MOREA: a GPU-accelerated Evolutionary Algorithm for Multi-Objective Deformable Registration of 3D Medical Images

Georgios Andreadis
Leiden University Medical Center
Leiden, The Netherlands
G.Andreadis@lumc.nl

Peter A.N. Bosman
Centrum Wiskunde & Informatica
Amsterdam, The Netherlands
Peter.Bosman@cwi.nl

Tanja Alderliesten
Leiden University Medical Center
Leiden, The Netherlands
T.Alderliesten@lumc.nl

ABSTRACT

Finding a realistic deformation that transforms one image into another, in case large deformations are required, is considered a key challenge in medical image analysis. Having a proper image registration approach to achieve this could unleash a number of applications requiring information to be transferred between images. Clinical adoption is currently hampered by many existing methods requiring extensive configuration effort before each use, or not being able to (realistically) capture large deformations. A recent multi-objective approach that uses the Multi-Objective Real-Valued Gene-pool Optimal Mixing Evolutionary Algorithm (MO-RV-GOMEA) and a dual-dynamic mesh transformation model has shown promise, exposing the trade-offs inherent to image registration problems and modeling large deformations in 2D. This work builds on this promise and introduces MOREA: the first evolutionary algorithm-based multi-objective approach to deformable registration of 3D images capable of tackling large deformations. MOREA includes a 3D biomechanical mesh model for physical plausibility and is fully GPU-accelerated. We compare MOREA to two state-of-the-art approaches on abdominal CT scans of 4 cervical cancer patients, with the latter two approaches configured for the best results per patient. Without requiring per-patient configuration, MOREA significantly outperforms these approaches on 3 of the 4 patients that represent the most difficult cases.

CCS CONCEPTS

• **Mathematics of computing** → **Evolutionary algorithms**; • **Applied computing** → **Health informatics**.

KEYWORDS

deformable image registration, multi-objective optimization, smart mesh initialization, repair method, GOMEA

1 INTRODUCTION

In recent decades, the field of radiation oncology has experienced rapid developments. Key to its modern practice are medical images acquired before, during, and after treatment. Although these images are already guiding clinical decision-making in many ways, the transfer of information between multiple images that feature large deformations or content mismatches has proven to be a hard challenge and has eluded widespread clinical adoption. In general, the challenge of Deformable Image Registration (DIR) is to find a realistic transformation that matches two or more image spaces to each other, as illustrated in Figure 1. Given this transformation, other metadata could be transferred between images, such as annotated contours [26] or 3D radiation dose distributions [28], opening up opportunities to make radiation treatment more precise [15].

The DIR problem consists of three main objectives: an image-based objective (for a visual comparison), a contour-based objective (for an assessment of object contour overlap), and a realism-based objective (to measure the energy required to perform the deformation). These objectives are conflicting, especially when large deformations and content mismatches are at play [1]. DIR is therefore an inherently multi-objective problem, making Evolutionary Algorithms (EAs) well-suited for its optimization [18].

A diverse set of approaches to DIR has emerged [5, 16, 40]. These all take a single-objective approach, requiring the user to choose the weights associated with the optimization objectives for each use, *a priori*. This can however hinder clinical adoption, since it has been shown that choosing good weights (and other parameters) for specific patients is difficult in general and can strongly influence registration quality [31]. Even when configured for the best results, many existing approaches struggle with large deformations and content mismatches between images because of limitations of their underlying transformation models and (often gradient-descent-based) optimization techniques. This shortcoming forms a second obstacle to their translation into clinical workflows. Therefore, there still is a need for a DIR approach that does not require *a priori* objective weight configuration *and* can tackle large deformations.

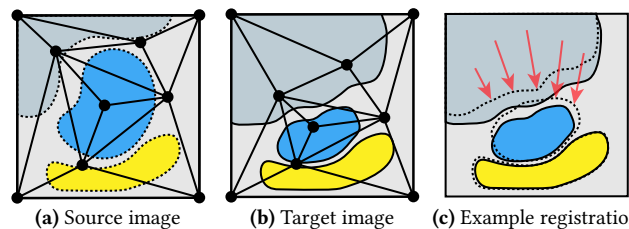**(a)** Source image   **(b)** Target image   **(c)** Example registration

Figure 1: Illustration of two images with large deformations and an example of a deformable image registration with MOREA's dual-dynamic mesh transformation model.

The need to configure objective weights *a priori* has previously been addressed by taking a multi-objective approach [2]. This removes the need to select weights for the optimization objectives in a scalarized problem formulation *a priori*, since a set of solutions can be produced that appropriately represents the trade-off between different conflicting objectives, allowing the user to select a solution from this set, *a posteriori*. To overcome the second obstacle, a flexible dual-dynamic triangular mesh transformation model that allows for inverse-consistent, biomechanical registration has been introduced [3]. This model can match structures on both images to capture large deformations. The Multi-Objective Real-Valued Genepool Optimal Mixing Evolutionary Algorithm (MO-RV-GOMEA) has proven to be effective at performing DIR with this model for 2D images by decomposing the problem into local, partial evaluations [9]. The Graphics Processing Unit (GPU) is exceptionally well-suited to execute these partial evaluations in parallel, yielding significant speed-ups [11]. Recently, first steps have been taken to extend this GPU-accelerated approach to 3D images [4], for which the benefits of partial evaluations may be even greater due to the increase in the amount of image information, leading to more, but also costlier partial evaluations. While this extended approach has been shown to be capable of solving simple registration problems of single objects, it misses several crucial components required to tackle clinical problems that feature multiple interacting objects.

In this work, we therefore introduce MOREA, the first EA-based Multi-Objective Registration approach capable of registering 3D images with large deformations using a biomechanical model, without requiring *a priori* configuration of objective weights. We summarize the motivation for our approach and its main contributions, below:

- One needs **complex, biomechanical models** to solve important, but difficult, problems in deformable medical image registration, especially for 3D medical imaging scans including multiple organs;
- As this induces complex, non-convex, multi-objective optimization problems, one needs **modern, GPU-parallelizable EAs** to do this at a clinically relevant scale;
- This work achieves all this for the **first time**–prior work does not scale due to several missing components;
- Our approach **outperforms state-of-the-art** alternative approaches on real-world complex clinical cases.

---

2 DEFORMABLE IMAGE REGISTRATION FOR LARGE DEFORMATIONS

In this section, we define the DIR optimization problem (Section 2.1) and examine existing approaches (Section 2.2).

2.1 Problem Definition

The problem of DIR for a pair of 2 images is to find a non-rigid transformation T that deforms a source image I_s to match a target image I_t as closely as possible [35]. We distinguish between *unidirectional* and *symmetric* registration: in unidirectional registration, only $T(I_s) \approx I_t$ is optimized, while in symmetric registration, $T'(I_t) \approx I_s$ is also optimized [35]. This can improve the physical

viability of the registration. Another desirable distinction for registrations is *inverse-consistency* [35], guaranteeing a one-to-one correspondence between any point in the source image and its corresponding point in the target image.

