RANDOM WALKER IMAGE REGISTRATION WITH COST AGGREGATION

Lisa Ying Wai Tang and Ghassan Hamarneh

Medical Image Analysis Lab, School of Computing Science, Simon Fraser University, Canada
{lisat, hamarneh}@sfu.ca

ABSTRACT

Random walker image registration (RWIR) [1] has recently been shown to be a promising method for deformable registration. The computational complexity of RWIR, however, depends on two factors: 1) the number of unknown variables, and 2) the number of discrete values each variable can take. Having large values for both factors increases flexibility in the spatial transformation model, which permits fine features in the deformation fields, thereby allowing for increased accuracy in the solution. However, this comes at a cost of increased memory requirements and computation. In this paper, we propose *cost aggregation* for RWIR to reduce the number of unknowns, such that we group nodes with similar data costs and obtain a non-uniform discretization of the image domain, through which we perform registration. As our experimental results show, we are able to achieve comparable registration accuracy than that achieved by a uniform discretization [1] while lowering complexity.

1. INTRODUCTION

The goal of deformable image registration is to find a spatial transformation $T : \Omega \mapsto \mathbb{R}^N$ that best aligns two images I_a and I_b of dimensionality N. A growing trend in registration is to employ a graph-based approach, which proceeds as follows. The image grid is represented as a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ whose nodes $\mathbf{x}_p \in \mathcal{V}$ represent image coordinates $\mathbf{x}_p$ and edges $e_{pq} \in \mathcal{E}$ encode grid connectivities. The feasible parameter space of $T(\mathbf{x})$ is discretized and represented by a set of K labels $\mathcal{L} = \{\mathbf{v}_1, \cdots, \mathbf{v}_k, \cdots\}$, where $\mathbf{v}_k \in \mathbb{R}^N$. Then, the registration task is cast as a graph-labeling problem where the total cost of label-assignments is measured by a Markov random field energy of the form:

$$\sum_{\mathbf{x}_p \in \mathcal{V}} \mathcal{D}(I_a(\mathbf{x}_p), I_b(\mathbf{x}_p + \mathbf{v}_k)) + \alpha \sum_{e_{pq} \in \mathcal{E}} \mathcal{R}(\mathbf{x}_p, \mathbf{x}_q, \mathbf{v}_j, \mathbf{v}_k)) \quad (1)$$

where $\mathcal{D}$ encodes an image dissimilarity term and $\mathcal{R}$ encodes spatial regularization of the labels (and thus the transformation it represents).

If we employ diffusion regularization [2], then (1) can be solved via a Gaussian Markov random field energy minimization [1–3], which can be optimized analytically with random walks. The overall registration framework, called random walk image registration (RWIR) [1, 4], not only allows for a globally optimal solution but allows for a probabilistic solution, such that a simplex vector (i.e. whose elements sum to unity) is assigned to each node, with entry k encoding the probability of assigning label k to that node.

Just like many other graph-based optimization formulations, e.g. [5, 6], the computational complexity of RWIR depends on: 1) the number of unknown variables, $V = |\mathcal{V}|$ and 2) the number of discrete values each variable can take, $L = |\mathcal{L}|$. In the context of registration, employing large V and L offers flexibility in the spatial transformation model, thus permitting fine features in the resolved deformation fields, which would allow for increased accuracy in the solution. However, this comes at a cost of increased memory requirements and computation.

To reduce the computational complexity of RWIR, we recently developed an approach that reduces L and proposed a strategy that explores the solution search space progressively, adding new labels to an initial label set that was generated from a coarse sampling of the search space [4]. By examining the probabilistic solution and generating labels at confident regions in the solution, our algorithm was capable of enumerating new candidate labels that can better represent the unknown solution search space than the standard coarse-to-fine sampling of the search space.

In this paper, we proceed in the complementary direction of decreasing V to reduce the computational complexity of RWIR. Specifically, we extend and adapt the idea of *cost aggregation* [7] and propose to group neighbouring graph nodes with similar data costs. Performing the aggregation over the entire graph thus forms a non-uniform discretization of the image domain, yielding a smaller graph that effectively decreases the computational complexity of RWIR (V is reduced). As we will demonstrate in Section 3, the proposed RWIR with cost aggregation can achieve registration accuracy comparable to that achieved by uniform discretization [1], but with much lower computational complexity.

