Quasi real-time sub-space 3D deformable fusion

Bharat Singh, Stavros Alchatzidis, Nikos Paragios

QUASI REAL-TIME SUB-SPACE 3D DEFORMABLE FUSION

Bharat Singh, Stavros Alchatzidis and Nikos Paragios

Ecole Centrale Paris

ABSTRACT

In this paper we propose a novel and robust approach for deformable fusion using a metric defined in an appropriate subspace which is adaptive to the image-content/image-modality. We adopt a graph-based formulation that assumes that intensities of corresponding pixels in the two image domains are related through an unknown piece-wise constant linear function. This relation is propagated to an appropriate sub-space (wavelets coefficients) where a criterion that couples the estimation of the deformation field with optimal transport function on the subspace and the smoothness of the deformation is considered. Message passing methods efficiently implemented using jump flooding are considered to get the optimal parameters of the deformation field and the transport function. Promising results are obtained in near real-time with a unique parameter setting.

Index Terms— deformable image registration, wavelets, similarity metric

1. INTRODUCTION

Deformable multi-modal fusion is a well explored topic in biomedical image analysis. It shares the challenges of mono-modal registration with the additional constraint of defining an appropriate similarity metric capturing the co-dependencies between the images to be registered.

Removing intensity variation through implicit image normalization was the first attempt to cope with multi-modal fusion using criteria like normalized cross correlation or correlation ratio that work well in general in the multi-modal setting. Exploring feature spaces beyond intensities that capture rich geometric information was a further attempt to cope with deformable fusion [1]. The use of statistical methods and in particular of mutual information [2] was a major breakthrough in the field because, the similarity criterion does not measure the quality of correspondences between individual pixels but a statistical entropy-driven criterion derived from the distributions of the image intensities before and after alignment. Numerous metrics have been inspired from mutual information, for example its normalized version, the Kullback-Leibler divergence or the Jensen-Renyi divergence. It should be noted that due to the non-linearity of the similarity criterion, the optimization component of the registration process often converges to a local minimum and the obtained solution depends heavily on the initial conditions/alignment between the two images.

The aim of this paper is to provide a novel, efficient, robust metric for general purpose multi-modal registration. To this end, we consider a robust/efficient sub-space local representation of the images on a wavelet base. Then assuming a local partition of the image domain, where there exists an unknown piece-wise constant linear mapping from one modality to another, we propagate this mapping to the subspace. An appropriate metric that depends on the unknown mapping between the two images is defined in the subspace. Such a propagation makes feasible the quantization of the mapping parameter through the subspace metric, which is propagated to a hierarchical/coarse-to-fine graphical model [3]. In this model both deformations (grid-like interpolations [4]) and local mapping are unknown and optimized together. The resulting MRF is sub-modular and message passing methods are considered to recover the joint solution of subspace-mapping/deformation producing quasi realtime results which are comparable to the state of the art.

The reminder of this paper is organized as follows, section 2 presents the metric and the registration formulation, while section 3 describes the the optimization framework. Implementation details and experimental validation are part of the section 4 while section 5 concludes the paper.

2. DEFORMABLE FUSION

Let us consider two images $s(\mathbf{x})$ and $t(\mathbf{x})$ corresponding to different modalities. Without loss of generality let us consider that the aim is to map s to t using a non linear transformation $d(\mathbf{x})$. Once a similarity metric is given, denoted with $\rho(s(\mathbf{x}), t(\mathbf{y}))$, the optimal solution of fusion refers to the transformation $d(\mathbf{x})$ that minimizes an image-driven term and a regularization constraint on the deformation space,

$$\min_{d} \int_{\Omega} \rho(s(d(\mathbf{x})), t(\mathbf{x}))d\mathbf{x} + \alpha \int_{\Omega} \phi(\nabla d(\mathbf{x}))d\mathbf{x}$$

where Ω is the image domain and ϕ imposes a smoothness constraint on the deformation field. Let us consider a partition of the image domain Ω into a set of sub-domains

This research was partially supported by European Research Council Starting Grant Diocles (ERC-STG-259112).