Registrations can generally be evaluated according to three classes of quality metrics. *Image intensity* metrics compare the predicted voxel intensity values of $T(I_s)$ to the voxel intensity values of I_t, using metrics such as cross-correlation or mutual information [23]. *Contour* metrics judge registration accuracy by applying T to pairs of sets of points, representing contours (C_s and C_t), and computing the distances between those point sets. One example is the Chamfer distance [19]: for each pair $\langle C_s, C_t \rangle$, the longest minimum distance is calculated between points in $T(C_s)$ and any point in C_t. DIR approaches can also use these contours at initialization time, to build transformation models for use during optimization. Finally, *deformation magnitude* metrics express registration realism by measuring the force needed to apply the deformation, using a physical model of the image space [20]. This can serve as a regularization mechanism, discouraging the registration to overfit.

2.2 Related Work

These three quality metrics are conflicting objectives that form a trade-off [1]. A number of single-objective registration approaches have emerged in recent years, typically attempting to deal with this trade-off by exploring different objective scalarizations. This however has the downside of having to set objective weights, *a priori*. We categorize these existing approaches broadly according to the above defined classes of quality metrics, into classes of approaches mainly optimizing for (1) intensity match, (2) contour match, and (3) both matches simultaneously. These and other features are compared for selected prominent approaches in Table 1.

An example of the first class, optimizing for intensity match, is the Elastix toolbox [24]. It uses a B-spline based transformation model, which uses Bézier curves to model physical space. With this model, Elastix optimizes for intensity, regularized by deformation magnitude metrics. While this is a good fit for many applications, we observe that registering more complex, large deformations with local discontinuities (such as studied in this work) can be difficult. The ANTs SyN registration approach [5] was conceived to model such large deformations, featuring symmetric, inverse-consistent, and intensity-based registration using time-varying velocity fields. A third intensity-based approach is the Demons algorithm [37], using principles from optical flow and Maxwell's Demons for inverse-consistent registration. A more recent version of this approach also has a mechanism to handle content mismatch between images [29]. Both the ANTs and Demons approach can in theory flexibly model large deformations, but lack biomechanical modeling capabilities and only operate on image intensity. This can hamper reliably producing anatomically valid registrations [26].

This is one reason to consider the second class of approaches. One of these approaches is the Thin-Plate Splines Robust Point Matching approach (TPS-RPM), which deforms contours using a thin-plate spline model [17]. Subsequent work validated this on an abdominal test case, registering a deforming bladder and two surrounding organs [39]. There is also a symmetric version of TPS-RPM, which improves robustness on large deformations [7]. Work

Feature	Elastix [24]	ANTs SyN [5]	Demons [37]	TPS-RPM [17]	ANACONDA [40]	MORFEUS [16]	MOREA *(this work)*
Image-based	✓	✓	✓	✗	✓	✗	✓
Contour-based	✗	✗	✗	✓	✓	✓	✓
Biomechanical	✗	✗	✗	✗	✗	✓	✓
Multi-objective	✗	✗	✗	✗	✗	✗	✓

Table 1: Comparison of selected prominent existing DIR approaches by supported registration features.

conducted in parallel also applies a similar model to contours for abdominal registration problems [34]. While large deformations can be modeled, the biomechanical plausibility of the transformation is not guaranteed, and objective weights still require configuration. Another contour-based approach is MORFEUS [16], which registers a mesh representation of imaged objects using a Finite Element Method (FEM) solver. It has shown promising results on brachytherapy applications in the abdomen [32]. Although MORFEUS uses a biomechanical model, which improves realism, it does not take image intensities into account, thus losing detail between object surfaces and relying too heavily on (user-supplied) contours.

Recent work has targeted this shortcoming by proposing a combined contour-based and image-based approach: the ANAtomically CONstrained Deformation Algorithm (ANACONDA) [40] optimizes a fixed scalarization of image and contour terms by using the quasi-Newton algorithm. This approach however lacks biomechanical modeling, and also introduces yet another parameter to configure. Other hybrid attempts have also emerged, such as a combination of the Demons approach with local FEM meshes [42], or the use of an image-based registration step to derive tissue elasticities that are later used in an FEM-based registration approach [25].

In general, we see a gap: an approach that includes all registration aspects in one model. As Table 1 shows, we target this gap with MOREA by being both image-based and contour-based, featuring biomechanical modeling, and exploiting the multi-objective nature of the DIR problem. These novelties are made possible by the flexibility and robustness of EAs, which are well-suited to optimize non-differentiable, multi-objective problems. Additionally, the objective functions include millions of image voxel values and are therefore expensive to compute, calling for hardware acceleration. Modern model-based EAs such as MO-RV-GOMEA feature excellent GPU compatibility, making them a good fit for this problem.

3 MO-RV-GOMEA

The structure of Black-Box Optimization (BBO) problems only gets revealed through repeated function evaluations. Gray-Box Optimization (GBO) problems, on the other hand, have a (partly) known problem structure, which can be exploited during optimization. The GOMEA suite of EAs has proven to be exceptionally well suited for efficiently solving both benchmark and real-world GBO problems [36]. Its extension to multi-objective, real-valued problems, MO-RV-GOMEA [10], has even found real-world adoption in clinical practice for prostate brachytherapy treatment planning [6, 12]. We give an overview of the key working principles of MO-RV-GOMEA here. A detailed description may be found in literature [13].

Non-dominated solutions are preserved across generations in an elitist archive with a pre-specified capacity [27]. Each generation starts with the selection of a subset of non-dominated solutions from the current population. This selection is clustered into k equally sized clusters. For each cluster, MO-RV-GOMEA employs a linkage model that describes dependence relations between variables using a set of dependent variable sets, called Family of Subset (FOS) elements. This linkage model can be learned during optimization in a BBO setting, but in MOREA, we employ a static linkage model based on topological proximity of variables (see Section 4.2.1). Variation then proceeds by considering variables in FOS elements jointly in a procedure called *optimal mixing*. In this step, distributions are estimated for each FOS element in each cluster, and new, partial solutions are sampled from these distributions. Newly sampled partial solutions are evaluated and accepted if their insertion into the parent solution results in a solution that dominates the parent solution or that is non-dominated in the elitist archive.

4 APPROACH

The approach outlined in this work builds on the recently proposed multi-objective approach for 3D images [4]. In this section, we present the new techniques we have added, in modeling the problem (Section 4.1), initializing the population of solutions (Section 4.2), and optimizing the deformations (Section 4.3).

