ is the transformation and h is a non-linear operator explaining the changes of appearance between them. The most common way to formulate the registration problem is through the definition of a distance between the source and the target image that is to be minimized in the entire domain Ω, or

$$E_{\text{data}}(\mathscr{T}) = \int_\Omega |g(\mathbf{x}) - h \circ f(\mathscr{T}(\mathbf{x}))| \mathrm{d}\mathbf{x}. \tag{2}$$

Recovering the optimal potential of this objective function is not straightforward. In the case of 2D images, two variables are to be determined while one constraint is available per pixel. The most basic approach to address this limitation is through the use of a regularization function on the space of unknown variables (Tikhonov, 1992), or

$$E_{\text{smooth}}(\mathscr{T}) = \int_\Omega \phi |\nabla \mathscr{T}(\mathbf{x})| \mathrm{d}\mathbf{x} \tag{3}$$

with ϕ being a convex function imposing smoothness on the deformation field for neighboring pixels. Such a term will make the estimation of the deformation field feasible assuming that the relationship between the signal intensities is known. This hypothesis is not realistic due to the fact that (i) when registering the same modalities this relationship depends on the parameters of the scanner which are not available and (ii) when registering different modalities in most of the cases, such an operator does not exist.

In order to overcome this constraint, in the most general case a similarity measure ρ is introduced to account for the intensity relation between the two images, or

$$E_{\text{data}}(\mathscr{T}) = \int_\Omega \rho_h(g(\mathbf{x}), f(\mathscr{T}(\mathbf{x}))) \mathrm{d}\mathbf{x}. \tag{4}$$

The definition of the ρ_h depends on the nature of the observed signals as well as the application itself. Once this measure is defined, the data term is combined with the smoothness one to determine the objective/cost function under consideration. Gradient-descent is the most common approach to perform the optimization, a method that has some strengths and known limitations. One can claim that this approach is convenient and it is often straightforward to implement. On the other hand, the problem is ill-posed due to the fact that the number of constraints is inferior to the number of variables to be determined. Furthermore, since the cost function is non-convex one cannot guarantee that the obtained solution will be the optimal one. Last, but not least, gradient numerical manipulation is not straightforward when projecting from the continuous space to the discrete one.

The above observations lead to a natural conclusion that one should seek (i) dimensionality reduction on the degrees of freedom of the model, (ii) more efficient optimization techniques both in terms of ability to approach the optimal solution with reasonable computational cost, and (iii) techniques that do not require continuous gradient manipulation in discrete spaces.

2.1. Continuous domain

Since we are interested in local registration, let us introduce a deformation grid $G : [1, K] \times [1, L]$ (usually $K \ll M$ and $L \ll N$) superimposed onto the image (no particular assumption is made on the grid resolution). The central idea of our approach is to deform the grid (with a 2D displacement vector $\mathbf{d_p}$ for each control point) such that the underlying image structures are perfectly

[1] For consistency reasons, we always use the term similarity measure, although measures such as the sum of squared differences are actually dissimilarity measures.