$\{\Omega_1, \cdots \Omega_n\}$. The union of these domains correspond to the image registration domain and we further assume that adjacent sub-domains could overlap. Then, given the optimal deformation, we consider a linear transport function relating the two signals, $\mathbf{x} \in \Omega_i : s(d(\mathbf{x})) = a_i t(\mathbf{x}) + b_i$, where a_i and b_i are sub-domain constants corresponding to the linear mapping between the two modalities. It should be noted that linearity is for the purpose of facilitating the introduction of the model and by no means a hard constraint. Given such a model, we can cast registration as a coupled inference process that aims at optimally determining the image partition sub-domains, the transport coefficients and the optimal deformation, or

$$\min_{n,a_i,b_i,\Omega_i,d} \sum_{i=1}^{n} \left\{ \int_{\Omega_i} \rho(s(d(\mathbf{x})), a_i t(\mathbf{x}) + b_i) d\mathbf{x} \right\} + \alpha \int_{\Omega} \phi(\nabla d(\mathbf{x})) d\mathbf{x}$$

2.1. Deformation Model/Image Partition

We consider a grid like representation endowed with an interpolation function. Let us consider the well established model of hierarchical free-form deformations which consists of superimposing a grid-like structure $G_{k,l}, (k,l) \in [1,K] \times [1,L]$ to the image. The deformation of a given image pixel is computed by a cubic B-Spline interpolation of the displacements of the "adjacent" grid nodes [4].

The advantage of such a deformation is that it significantly reduces the degrees of freedom in registration, guarantees smoothness of the deformation field and provides a natural partition of the image domain. Let us consider that each control point is endowed with a rectangular support domain $\mathcal{R}_{k,l}$ that is centered to the control point. In such a case the registration functional can be expressed as

$$\min_{a_{k,l},b_{k,l},d_{k,l}} \sum_{k=1}^{K} \sum_{l=1}^{L} \left\{ \int_{\mathcal{R}_i} \rho(s(d(\mathbf{x})), a_{k,l} t(\mathbf{x}) + b_{k,l}) d\mathbf{x} \right\} + \alpha \int_{\Omega} \phi(\nabla d(\mathbf{x})) d\mathbf{x}$$

The number of sub-domains, and their corresponding integrals are not considered as unknowns.

2.2. Similarity Metric

Without loss of generality, let us consider a patch around a control point in a 1-D signal with a support domain $\mathcal{R}_k$. Let us also assume that the control point is in the source image and its intensity distribution is given by $s(d(\mathbf{x}))$. The intensity distribution of the corresponding patch in the target image is given by $a_k t(\mathbf{x}) + b_k$. Our aim is to define a similarity criterion which is minimized when the source image aligns with the target image, irrespective of a_k and b_k. In a discrete domain, let us assume the length of the source and the target to be N. For a signal f of size N, the discrete Haar wavelet transform is defined as, $dwt[i] = \frac{f[2i]+f[2i+1]}{\sqrt{2}}$, when $0 \leq i < \frac{N}{2}$ and $dwt[i] = \frac{f[2i-N]-f[2i+1-N]}{\sqrt{2}}$, when $\frac{N}{2} \leq i < N$, where dwt represents the first level wavelet coefficients. Such an operation is applied recursively on the first half of the signal. It is clear that all the coefficients (other than the first level wavelet coefficients, LL1 in case of 2 levels) comprise of a local difference of intensities in the input signal at various scales. In the process, the constant a_k is scaled up and b_k gets canceled. Applying discrete wavelet transform over the support domain of the control point, would lead to wavelet coefficients of the form $s'(d(\mathbf{x}))$ and $a_k t'(\mathbf{x})$, where s' and t' are the wavelet coefficients of the source and the target respectively. A linear relationship between the intensities of the source and the target images would also be propagated to the wavelet space. The relation between wavelet coefficients would then be given by, $s'(d(\mathbf{x})) = a_k t'(\mathbf{x})$. After taking magnitude of the coefficients and applying logarithm, we have $log(|s'(d(\mathbf{x}))|) = log(|t'(\mathbf{x})|) + log(a_k)$. Using this property, we define the similarity metric $\rho(s(d(\mathbf{x})), t(\mathbf{x})))$ as,