4.1 Modeling

4.1.1 Enhancing realism with tissue-specific elasticities. Adjacent work has indicated that using tissue-specific elasticities, instead of assuming one homogeneous elasticity for the entire image region, can enhance the realism of resulting deformations [32, 41]. Following this insight, we extend the deformation magnitude objective used in existing work [4] by computing an elasticity factor for each tetrahedron, based on its underlying image region. Implementation details for this computation are provided in Appendix A of the supplementary material. We observe in exploratory experiments that this leads to better registration outcomes (see Appendix C.3.1).

To compute the deformation magnitude objective, we consider all corresponding edges e_s and e_t of each tetrahedron $\delta \in \Delta$, belonging to the mesh on the source image and the mesh on the target image, respectively. This includes 4 spoke edges that better capture flattening motion, giving a total of 10 edges per tetrahedron [4]. Given the tetrahedron-specific elasticity constant c_δ, the objective is computed as follows:

$$f_{magnitude} = \frac{1}{10|\Delta|} \sum_{\delta \in \Delta} \left[\sum_{(e_s, e_t) \in E_\delta} c_\delta (\|e_s\| - \|e_t\|)^2 \right]$$

4.1.2 Robustly estimating image similarity. The intensity objective we use is defined as a voxel-to-voxel comparison by taking the sum of squared intensity differences, with special handling for comparisons of foreground (i.e., non-zero intensity) and background (i.e., zero intensity) voxels. We use a random sampling technique that is

well-suited for GPU acceleration (defined in detail in Appendix A). Using the set of all sampled image points on both images, P_s and P_t, and image intensities of source image I_s and target image I_t, the objective is defined as follows:

$$f_{intensity} = \frac{1}{|P_s| + |P_t|} \left[\sum_{p_s \in P_s} h(p_s, T(p_s)) + \sum_{p_t \in P_t} h(p_t, T'(p_t)) \right]$$

$$h(p_s, p_t) = \begin{cases} (p_s - p_t)^2 & p_s > 0 \wedge p_t > 0 \\ 0 & p_s = 0 \wedge p_t = 0 \\ 1 & \text{otherwise} \end{cases}$$

4.1.3 Approximating the guidance error. In contrast to previous work where an exact guidance measure was used as one of the objectives [4], in this work we have opted to introduce a measure that is an approximation thereof that can be much more efficiently computed using the GPU-accelerated sampling method that we already use for the calculation of the values for the image similarity objective. Preliminary experiments showed very similar results (when looking at the voxel displacement fields), also because a perfect guidance error is not necessarily the best solution. In Appendix A, we provide details regarding the implementation.

MOREA's guidance objective is computed at positions P_s and P_t, using the set G of all point set pairs $\langle C_s, C_t \rangle_i$ and the minimal point-to-point-set distance $d(p, C)$. The total number of guidance points is indicated as $|G_s|$ and $|G_t|$, and a truncation radius as r. The guidance objective is now defined as follows:

$$f_{guidance} = \frac{1}{|P_s| + |P_t|} \sum_{\langle C_s, C_t \rangle \in G} \left[\frac{|C_s|}{|G_s|} g(P_s, T, C_s, C_t) + \frac{|C_t|}{|G_t|} g(P_t, T', C_t, C_s) \right]$$

$$g(P, \Phi, C, C') = \sum_{\substack{p \in P \\ d(p,C) < r}} \left[\frac{r - d(p,C)}{r} (d(p,C) - d(\Phi(p), C'))^2 \right]$$

4.1.4 Rapidly computing constraints. MOREA's solutions represent meshes with hundreds of points, which can easily get entangled into folded configurations. Such constraint violations should be prevented, to uphold the guarantee of inverse-consistency. Prior work [4] used a strategy that proved error-prone in more complex meshes. MOREA includes a novel fold detection method that is based on an observed phenomenon: a mesh fold will cause the sign
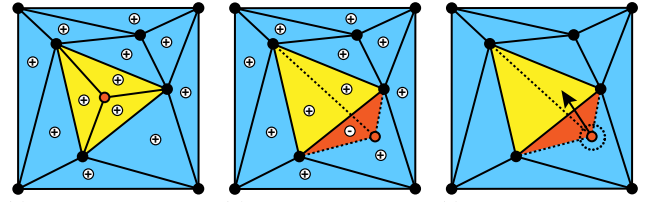
(a) The initial configuration, with positive area signs for each triangle. **(b)** The fold, detected by a sign change in the folded (red) triangle. **(c)** The repair method, resolving the fold by moving the red point.

Figure 2: 2D illustration of a mesh configuration with and without a constraint violation (fold). One of the triangles is folded, due to the red point having moved outside the central triangle, colored yellow. The folded area is colored red.

of at least one tetrahedron's volume to change, as illustrated in Figure 2 (the figure is in 2D, but this also holds in 3D). Our method uses this phenomenon to detect folds and to measure their severity, opening up repair opportunities (see Section 4.3.1). Implementation details for our method are provided in Appendix A.

4.2 Initialization of Registration Solutions

Significant performance gains can be obtained if the initial guesses given to the optimizer are closer to desirable objective space regions than a random guess or grid-like initializations [8]. We introduce two techniques that provide such initial guesses.

4.2.1 Exploiting problem structures with mesh initialization. We initialize the meshes to align with objects in the image, adapting an existing method for 2D images [8] and expanding it to facilitate parallelization on the GPU. First, we place points on the contours of objects in the source image to capture their shape (see Fig. 3a). We choose these points by greedily taking a spread-out subset from the contour annotations also used for the guidance objective, as well as a small fraction of randomly chosen points across the image. Then, we perform a Delaunay tetrahedralization on these points, using the TetGen suite [22] (see Fig. 3b). This yields a mesh that we duplicate to the target image space to complete the dual-dynamic transformation model.

As laid out in Section 3, MO-RV-GOMEA evaluates groups of variables (i.e., FOS elements) jointly during variation. Exploratory experiments have shown that using edges as FOS elements (i.e., groups of two connected points, with the variables encoding their coordinates), is beneficial for this problem. If two FOS elements

(a) Points placed on potentially interesting positions. **(b)** Custom mesh derived from these points. **(c)** Edges selected for variation through set cover. **(d)** Interaction graph (blue) built between selected edges. **(e)** Graph coloring computed on interaction graph.
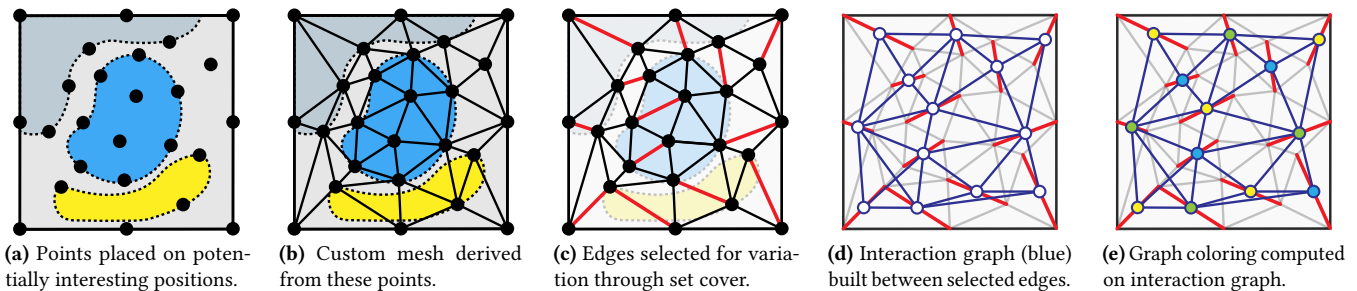
Figure 3: 2D illustration of the mesh initialization process, which produces a custom mesh and determines which groups of edges (i.e., FOS elements) can be optimized in parallel. Selected edges are highlighted in red, interaction edges in blue.

