A first step toward uncovering the truth about weight tuning in deformable image registration

Kleopatra Pirpinia*[a], Peter A. N. Bosman[b], Jan-Jakob Sonke[a], Marcel van Herk[c,d],
Tanja Alderliesten[e]

[a] Dept. of Radiation Oncology, Netherlands Cancer Institute (NKI), P.O. Box 90203, 1006 BE Amsterdam, The Netherlands
[b] Centrum Wiskunde & Informatica (CWI), P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
[c] Institute of Cancer Sciences, University of Manchester, Wilmslow Road, Manchester, M20 4BX Manchester, United Kingdom
[d] Dept. of Biomedical Engineering and Physics, Academic Medical Center (AMC), P.O. Box 22660, 1100 DD Amsterdam, The Netherlands
[e] Dept. of Radiation Oncology, Academic Medical Center (AMC), P.O. Box 22660, 1100 DD Amsterdam, The Netherlands

ABSTRACT

Deformable image registration is currently predominantly solved by optimizing a weighted linear combination of objectives. Successfully tuning the weights associated with these objectives is not trivial, leading to trial-and-error approaches. Such an approach assumes an intuitive interplay between weights, optimization objectives, and target registration errors. However, it is not known whether this always holds for existing registration methods. To investigate the interplay between weights, optimization objectives, and registration errors, we employ multi-objective optimization. Here, objectives of interest are optimized simultaneously, causing a set of multiple optimal solutions to exist, called the optimal Pareto front. Our medical application is in breast cancer and includes the challenging prone-supine registration problem. In total, we studied the interplay in three different ways. First, we ran many random linear combinations of objectives using the well-known registration software `elastix`. Second, since the optimization algorithms used in registration are typically of a local-search nature, final solutions may not always form a Pareto front. We therefore employed a multi-objective evolutionary algorithm that finds weights that correspond to registration outcomes that do form a Pareto front. Third, we examined how the interplay differs if a true multi-objective (i.e., weight-free) image registration method is used. Results indicate that a trial-and-error weight-adaptation approach can be successful for the easy prone to prone breast image registration case, due to the absence of many local optima. With increasing problem difficulty the use of more advanced approaches can be of value in finding and selecting the optimal registration outcomes.

Keywords: Deformable image registration, multi-objective optimization, elastix, weights, evolutionary algorithms

1. INTRODUCTION

Deformable image registration is the task of finding the optimal nonlinear transformation to align two images. Of interest in deformable registration is not only similarity, but also anatomical correctness achieved by realistic deformations. Often, deformable image registration is formulated and solved as a single-objective optimization problem. As such, there is one objective function to be optimized, which is most commonly defined as a weighted linear combination of two or more terms that describe the objectives of interest, such as image similarity and deformation magnitude. Before the start of the registration, many parameters have to be manually defined and tuned. An important subset of these parameters is the set of weights associated with the terms that constitute the objective function, because they implicitly define the trade-off between the often conflicting objectives. Therefore, successfully tuning the weights that express the right trade-off is very important for finding the best registration outcome. However, the relation between the different objectives in deformable image registration and thus the nature of these trade-offs, are not fully understood. As a consequence, there is

* E-mail: k.pirpinia@nki.nl; Telephone: +31 (0)20 512 2235

little to no insight into how these weights can be determined a priori. A spatially adaptive regularization technique where the weight associated with the regularization term is adaptively determined during the optimization run, has been proposed.[1] Other methods employ machine-learning techniques to learn hundreds of registration parameters for a specific task.[2] These methods have been tested on synthetic data as well as on brain image registration problems with little amount of deformation but a high degree of regularization needed. In this work, we focus on another class of problems, the challenge of which lies in the large amount of deformation that is required in order to achieve a good registration outcome. Furthermore, for the vast majority of the currently widely used registration methods, one still must rely on intuition to determine the weight that best expresses the preference of the user. As a consequence, weights are usually tuned manually for each specific registration task via trial-and-error.

The success of such trial-and-error approaches relies on a sensible and smooth relation between the conflicting objectives, which corresponds to expected changes when weights change. However, since there is little to no insight into the nature of these inherent trade-offs, it is not guaranteed that trial-and-error approaches always succeed, especially in hard deformable image registration cases, such as prone-supine breast image registration. For this reason, we approach deformable image registration from a multi-objective optimization standpoint. We aim to study the interplay between key objectives used in deformable image registration by comparing weight distributions, registration outcomes and associated registration error distributions obtained in three different ways, on three problems of increasing difficulty.