$$\frac{1}{N} \sum_{i=1}^{N} (log(|s'(d(\mathbf{x}))|) - log(|t'(\mathbf{x})|) - log(a_k))$$

In the metric which we have defined, the factor due to linear intensity variation appears as a constant instead of a multiplicative factor. This constant is small when compared to scale of the signals and is spatially stationary. Therefore, it does not alter the nature of the distribution of the similarity metric. Towards accounting for more complex signal variations, we assume a non-linear exponential relationship between the wavelet coefficients, $s'(d(\mathbf{x}))^{\gamma_k} = a_k t'(\mathbf{x})$, where γ_k is the exponential variation in the wavelet coefficients. To compensate for this non-linear change, we define the similarity cost $\rho'(s(d(\mathbf{x})), t(\mathbf{x})), c_k)$ as,

$$\frac{1}{N} \sum_{i=1}^{N} (c_k \cdot \gamma_k \cdot log(|s'(d(\mathbf{x}))|) - log(|t'(\mathbf{x})|) - log(a_k))$$

where ρ' is the joint distribution of the similarity metric with the non-linear exponential constant c_k.

3. GRAPH BASED MULTI-MODAL SUBSPACE FUSION

Inspired by the work of [3], we reformulate registration as a discrete labeling problem. Let us consider a discrete set of labels $\mathcal{L} = \{l^1, l^2, ..., l^i\}$ which correspond to a quantized version of the deformation space $\theta = \{\mathbf{d}^1, ..., \mathbf{d}^i\}$ and a second label set $\otimes = \{\omega^1, \cdots, \omega^j\}$ corresponding to a quantized version of the non-linear transport coefficient c_k, $\Gamma = \{\gamma^1, \cdots, \gamma^j\}$. Based on the label assigned to each

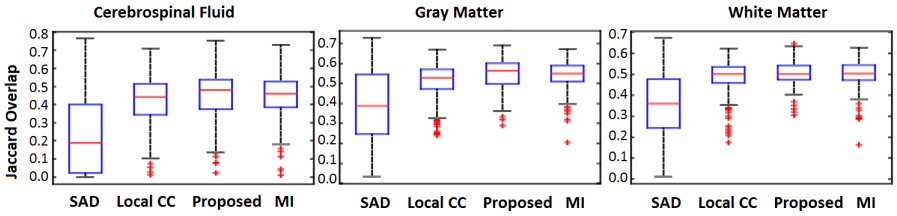

Fig. 1. Jaccard Overlap for different metrics on IBSR dataset

control point $\mathbf{p}$, a dense deformation field is computed as $\mathcal{D}(\mathbf{x}) = \sum_{\mathbf{p}\in G} \eta(|\mathbf{x}-\mathbf{p}|)\mathbf{d}^{l_p}$, where η is the projection function used to interpolate the deformation of the control points. In this context, registration can be formulated as a labelling problem which can be efficiently solved using markov random fields [3] as

$$E_{MRF}(\mathbf{l},\omega) = \sum_{\mathbf{p}\in G} V_{\mathbf{p}}(l_{\mathbf{p}},\omega_{\mathbf{p}}) + \sum_{\mathbf{p}\in G}\sum_{\mathbf{q}\in\mathcal{N}(G)} V_{\mathbf{pq}}(l_{\mathbf{p}},l_{\mathbf{q}})$$