are completely independent because their variables are not needed for the partial evaluation of each set, variation and evaluation for these FOS elements can be done in parallel. We conduct two further steps to facilitate parallel evaluation and optimization on the GPU. First, we execute a greedy set cover algorithm[1] to find a subset of edges that covers all points (see Fig. 3c), so that each variable (point coordinate) undergoes variation. We could alternatively use all edges, but this would lead to points being included in several FOS sets and thus undergoing variation multiple times per generation. For parallelization purposes, it is more efficient to select an (approximately) minimal set of edges.

Given the edge subset found by the set cover, we now determine which FOS elements can be safely optimized in parallel. For this, we build an interaction graph based on topological proximity [11], where two elements are connected if their sets of *dependent tetrahedra* overlap, i.e., the tetrahedra that are reevaluated when an element is changed (see Fig. 3d). Given this graph, parallel groups are created with the DSATUR graph coloring algorithm [14] (see Fig. 3e). The dependent tetrahedra of each parallel group can be evaluated in parallel on the GPU, which has been proven to lead to speed-ups of more than 100x on 2D images [11].

Tetrahedral mesh quality can further be improved by specifying surfaces that should be included in the generated mesh. We apply this principle to the bladder by generating a surface mesh using the Marching Cubes algorithm. We then specify its triangular surfaces as constraints to the mesh generation algorithm, ensuring that bladder surface triangles are included in the mesh. Exploratory experiments show superior performance when using this step (see Section B.3.1 of the supplementary material).

4.2.2 Ensuring diversity in initial population. To promote diversity in the initial population, prior work generates random deviations for each point in the mesh, starting at a grid-initialized solution [4]. We observe that this method can produce many folded mesh configurations in generated solutions, which get discarded and thus hamper convergence speed. In this work, we use a radial-basis-function approach to introduce large deformations free of mesh folds. Implementation details on how these fields are generated and applied to solution meshes are provided in Appendix A.

4.3 Repairing and Steering

During optimization, we apply two techniques to improve the quality of solutions obtained, and the time needed to reach them.

4.3.1 Repairing infeasible solutions. By default, infeasible solutions (i.e., solutions with either of the two meshes having one or more folds) are discarded. This, however, can hamper the creation of high-quality offspring, as infeasible solutions may still provide useful information for higher-quality search space regions. We therefore devise a repair method that attempts to reverse folds on a point-by-point basis. For each point in a folded tetrahedron, the method mutates the point using a Gaussian distribution scaled by its estimated distance to the surrounding 3D polygon. After 64 samples, the change with the best constraint improvement is selected, if present. If all samples result in a deterioration, repair is aborted. The repair process for one point is illustrated in Figure 2c.

4.3.2 Applying pressure with adaptive steering. In general, an approximation set should be as diverse as possible while resembling the Pareto set as closely as possible. In practice, however, not all regions of the Pareto front are of equal interest to users. A user conducting medical DIR for images with large deformations is typically not interested in solutions with a small deformation magnitude. The user is actually most interested in solutions with good guidance objective values, and we would like the algorithm to *steer* its search towards that region in the objective space. Following earlier work [1], we implement an adaptive steering strategy, which steers the front towards high-quality guidance solutions after an exploration period of 100 generations. Given the best guidance objective value s_G of any solution in the elitist archive, we only preserve solutions with guidance objective values between $[s_G; 1.5s_G]$, i.e., this becomes a hard constraint.

5 EXPERIMENTS

We compare MOREA to existing state-of-the-art registration approaches. Due to the complexity of the problem, we do not impose one time limit on all approaches, but rather ensure that they have (reasonably) converged. We repeat all approaches with all configurations 5 times, seeded reproducibly. All MOREA registration runs are run on Dell Precision 7920R machines with NVIDIA RTX A5000 GPUs. Additional information on experimental setup and results is provided in supplementary material.

5.1 Registration Problems

We test all approaches on 4 clinical registration problems with large deformations (see Table 2). We retrospectively select two successive Computerized Tomography (CT) scans of the abdominal area of cervical cancer patients, acquired for radiation treatment planning purposes, with a Philips Brilliance Big Bore scanner. On the first CT scan, the bladder of the patient is filled, and on the second scan, the bladder is empty and thus has shrunken significantly. This large deformation is challenging to register correctly while respecting the surrounding organs (e.g., rectum and bowel) and bony anatomy. Patients 1–3 represent common cases in clinical practice, exhibiting large deformations and little to no margin between bladder and bowel in the full-bladder scan. The bladder of Patient 4 largely preserves its shape and exhibits a wide margin between bladder and bowel, making registration easier. This case, however, is also rarer in practice, and therefore less representative.

The axial slices of the CT scans have a thickness of $3mm$, with in-slice resolutions ranging between $(0.86, 0.86)mm$ and $(1.07, 1.07)mm$. Each scan is resampled to $(1.5, 1.5, 1.5)mm$ for consistency. Afterward, each scan pair is rigidly registered (i.e., translated, rotated, or scaled linearly) to align the bony anatomies of both scans, using bone contours delineated by a radiation therapy technologist (RTT). Each pair is cropped to an axis-aligned bounding box surrounding the bladder with a $30mm$ margin, taking the maximal bounds from both images. This restricts the registration to the region where treatment was delivered, including the surrounding organs at risk.