We note that the idea of reducing V in registration has been explored in [8, 9], where authors propose to non-uniformly divide the image domain and group nodes to produce a smaller graph representation, thereby reducing V. Nevertheless, the grouping of graph nodes is carried

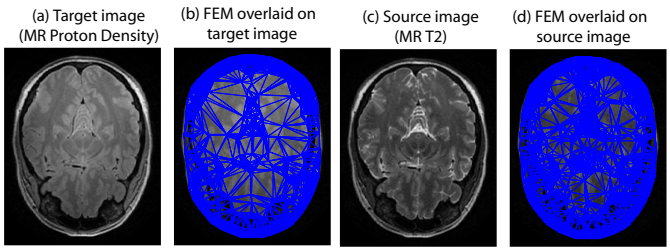

Fig. 1. Intensity-based meshing employed in [2] results in asymmetric discretization of the image domain that is dependent on the image selected as the reference.

out within each image independently. The image domain subdivision and the labelling problem were also preformed independently. These methods thus implicitly assume that the label boundaries (which represents T) are governed by a single image, even though T should be dependent on both the source and target images. As a result, the labelling boundaries may be biased by the image boundaries. Similarly, Popuri et al. [2] proposed to non-uniformly discretize the image domain with a finite-element mesh (FEM) with a meshing strategy, which nonetheless examines information derived only from one image. As their meshing strategy is based on image intensity gradients, it is also unclear how their approach generalizes to multi-modal registration (i.e. switching the source and target images may lead to succinctly different subdivisions, as shown in Fig. 1).

In summary, the approach of grouping pixels into patches based on local intensities [2, 8, 9] had inherently assumed that the label similarities are governed by local similarities within a single image. Conversely, our approach of aggregating data costs subdivide the image domain based on image similarity that is derived from *both* images, which would better relate to the actual registration objective.

2. METHODS

We first review the RWIR algorithm. Then, we present our strategy for integrating cost aggregation into RWIR and a method for interpolating probabilities from the solution resolved for the aggregated graph. To allow for large deformations and to ensure that the obtained solution is diffeomorphic, we embed our ideas in the popular *compositive* registration framework [10]. The overall proposed method is outlined in Algorithm 1.

2.1. Review of random walker image registration

Under RWIR, grid connectivities between the V graph nodes are encoded with a Laplacian matrix L while the data term $\mathcal{D}$ in (1) is encoded with a $V \times K$ matrix $\mathbf{D}$ of data likelihood potentials where $\mathbf{D} = [d_{pk}]_{V \times K}$ represents the data-likelihood of assigning label k to node $\mathbf{x}_p$. Usually, d_{pk} is assumed to be

Algorithm 1 Compositive RWIR with cost aggregation

$T^0 = Id$ `//Initial velocity field`
$RES := \{1, 2, 3, \cdots\}$
$RANGE := \left\{\mathbf{d}_{max}, \frac{\mathbf{d}_{max}}{2}, \frac{\mathbf{d}_{max}}{3}, \cdots\right\}$
`//RES and RANGE determine sampling resolution of the solution space`
for $r = 1 \cdots \text{length}(RES)$ **do**
  Sample search space in $RANGE(r)$ at resolution $RES(r)$ to obtain $\mathcal{L}^r$
  $I_b \leftarrow I_b \circ \exp(T^{r-1})$
  Compute $\mathbf{D}$ using (2) with $\mathcal{L}^r$, I_a, I_b
  Compute $\mathcal{V}^A$ using (4)
  Aggregate data term costs using (5)
  Calculate graph edges and graph weights
  Solve (3) using $\mathbf{D}^A$, $\mathcal{V}^A$, $\mathcal{E}^A$, and $\mathcal{W}^A$
  Compute update T from $\mathbf{U}^A$ using (8)
  Smooth update $T \leftarrow K_{fl.} \star T$ to impose fluid-based regularization
  Update the velocity field $T^r \leftarrow \log(\exp(T^{r-1}) \circ \exp(T))$ via the squaring-and-scaling procedure [10]
  To impose elastic-like regularization, smooth the velocity field $T^r \leftarrow K_{el.} \star T^r$
end for
Compute the final transform as $T = \exp(T^r)$

spatially independent of d_{qk} and may be defined as:

$$d_{pk} = \exp(-\mathcal{D}(F(\mathbf{x}_p), M(\mathbf{x}_p + \mathbf{v}_k)), \tag{2}$$

such that low image dissimilarities between the transformed image $M(\mathbf{x}_p + \mathbf{v}_k)$ and $F(\mathbf{x}_p)$ yield high potential values. One then finds a probabilistic field $\mathbf{U} : \Omega \mapsto \mathbf{u}$ ($\mathbf{u} \in \Delta_K$, the K-dimensional unit simplex) that minimizes:

$$E(\mathbf{U}_k) = \sum_{j=1, j\neq k}^{K} \mathbf{U}_k^T \Lambda_j \mathbf{U}_k + (1 - \mathbf{U}_k)^T \Lambda_k (1 - \mathbf{U}_k) + \alpha \mathbf{U}_k^T L \mathbf{U}_k, \tag{3}$$

where $\mathbf{U}_k$ denotes the k-th component of $\mathbf{U}$; Λ_j is a diagonal matrix with entries $[d_{1k}, \cdots, d_{Vk}]$. As [1, 2] present, the first two terms in (3) encourage $\mathbf{U}$ to be similar to $\mathbf{D}$, while the last term enforces diffusion regularization on $\mathbf{U}$. When $\mathbf{D}$ is normalized to rows with unity sum, the RHS of (3) is positive definite and allows for a unique and global solution for $\mathbf{U}_k$ [11].

2.2. RWIR with Cost Aggregation

In this paper, we extend *cost aggregation* [7] to RWIR. Neighbouring nodes with similar data costs are merged into a single node s. Repeating this merge operation over the entire graph thus yields a set of aggregation nodes, $s \in \mathcal{V}^A$, where $|\mathcal{V}^A| \ll |\mathcal{V}|$. Edges between these nodes, $\mathcal{E}^A$ and corresponding edge weights $\mathcal{W}_A$ are then determined based on spatial proximity and data cost similarity, respectively. Labelling is then performed on the smaller graph $\mathcal{V}^A$, using a Laplacian matrix L_A defined by $\mathcal{V}^A$, $\mathcal{E}^A$, and $\mathcal{W}^A$. Due to

cost aggregation, the graph size is significantly reduced, as we will demonstrate in Section 3.

To compute $\mathcal{V}^A$, we adapted the simple linear iterative clustering (SLIC) method of [12]. Based on k-means clustering, SLIC is a highly efficient superpixel generation algorithm that has $O(n)$ complexity, where n is the number of voxels in the image. As initialization, b clusters are uniformly placed on the image grid with γ voxels in spacing. Then, SLIC assigns each node $\mathbf{x}_p$ to the closest cluster based on an *aggregation distance* [12] (more details below). For computational efficiency, only the clusters in close spatial proximity to $\mathbf{x}_p$ will be considered for assignment (as opposed to considering all clusters). Like k-means, the algorithm then iterates between updating the clusters and re-calculating the node-cluster assignments.

In our framework, we encourage node $\mathbf{x}_p$ to be assigned to a spatially near cluster if their data cost profiles (i.e. image dissimilarity costs) are similar. We thus define aggregation distance as:

$$AD = \sqrt{||\mathbf{D}_p - \mathbf{D}_q||^2 + \beta(||\mathbf{x}_p - \mathbf{x}_q||/\gamma)^2} \tag{4}$$

where $\mathbf{D}_p := [d_{p1} \cdots d_{pk}]$. In words, the first term measures the distance between data similarity profiles of node p and q, the second term measures their spatial distance, and β is a weight that has the effect of regularizing the sizes and shapes of each aggregation node.

With $\mathcal{V}^A$ computed, we next compute the corresponding data costs $\mathbf{D}^A$ and connectivities $\mathcal{E}^A$. Specifically, the data cost associated to each aggregation $s \in \mathcal{V}^A$ and label k is computed as a distance-weighted sum of data costs of all nodes belonging to s, i.e.:

$$d^A_{sk} = \sum_{p \in s} \exp(-||\mathbf{x}_p - \emptyset_s||) d_{pk} \tag{5}$$

where $\emptyset_s$ is the centroid of $s \in \mathcal{V}^A$ and $p \in s$ denotes the nodes that have been aggregated into s. Consequently, nodes closer to $\emptyset_s$ will have larger influence on the data costs of s. An edge is then defined between s and t ($t \in \mathcal{V}^A$), if their centroids are spatially close, i.e. $||\emptyset_s - \emptyset_t|| < \kappa$, where $\kappa \in [3, 5]$ mm. Finally, the edge weight between s and t is computed as:

$$\mathcal{W}^A_{st} = \exp(-||\emptyset_s - \emptyset_t||). \tag{6}$$

such that nodes with spatially close centroids have higher edge weights.