The unary potential corresponding to the deformation and the non-linear correction constant c_k is given by

$$V_{\mathbf{p}}(l_{\mathbf{p}},\omega_{\mathbf{p}}) = \int_{\Omega_{\mathbf{p}}} \rho'(s(\mathbf{d}^{l_{\mathbf{P}}}+\mathbf{x}), t(\mathbf{x}), \gamma^{\omega_p}) d\mathbf{x}$$

In order to impose smoothness on the deformation field, we use the pair-wise terms of the graphical model as mentioned in [3]. For imposing local-consistency on the non-linearity correction, we can impose on the pair-wise term a Potts like model between the distance of the retained coefficients. However, due to the robustness of the metric and its hierarchical nature such a constraint is not necessary.

For recovering the optimal parameters of the MRF, message passing methods efficiently implemented using jump flooding [5] are considered. Further, we use a limited label set and perform several optimization cycles. The control grid is reset after every optimization cycle and the spacing between different labels is reduced by a factor to capture finer displacements.

4. EXPERIMENTS AND RESULTS

The validation of deformable multi-modal registration is a challenging task since the the deformation field for the images is generally not available. We demonstrate the performance and the robustness of the proposed method by providing statistical and visual results on mono-modal and multi-modal datasets.

4.1. Inter Subject Deformable Brain Registration

We performed 306 pairwise registrations on 18 MR-T1 brain images available at Internet Brain Segmentation Repository (http://www.cma.mgh.harvard.edu/ibsr/). In order to evaluate performance, we deform the segmentation available for 3 different brain parts, and compute the Jaccard Overlap between the deformed source segmentation and the target segmentation. From Fig. 1 we can infer that, our metric performs better in terms of median scores among all metrics for "Cerebrospinal Fluid" and "Gray Matter", while being at par with other metrics on "White Mater". In order to compare our approach with state of the art methods, we evaluate 10 registration algorithms [6, 7, 8, 9, 10, 4, 11, 1, 3], on LONI-LPBA40 dataset [12]. 1560 pairwise registrations on 40 skull stripped brain images (181x217x181) were performed and average Jaccard Overlap over 56 segmentation labels was computed. Fig. 2. It can be seen that the proposed approach is comparable to state of the art methods in terms of accuracy while being orders of magnitude faster.

4.2. Multi-Modal Brain Registration

In order to test our metric on multi-modal images, we performed registration on multi-modal brain images from different modalities like MR-T1, MR-T2, CT, PET. The images are part of the online dataset provided by Vanderbilt University (http://www.insight-journal.org/rire/). Registration results of our algorithm are shown in Fig. 2.

4.3. Implementation Details

We tested our approach on a Nvidia Geforce GTX 480 GPU. For estimation of γ, c was varied from 0.1 to 4.9 with increments of 0.15. We add 1 to the absolute value of wavelet coefficients to ensure that the logarithm is positive. In case of mono-modal images it is not necessary to estimate γ. Our metric is free from sequential dependencies and is easily parallelizable. Parallelism is restricted in information theoretic measures due to atomic operations in histogram binning. Further, the computational complexity of our metric is independent of grid resolution unlike information theoretic measures, where run time increases at a finer grid resolution (due to fixed size joint histogram computation).

Experimental validation results presented–other than the proposed method–are courtsey, Ou, Y, 2012. Development and validations of a deformable registration algorithm for medical images