Contours of key organs in each scan have been annotated by an RTT and verified by a radiation oncologist. The sets of points defining these contours serve as input to the guidance objective
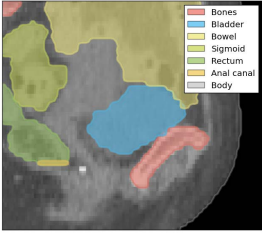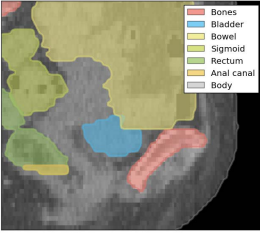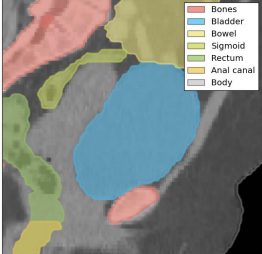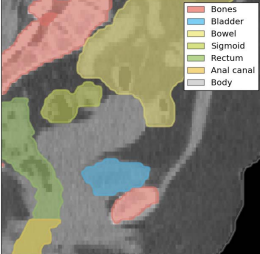
---

[1]Source: https://github.com/martin-steinegger/setcover
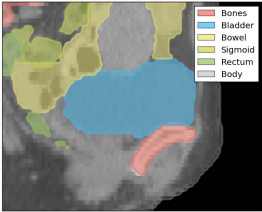
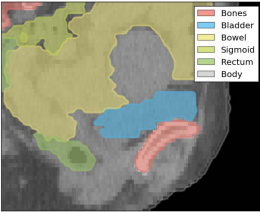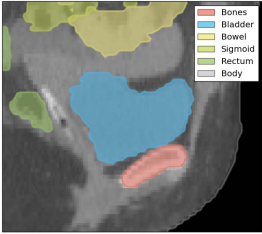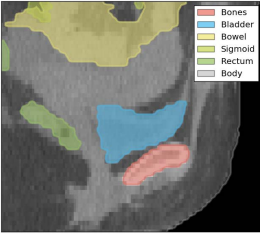| Instance	Source	Target
Patient 1		
Patient 2		
Patient 3		
Patient 4		

Table 2: Sagittal slices of all registration problems, with organs contoured in different colors.

of MOREA. We also use these clinical contours to generate binary masks for each organ and the bones by filling 2D polygonal estimates formed by contours on each slice. As is common in practice, these contours can overlap, since organs are delineated independently and are often surrounded by a small safety margin. Registration approaches therefore need to be robust enough to handle this overlap. Several anatomically relevant corresponding landmarks have been annotated by an RTT and verified by a radiation oncologist on both scans, for evaluation purposes (see Appendix D).

5.2 Registration Approaches

We consider a number of existing, popular registration approaches for which executable code is available. For these approaches, we follow a two-phase configuration process. First, we explore relevant coarse-grained settings for a single patient scan pair (of Patient 1), to find a suitable configuration for the imaging modality and problem difficulty. Then, we conduct fine-grained configuration on the remaining settings (e.g., objective scalarization weights) for each

patient scan pair. We describe the resulting configuration for each approach below, including the general coarse-grained configuration of MOREA. A detailed overview of how we reached these configurations, with additional configuration experiments, can be found in Appendix C of the supplementary material.

5.2.1 Elastix. We configure Elastix to conduct a regularized, multi-resolution [38] image registration. Recommended settings[2] did not yield satisfactory results on our scans, therefore we first register composite mask images onto each other for each patient. This is used as starting point for optimization on the original image intensities. As a fine-grained configuration step for each patient, we configure the weight assigned to the deformation magnitude objective in a fixed sweep of exponentially increasing weights of $[0, 0.001, 0.01, \ldots, 10.0]$, as is done in related work [7].

5.2.2 ANTs SyN. For the ANTs SyN algorithm, the recommended settings[3] for multi-resolution registration also were not satisfactory, which led us to conduct initial configuration experiments with several key parameters, listed in the supplementary material. We also add a composite mask in an additional image channel that is registered alongside the image. For each patient, we test the same regularization weight of the overall deformation by testing the same weights as for Elastix.

5.2.3 This work: MOREA. MOREA uses a single-resolution approach and is configured to generate a mesh of 600 points (i.e., the problem is 3600-dimensional), using the strategies for mesh generation described in Section 4.2. We set the elitist archive capacity to 2000 and use 10 clusters during optimization, with a runtime budget of 500 generations, during which the EA converges (see Appendix D). As MOREA is a multi-objective approach returning an approximation set of registrations, we do not need to configure it further for each patient.

5.3 Evaluation of Registrations

Solutions to complex registration problems, such as the problems in this study, require a multi-faceted evaluation. Below, we outline two main methods for evaluating registrations: surface-based accuracy and visual inspection. Additional methods are described in Appendix Section B.2 and applied in Appendices C and D.

5.3.1 Surface-based registration accuracy. A key part of evaluating registration accuracy is to assess how well the surfaces (contours) of objects align [15]. We use the Hausdorff distance, which represents the largest minimal distance between any two points on two object surfaces. This can be interpreted as the severity of the worst surface match. To account for potential deformation inaccuracies at the border regions of the image, we discard a margin of $15\,mm$ on each side for the computation of this metric. Since this is smaller than the earlier cropping margin of $30\,mm$, the bladder and regions around it are left untouched by this second crop.

5.3.2 Visual inspection. Surface-based accuracy analysis is complemented by a visual inspection, since a registration with a good contour match can still have undesirable deformations in regions

---

[2]Based on an official parameter settings database: https://elastix.lumc.nl/modelzoo/
[3]Based on technical documentation: https://github.com/ANTsX/ANTs/wiki/Anatomy-of-an-antsRegistration-call

(a) Patient 1   **(b)** Patient 2   **(c)** Patient 3   **(d)** Patient 4
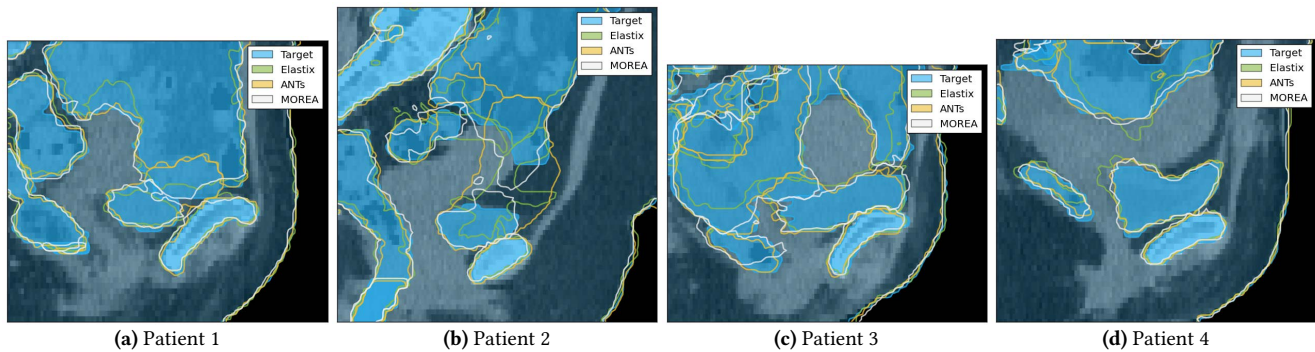
Figure 4: A selection of the best predicted deformations for each patient, represented by deformed contours rendered onto the target image with its reference contours (i.e., target in blue). Annotated slices showing all organs are provided in Table 2.