2.3. Computing the transformation

Having solved (3), we obtain the probability matrix $\mathbf{U}^A$ for the aggregated graph. The next step is to derive a probabilistic vector for each node in the original graph, i.e. $p \in \mathcal{V}$, so that a dense displacement field is obtained and the source image can be transformed in the end. In doing so, we first compute:

$$\mathbf{U}_{pk} = \sum_{t \in \mathcal{N}_s} \frac{\exp(-||\mathbf{x}_p - \emptyset_t||)}{Z} \mathbf{U}^A_{tk} \tag{7}$$

where $\mathcal{N}_s$ denotes the neighbourhood of s and Z is a normalization constant, i.e. $Z = \sum_{t \in \mathcal{N}_s} \exp(-||\mathbf{x}_p - \emptyset_t||)$. We then compute the transformation as:

$$T(\mathbf{x}_p) = \sum_{j=1}^{J} g(\mathbf{U}_p, j) \mathbf{v}_{h(\mathbf{U}_p, j)} \tag{8}$$

where h returns the index to the j-th top probable label as defined by $\mathbf{U}_p$ and g returns the corresponding probability. Note that the proposed probability-weighted combination of the top-probable labels in (8) allows for the introduction of new labels via probability-based interpolation.

2.4. Compositive registration

As noted in [2], the energy minimization in (3) does not guarantee a diffeomorphic solution. To ensure diffeomorphism on T, we embed the proposed cost aggregation method into a compositive registration framework [10], as summarized in Algorithm 1. At a high level, this framework incrementally updates the transformation with an update field, which we compute efficiently using the proposed RWIR with cost aggregation method.

3. RESULTS

Two multi-modal datasets were employed to evaluate the proposed method: 1) Proton Density and T2 weighted brain images from Retrospective Image Registration Evaluation (RIRE)[1]; 2) T1 and T2 weighted thigh images from the visible human (VH) project[2].

We compared the proposed cost aggregation (CA) strategy with the multi-resolution (MR) framework originally proposed in [1]. For a direct and fair comparison, we did not use progressive search [4] in both frameworks; hence, the number of labels allowed for each iteration was kept equal under both frameworks. Lastly, we empirically set $J = 8$, and employed the normalized local correlation as the image similarity measure $\mathcal{D}$ in all experiments.

In evaluating the registration accuracy achieved by both frameworks, we performed synthetic experiments, where in each trial, we attempted to recover a random warp that had been introduced to a registered pair. The warps were generated by randomly displacing control points of a B-spline free-form-deformation model, where the magnitude of displacements (in voxels) were sampled from $\mathcal{N}(8, 2)$.

Fig. 2a reports the mean end-point-error (MEPE) obtained for trials performed on the RIRE dataset. Clearly, our proposed method (MR+CA) achieved lower errors in all resolutions than those achieved with a standard MR scheme, but at a significantly lower computational cost (Fig. 2c). As shown in Fig. 2b, this was true under various parameter settings

[1] http://www.insight-journal.org/rire/

[2] ftp://nlmpubs.nlm.nih.gov/visible/bitmaps/mri/

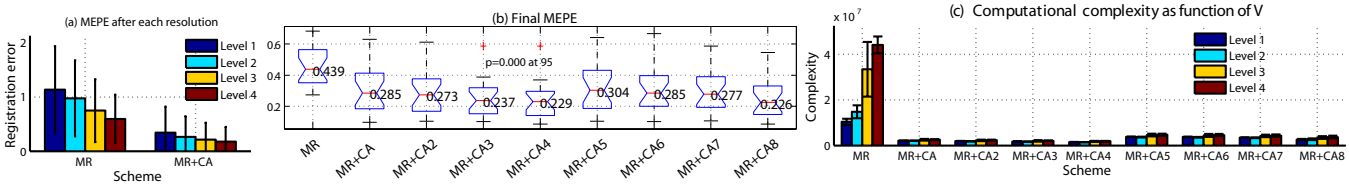

Fig. 2. Effects of proposed cost aggregation strategy as evaluated from 50 trials ran on RIRE images using compositive multi-resolution RWIR (MR-RWIR). (a) MEPE after each level of a 4-level registration scheme with (MR+CA1) and without cost aggregation (MR). Note that in all levels, MR+CA achieved lower error than MR. (b) Final MEPE of MR and of MR+CA (using 8 different parameter settings). (c) Computational complexity of matrix inversion for solving RWIR in each level: $\mathcal{V}$ (number of unknowns) $\times$ $\mathcal{L}$ (label set size). Note that complexity of RWIR without CA was substantially higher than the cases with CA.