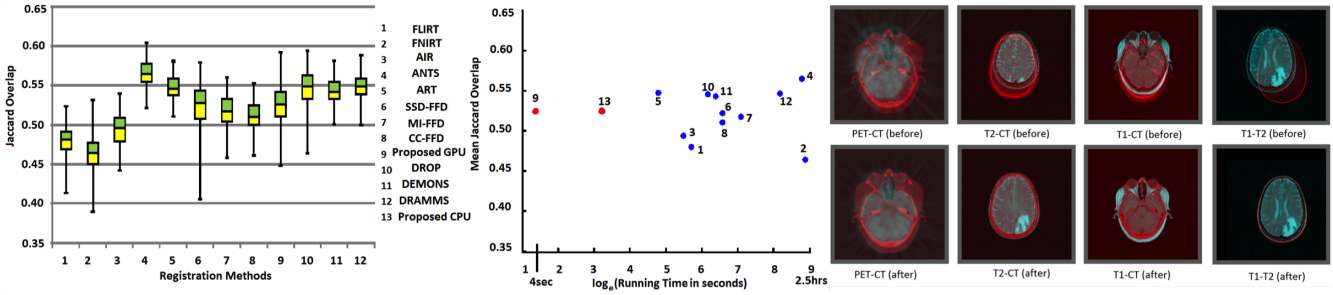

Fig. 2. Average Jaccard Overlap over 56 ROIs for different registration algorithms is shown in the figure on the left. We show the running time of different algorithms in center figure. The red dots indicate the performance of the proposed method. Qualitative registration results on multi-modal images are shown in the figure to the right.

5. DISCUSSION

In this paper we have proposed a novel intensity based similarity metric for mono/multi-modal image registration that once combined with a graphical model and efficient discrete optimization is comparable to the state of the art in terms of accuracy while being an order of magnitude faster. The experimental validation shows that the trade-off between the performance/complexity of the proposed method is exceptional. The underlying metric computation complexity is O(n) and benefits greatly from implementation on modern parallel architectures (GPU). Experimental results on representative multi-modal cases demonstrate the interest in our approach. The exploration of alternative sub-space representations as well as more complex transport functions can be an interesting theoretical extension of the method.

6. REFERENCES

[1] Yangming Ou, Aristeidis Sotiras, Nikos Paragios, and Christos Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," *Medical image analysis*, vol. 15, no. 4, pp. 622, 2011.

[2] P. Viola and W.M. Wells III, "Alignment by maximization of mutual information," *International journal of computer vision*, vol. 24, no. 2, pp. 137–154, 1997.

[3] B. Glocker, N. Komodakis, G. Tziritas, N. Navab, N. Paragios, et al., "Dense image registration through MRFs and efficient linear programming," *Medical image analysis*, vol. 12, no. 6, pp. 731–741, 2008.

[4] D. Rueckert, L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[5] S. Alchatzidis, A. Sotiras, and N. Paragios, "Efficient parallel message computation for map inference," in *ICCV*, 2011, pp. 1379–1386.

[6] Mark Jenkinson, Stephen Smith, et al., "A global optimisation method for robust affine registration of brain images," *Medical image analysis*, vol. 5, no. 2, pp. 143–156, 2001.

[7] J Andersson, S Smith, and M Jenkinson, "FNIRT-FMRIBs non-linear image registration tool," *Human Brain Mapping*, pp. 15–19, 2008.

[8] Roger P Woods, Simon R Cherry, John C Mazziotta, et al., "Rapid automated algorithm for aligning and reslicing PET images.," *Journal of computer assisted tomography*, vol. 16, no. 4, pp. 620, 1992.

[9] Babak A Ardekani, Stephen Guckemus, Alvin Bachman, Matthew J Hoptman, Michelle Wojtaszek, Jay Nierenberg, et al., "Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain mri scans," *Journal of neuroscience methods*, vol. 142, no. 1, pp. 67–76, 2005.

[10] Brian B Avants, CL Epstein, M Grossman, and James C Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[11] Tom Vercauteren, Xavier Pennec, Aymeric Perchant, Nicholas Ayache, et al., "Diffeomorphic demons: Efficient non-parametric image registration.," *NeuroImage*, vol. 45, no. 1, 2008.

[12] David W Shattuck, Mubeena Mirza, Vitria Adisetiyo, Cornelius Hojatkashani, Georges Salamon, Katherine L Narr, Russell A Poldrack, Robert M Bilder, and Arthur W Toga, "Construction of a 3D probabilistic atlas of human cortical structures," *Neuroimage*, vol. 39, no. 3, pp. 1064, 2008.