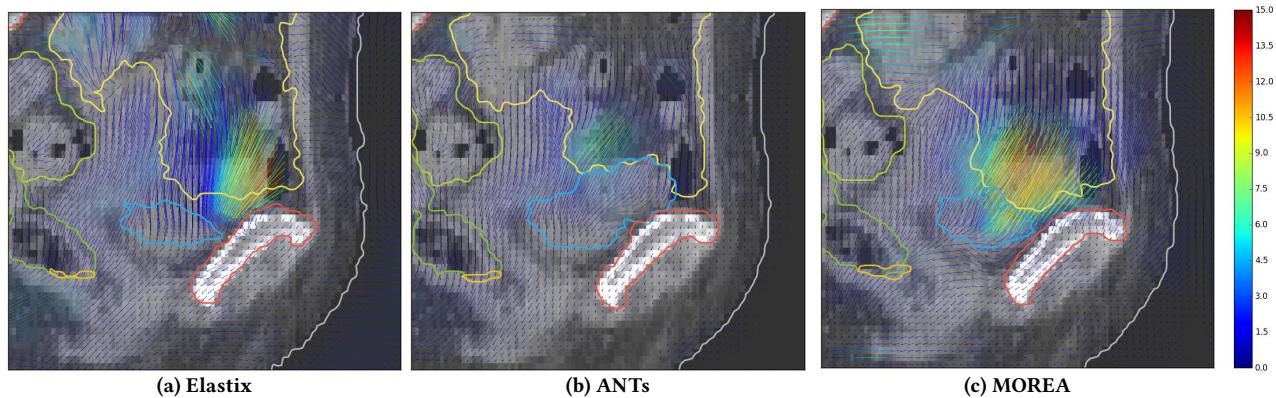**(a) Elastix**   **(b) ANTs**   **(c) MOREA**

Figure 5: Forward deformation vector fields and deformed contours of selected predicted deformations on Patient 1, for all 3 approaches (down-sampled for visibility). Arrow colors represent deformation magnitudes, in voxels (1 voxel = 1.5mm).

between contours. This inspection includes viewing slices of the target image overlaid with the source contours transformed using the computed forward DVF of the registration. To also inspect the deformation between contours, we also visualize the full deformation: First, we render the DVF itself with a quiver plot. Second, we overlay a regular grid onto a slice and deform it with the DVF, which gives a different perspective.

5.4 Comparison of Registrations

All registration solutions from all approaches are compared using the same evaluation pipeline, to ensure a fair comparison. Each approach is configured to output its registrations in the form of a forward and an inverse DVF, which define the deformation on the source and the target image, respectively. Existing approaches either directly or indirectly can be configured to output such DVFs. For MOREA, we rasterize the deformation encoded by the two deformed meshes of a solution, using an existing rasterization method [21]. Since we are comparing single-objective approaches to a multi-objective approach (MOREA), we need to select solutions from MOREA's approximation set. We conduct this *a posteriori* selection by starting at the solution with the best guidance objective value and manually navigating through the approximation front to find a solution with a good trade-off between contour quality and realism.

We also conduct statistical testing using the two-sided Mann-Whitney U test (a standard non-parametric test) to compare MOREA

Problem	MOREA vs. Elastix	MOREA vs. ANTs
Patient 1	**0.011** (+)	**0.007** (+)
Patient 2	**0.007** (+)	**0.007** (+)
Patient 3	**0.012** (+)	**0.007** (+)
Patient 4	**0.007** (+)	0.195 (−)

Table 3: p-values of pair-wise comparisons of Hausdorff distances for the bladder between approaches. A plus (+) indicates a better mean with MOREA, a minus (−) the opposite. Significant results are highlighted.

to ANTs and Elastix. The Hausdorff distance of the bladder contour is used as the test metric, as it describes the largest deforming organ. To correct for multiple tests in the pair-wise comparisons, we apply Bonferroni correction to the α-level and reduce it from 0.05 to 0.025.

6 RESULTS AND DISCUSSION

Figure 4 shows selected outcomes from each per-patient fine-grained configuration experiment, along with a solution from MOREA's approximation front for each patient. For Elastix, we select the runs with regularization weights 1.0, 1.0, 10.0, and 10.0 on Patients 1–4, respectively, and for ANTs, we select all runs with weight 0. The full results of our configuration experiments
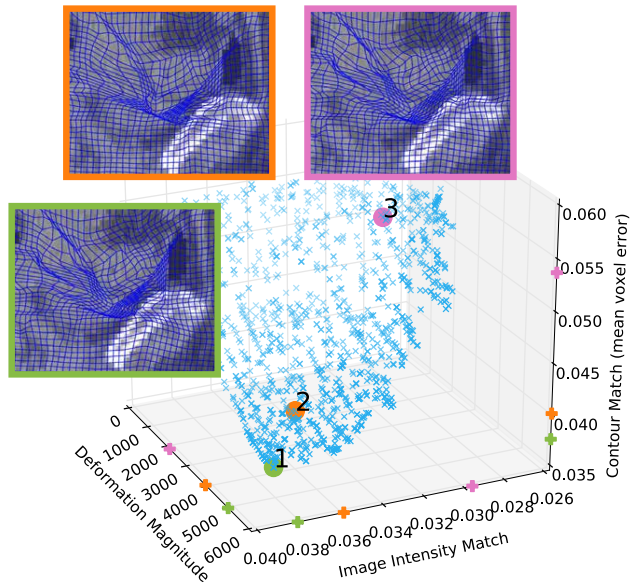
Figure 6: Approximation front produced by MOREA on Patient 1. We render 3 zoomed-in registration solutions.

for both existing approaches can be inspected in Sections B.1.2 and B.2.2 of the supplementary material. Convergence plots for Patient 1, which show how all approaches have converged to the results presented here, can be found in Appendix D. As described in Section 5.1, there is an intrinsic difference in difficulty between the scans. In general, we observe MOREA generally outperforming other approaches on the more difficult patients (1–3), as can be seen visually in the deformed contours shown in Figure 4 and in the additional renders and analyses provided in Appendix D.

For **Patient 1**, we also render DVF slices in Figure 5, showing the transformation computed for each region of one slice. We observe that the deformations returned by Elastix and ANTs only deform the top region of the bladder. MOREA is the only approach which distributes this deformation across the entire bladder, which is a more realistic deformation in this flexible volume. Figure 6 plots the approximation set that is produced by MOREA on Patient 1, highlighting 3 solutions with slightly different deformations. This illustrates the range of solutions presented to the user for selection, all of which spread the deformation across the bladder.