($\beta = \{1000, 2000, 3000, 4000\}$ and $\gamma = \frac{\text{ImgWidth}}{12}$ in CA1 to CA4, respectively; CA5 to CA8, $\gamma = \frac{\text{ImgWidth}}{10}$). Fig. 3a reports results from synthetic experiments on the VH images. This time, we repeated the proposed CA with 16 parameter settings. Again, as evident from the Fig. 3b, the proposed CA achieved lower registration error while requiring smaller V. In both sets of experiments, we found no statistically significant difference between the results obtained under the parameter settings tested, thus suggesting that our method is relatively insensitive to these parameters once they are set in the appropriate ranges.

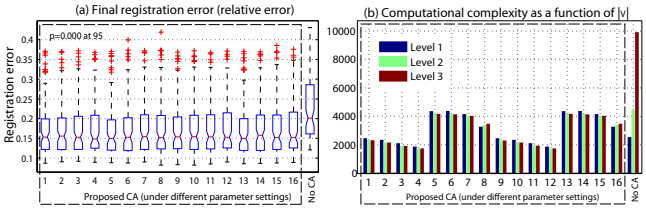

Fig. 3. Effects of proposed cost aggregation strategy as evaluated from 100 trials ran on VH images. (a) Final registration error obtained by RWIR with and without CA. For RWIR with CA, the parameter settings in 1-4 were $\beta = \{1000, 2000, 3000, 4000\}$ with $\gamma = \frac{\text{ImgWidth}}{10}$, $J = 8$; in settings 5-8: changed $\gamma = \frac{\text{ImgWidth}}{12}$; settings 9-16: changed $J = 12$. (b) Average value of V required by each scheme.

4. CONCLUSIONS

We made two improvements to the original RWIR framework [1]: 1) we applied cost aggregation to RWIR to obtain a non-uniform discretization of the image domain and developed methods to compute the transformation from the solution of the aggregated graph, thereby dramatically reducing computational complexity of [1]; and 2) we embedded the strategy into a compositive registration framework that generates diffeomorphic transformation. Our experiments on various multi-modal image pairs showed that our proposed framework performed registration more efficiently without comprising registration accuracy. One future work includes incorporating our proposed framework with precomputation [13].

Acknowledgement. LT and GH are grateful to NSERC for partially funding this work.

5. REFERENCES

[1] D Cobzas and A Sen, "Random walks for deformable image registration," in *MICCAI (2)*, 2011, pp. 557–565.

[2] K Popuri, D Cobzas, and M Jagersand, "A variational formulation for discrete registration," in *MICCAI (3)*, 2013, pp. 187–194.

[3] R Shen, I Cheng, X Li, and A Basu, "Stereo matching using random walks," in *ICPR*, 2008, pp. 1–4.

[4] L Tang and G Hamarneh, "Random walks with efficient search and contextually adapted image similarity for deformable registration," in *MICCAI (2)*, 2013, vol. 8150, pp. 43–50.

[5] L Tang, G Hamarneh, and R Abugharbieh, "Reliability-driven, spatially-adaptive regularization for deformable registration," in *WBIR*, 2010, pp. 173–185.

[6] L Tang, T Bressmann, and G Hamarneh, "Tongue contour tracking in dynamic ultrasound via higher-order MRFs and efficient fusion moves," *Medical Image Analysis*, vol. 16, no. 8, pp. 1503 – 1520, 2012.

[7] D Min, J Lu, and M Do, "A revisit to cost aggregation in stereo matching: How far can we reduce its computational redundancy?," in *ICCV*, 2011, pp. 1567–1574.

[8] M Heinrich, M Jenkinson, B Papiez, F Glesson, M Brady, and J Schnabel, "Edge- and Detail-Preserving Sparse Image Representations for Deformable Registration of Chest MRI and CT Volumes," in *IPMI*, 2013, pp. 463–474.

[9] F Amat, E Myers, and P Keller, "Fast and robust optical flow for time-lapse microscopy using super-voxels," *Bioinformatics*, vol. 29, no. 3, pp. 373–380, 2013.

[10] T Vercauteren, X Pennec, A Perchant, and N Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, Supplement 1, pp. S61 – S72, 2009.

[11] C Couprie, L Grady, L Najman, and H Talbot, "Power watershed: A unifying graph-based optimization framework," *IEEE TPAMI*, vol. 33, no. 7, pp. 1384–1399, 2011.

[12] R Achanta, A Shaji, K Smith, A Lucchi, P Fua, and Sabine Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE TPAMI*, vol. 34, no. 11, pp. 2274–2282, 2012.

[13] S Andrews, L Tang, and G Hamarneh, "Fast random walker image registration using precomputation," in *ISBI*, 2013, pp. 1–8.