Patient 2, which features the largest volume change in the bladder, seems to prove the most difficult: MOREA comes closest to modeling its deformation (see Fig. 4), although this comes at the cost of the bowel also being moved downwards. A probable cause is the little space (i.e., margin) left between the two organs in the source image. Here, MOREA's result exposes a more fundamental problem that affects all approaches: structures separated by little to no margin in one image cannot be separated in the other image with a transformation model consisting of a single mesh. The change of bladder shape in **Patient 3** is less severe than for Patient 2, but still proves challenging for Elastix and ANTs (see Fig. 4). Especially the back region (located left of the image center) does not match the target. **Patient 4** represents a relatively easy registration problem, with little change in the shape of the bladder and a clear margin between bladder and bowel (see Fig. 2). On this problem, visual

inspection shows that ANTs and MOREA both find a good bladder contour fit, while Elastix struggles with both bladder and bowel.

Examining these results quantitatively, we conduct significance tests on the Hausdorff distance of the bladder, listed in Table 3. In all patients, the contour match of the bladder as deformed by MOREA is significantly superior to Elastix's contour match. ANTs models the contour of the bladder significantly less accurately than MOREA in 3 out of 4 cases, with the fourth case (Patient 4) not having a significantly different result. Appendix D lists significance test results for all organs, which confirm these trends, but also show that MOREA's Hausdorff distance can sometimes be significantly higher than that of ANTs or Elastix. This does not however need to imply worse registration performance, as a qualitative analysis shows. For example, the deformed shape of the sigmoid of Patient 2 found by ANTs is strongly off (see Figure 4). However, its metric value is deemed significantly better than MOREA's, even though MOREA is closer to the target in terms of general shape.

7 CONCLUSIONS

This work uniquely brings multiple lines of research in the field of deformable image registration together. We have introduced a registration approach, MOREA, that is both contour-based and image-based, uses a biomechanical model, and performs multi-objective optimization. This combination uniquely positions MOREA to tackle challenging 3D image registration problems with large deformations and content mismatches. MOREA was built on the MO-RV-GOMEA model-based evolutionary algorithm with several problem-specific extensions, such as GPU acceleration, solution repair, and object-aligned mesh initialization. Our experiments have shown promising results on 4 cervical cancer patient scans, reaching higher contour registration accuracy than two state-of-the-art approaches on 3 of the 4 patients, representing the most difficult cases. Importantly, the deformation produced by MOREA seems to be more uniformly spread across objects than the deformations produced by existing approaches, which is deemed to be more realistic.

Solutions obtained by MOREA still contain local inaccuracies which does leave room for improvement, in particular in regions where organs interface. In fact, the results of this study expose a more fundamental problem in DIR, which is the inability of typical DIR models to capture local discontinuities and content mismatches. This motivates future research into the modeling of independent organ motion, following recent work on this topic [30, 33]. MOREA's extensible, biomechanical model could be well-positioned for expansions to capture these phenomena. Given such an expanded approach, a larger validation study, with more patients and involving domain experts, could help close the gap to clinical practice.

REFERENCES

[1] T. Alderliesten, P. A. N. Bosman, and A. Bel. 2015. Getting the most out of additional guidance information in deformable image registration by leveraging multi-objective optimization. In *SPIE Medical Imaging 2015: Image Processing*. 94131R.

[2] T. Alderliesten, J. J. Sonke, and P. A. N. Bosman. 2012. Multi-objective optimization for deformable image registration: proof of concept. In *SPIE Medical Imaging 2012: Image Processing*, Vol. 8314. 831420.

[3] T. Alderliesten, J. J. Sonke, and P. A. N. Bosman. 2013. Deformable image registration by multi-objective optimization using a dual-dynamic transformation model to account for large anatomical differences. In *SPIE Medical Imaging 2013: Image Processing*, Vol. 8669. 866910.

[4] G. Andreadis, P. A. N. Bosman, and T. Alderliesten. 2022. Multi-objective dual simplex-mesh based deformable image registration for 3D medical images - proof of concept. In *SPIE Medical Imaging 2022: Image Processing*. 744–750.

[5] B. B. Avants, C. L. Epstein, M. Grossman, and J. C. Gee. 2008. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* 12, 1 (2008), 26–41.

[6] D. L. J. Barten, B. R. Pieters, A. Bouter, M. C. van der Meer, S. C. Maree, K. A. Hinnen, H. Westerveld, P. A. N. Bosman, T. Alderliesten, N. van Wieringen, and A. Bel. 2023. Towards artificial intelligence-based automated treatment planning in clinical practice: A prospective study of the first clinical experiences in high-dose-rate prostate brachytherapy. *Brachytherapy* In Press (2023).

[7] L. Bondar, M. S. Hoogeman, E. M. Vásquez Osorio, and B. J.M. Heijmen. 2010. A symmetric nonrigid registration method to handle large organ deformations in cervical cancer patients. *Medical Physics* 37, 7 (2010), 3760–3772.

[8] P. A. N. Bosman and T. Alderliesten. 2016. Smart grid initialization reduces the computational complexity of multi-objective image registration based on a dual-dynamic transformation model to account for large anatomical differences. In *SPIE Medical Imaging 2016: Image Processing*. 978447.

[9] A. Bouter, T. Alderliesten, and P. A. N. Bosman. 2017. A novel model-based evolutionary algorithm for multi-objective deformable image registration with content mismatch and large deformations: benchmarking efficiency and quality. In *SPIE Medical Imaging 2017: Image Processing*, Vol. 10133. 1013312.

[10] A. Bouter, T. Alderliesten, and P. A. N. Bosman. 2021. Achieving highly scalable evolutionary real-valued optimization by exploiting partial evaluations. *Evolutionary Computation* 29, 1 (2021), 129–155.

[11] A. Bouter, T. Alderliesten, and P. A. N. Bosman. 2021. GPU-Accelerated Parallel Gene-pool Optimal Mixing applied to Multi-Objective Deformable Image Registration. In *IEEE Congress on Evolutionary Computation*. 2539–2548.

[12] A. Bouter, T. Alderliesten, B. R. Pieters, A. Bel, Y. Niatsetski, and P. A. N. Bosman. 2019. GPU-accelerated bi-objective treatment planning for prostate high-dose-rate brachytherapy. *Medical Physics* 46, 9 (2019), 3776–3787.

[13] A. Bouter, N. H. Luong, C. Witteveen, T. Alderliesten, and P. A. N. Bosman. 2017. The multi-objective real-valued gene-pool optimal mixing evolutionary algorithm. In *Proceedings of the 2017 Genetic and Evolutionary Computation Conference*. 537–544.

[14] D. Brélaz. 1979. New methods to color the vertices of a graph. *Commun. ACM* 22, 4 (1979), 251–256.

[15] K. K. Brock, S. Mutic, T. R. McNutt, H. Li, and M. L. Kessler. 2017. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132: Report. *Medical Physics* 44, 7 (2017), e43–e76.

[16] K. K. Brock, M. B. Sharpe, L. A. Dawson, S. M. Kim, and D. A. Jaffray. 2005. Accuracy of finite element model-based multi-organ deformable image registration. *Medical Physics* 32, 6 (2005), 1647–1659.

[17] H. Chui and A. Rangarajan. 2000. A new algorithm for non-rigid point matching. In *IEEE Conference on Computer Vision and Pattern Recognition*. 44–51.

[18] K. Deb. 2001. *Multi-Objective Optimization using Evolutionary Algorithms*. Wiley.

[19] M. Faisal Beg, M. I. Miller, A. Trouvétrouv, and L. Younes. 2005. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. *International Journal of Computer Vision* 61, 2 (2005), 139–157.

[20] B. Fischer and J. Modersitzki. 2008. Ill-posed medicine - An introduction to image registration. *Inverse Problems* 24, 3 (2008), 1–16.

[21] J. Gascon, J. M. Espadero, A. G. Perez, R. Torres, and M. A. Otaduy. 2013. Fast deformation of volume data using tetrahedral mesh rasterization. In *Proceedings - SCA 2013: 12th ACM SIGGRAPH / Eurographics Symposium on Computer Animation*. 181–186.

[22] S. Hang. 2015. TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator. *ACM Trans. Math. Software* 41, 2 (2015), 1–36.

[23] F. Khalifa, G. M. Beache, G. Gimel'farb, J. S. Suri, and A. S. El-Baz. 2011. State-of-the-Art Medical Image Registration Methodologies: A Survey. In *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*. Springer, 235–280.

[24] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P.W. Pluim. 2010. Elastix: A toolbox for intensity-based medical image registration. *IEEE Transactions on Medical Imaging* 29, 1 (2010), 196–205.

[25] M. Li, E. Castillo, X. L. Zheng, H. Y. Luo, R. Castillo, Y. Wu, and T. Guerrero. 2013. Modeling lung deformation: A combined deformable image registration method with spatially varying Young's modulus estimates. *Medical Physics* 40, 8 (2013), 1–10.

[26] G. Loi, M. Fusella, E. Lanzi, E. Cagni, C. Garibaldi, G. Iacoviello, F. Lucio, E. Menghi, R. Miceli, L. C. Orlandini, Antonella Roggio, Federica Rosica, Michele Stasi, Lidia Strigari, Silvia Strolin, and Christian Fiandra. 2018. Performance of commercially available deformable image registration platforms for contour propagation using patient-based computational phantoms: A multi-institutional study. *Medical Physics* 45, 2 (2018), 748–757.

[27] H. N. Luong and P. A. N. Bosman. 2012. Elitist Archiving for Multi-Objective Evolutionary Algorithms: To Adapt or Not to Adapt. In *Proceedings of the 12th Conference on Parallel Problem Solving from Nature*. 72–81.

[28] R. Mohammadi, S. R. Mahdavi, R. Jaberi, Z. Siavashpour, L. Janani, A. S. Meigooni, and R. Reiazi. 2019. Evaluation of deformable image registration algorithm for determination of accumulated dose for brachytherapy of cervical cancer patients. *Journal of Contemporary Brachytherapy* 11, 5 (2019), 469–478.

[29] S. Nithiananthan, S. Schafer, D. J. Mirota, J. W. Stayman, W. Zbijewski, D. D. Reh, G. L. Gallia, and J. H. Siewerdsen. 2012. Extra-dimensional Demons: A method for incorporating missing tissue in deformable image registration. *Medical Physics* 39, 9 (2012), 5718–5731.

[30] D. F. Pace, M. Niethammer, and S. R. Aylward. 2012. Sliding Geometries in Deformable Image Registration. In *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging*. 141–148.

[31] K. Pirpinia, P. A. N. Bosman, C. E. Loo, G. Winter-Warnars, N. N. Y. Janssen, A. N. Scholten, J. J. Sonke, M. van Herk, and T. Alderliesten. 2017. The feasibility of manual parameter tuning for deformable breast MR image registration from a multi-objective optimization perspective. *Physics in Medicine and Biology* 62, 14 (2017), 5723–5743.

[32] B. Rigaud, A. Klopp, S. Vedam, A. Venkatesan, N. Taku, A. Simon, P. Haigron, R. De Crevoisier, K. K. Brock, and G. Cazoulat. 2019. Deformable image registration for dose mapping between external beam radiotherapy and brachytherapy images of cervical cancer. *Physics in Medicine and Biology* 64, 11 (2019), 115023.

[33] L. Risser, F. X. Vialard, H. Y. Baluwala, and J. A. Schnabel. 2013. Piecewise-diffeomorphic image registration: Application to the motion estimation between 3D CT lung images with sliding conditions. *Medical Image Analysis* 17, 2 (2013), 182–193.

[34] B. Schaly, J. A. Kempe, G. S. Bauman, J. J. Battista, and J. van Dyk. 2004. Tracking the dose distribution in radiation therapy by accounting for variable anatomy. *Physics in Medicine and Biology* 49, 5 (2004), 791–805.

[35] A. Sotiras and N. Paragios. 2012. *Deformable Image Registration: A Survey*. Technical Report. Center for Visual Computing, Department of Applied Mathematics, Ecole Centrale de Paris, Equipe GALEN, INRIA Saclay.

[36] D. Thierens and P. A. N. Bosman. 2011. Optimal mixing evolutionary algorithms. In *Proceedings of the 2011 Genetic and Evolutionary Computation Conference*. 617–624.

[37] J.-P. Thirion. 1998. Image matching as a diffusion process: an analogy with Maxwell's Demons. *Medical Image Analysis* 2, 3 (1998), 243–260.

[38] M. Unser, A. Aldroubi, and C. R. Gerfen. 1993. Multiresolution image registration procedure using spline pyramids. In *SPIE Mathematical Imaging: Wavelet Applications in Signal and Image Processing*, Vol. 2034. 160–170.

[39] E. M. Vásquez Osorio, M. S. Hoogeman, L. Bondar, P. C. Levendag, and B. J. M. Heijmen. 2009. A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy. *Medical Physics* 36, 7 (2009), 2848–2859.

[40] O. Weistrand and S. Svensson. 2015. The ANACONDA algorithm for deformable image registration in radiotherapy. *Medical Physics* 42, 1 (2015), 40–53.

[41] S. Wognum, L. Bondar, A. G. Zolnay, X. Chai, M. C. C. M. Hulshof, M. S. Hoogeman, and A. Bel. 2013. Control over structure-specific flexibility improves anatomical accuracy for point-based deformable registration in bladder cancer radiotherapy. *Medical Physics* 40, 2 (2013), 1–15.

[42] H. Zhong, J. Kim, H. Li, T. Nurushev, B. Movsas, and I. J. Chetty. 2012. A finite element method to correct deformable image registration errors in low-contrast regions. *Physics in Medicine and Biology* 57, 11 (2012), 3499–3515.