2. METHODS

2.1 Registration as an optimization problem

Deformable image registration is predominantly formulated as a single-objective optimization problem, where the goal is to find the optimal transformation from one image (referred to as the moving image) to the other image (referred to as the fixed image). The cost function to be optimized can consist of multiple terms that express objectives of interest, such as image similarity, smoothness of the transformation, volume preservation etc. A cost function for deformable image registration that is to be *minimized* in case of two objectives is typically of the following form:[3]

$$C(x) = S(I_f, I_m \circ T(x)) + \lambda\, R(T(x)). \tag{1}$$

The first term S is the objective describing the degree of (dis)similarity between the two images. It is dependent on the fixed and moving images I_f, I_m, and the transformation T that deforms the moving image and is parameterized by x. The second term R is related to the deformation magnitude, which it controls by regularizing the transformation T and penalizing large deformations. λ is the weight associated with the deformation objective that determines the trade-off between deformation magnitude and dissimilarity. These two objectives are conflicting; lower dissimilarity often requires a highly non-regular deformation field, but by allowing deformation fields to become too irregular, optimization may lead to overfitting the dissimilarity objective with undesirable nonrealistic deformations. Therefore, dissimilarity needs to be minimized while also the objective related to the deformation magnitude remains as low as possible. It is clear then that the underlying optimization problem in deformable image registration is actually multi-objective, i.e., there are multiple, conflicting, objectives that need to be optimized simultaneously, and optimal decisions need to be taken in the presence of trade-offs between these objectives.

2.1.1 Multi-objective deformable image registration

In multi-objective optimization, a vector of m objective functions $F(x) = (f_0(x), f_1(x), \ldots, f_{m-1}(x))$ needs to be optimized. These objectives are usually conflicting; as a result, there is no unique optimal solution to this problem and the notion of optimality is redefined. To this end, the concepts of Pareto domination and Pareto optimality are introduced. Without loss of generality, we assume that all objectives are formulated in such a way that they need to be minimized. A solution x_1 is said to (Pareto) dominate a solution x_2 if and only if $f_i(x_1) \leq f_i(x_2)$ holds for all $i \in \{0, 1, \ldots, m-1\}$ and $f_i(x_1) < f_i(x_2)$ holds for at least one $i \in \{0, 1, \ldots, m-1\}$. A Pareto set of size n is a set of solutions $x_j \in \{0, 1, \ldots, n-1\}$ in which no solution dominates any other solution. A Pareto front corresponding to a Pareto set is the mapping of the solutions to the m-dimensional objective space, i.e., the set of all m-dimensional vectors $F(x_j), j \in \{0, 1, \ldots, n-1\}$. A solution x_1 is said to be Pareto optimal if and only if there is *no* other solution x_2 that (Pareto) dominates it. Further, the optimal Pareto set is the set of all Pareto-optimal solutions and the optimal Pareto front is the Pareto front that corresponds to the optimal Pareto set.
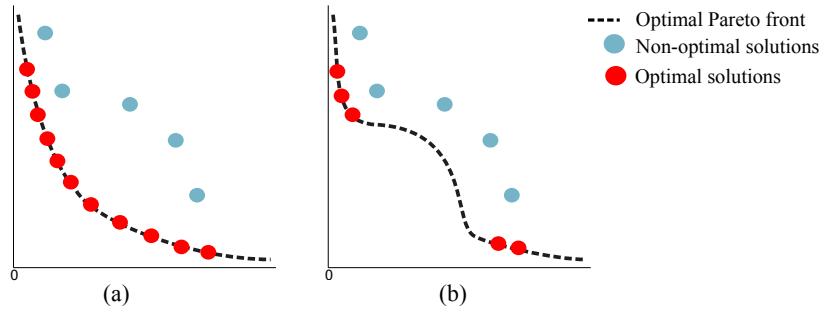
Figure 1. Schematic visualization of (a) a convex Pareto front, (b) Pareto front with a concave part in a bi-objective problem. Optimal solutions corresponding to different linear combinations are depicted in red.

Following from the above, in a bi-objective problem the objective space is a plane and each solution can be represented as a point in that plane. The optimal Pareto front can then be visualized as a curve in the plane. A schematic visualization of optimal Pareto fronts and (non-)optimal solutions in objective space can be seen in Figure 1.

One of the most common methods for solving multi-objective optimization problems, including deformable image registration, is linear scalarization. With linear scalarization multiple objectives are combined into a single-objective function via a weighted linear combination (as, e.g., in (1)), which is subsequently solved by single-objective optimization algorithms. A prerequisite then is to a priori define the weights that represent the desired trade-off between the objectives. However, having limited knowledge of the problem, it is not trivial for the decision maker to choose the weight that best expresses their preference. Furthermore, such an approach will be highly sensitive to the shape of the optimal Pareto front. If the optimal Pareto front is convex, then *any* weight combination will yield an outcome on the optimal Pareto front, assuming that the optimization algorithm used can find the global optimum. However, it is proven that linear scalarization cannot reach optimal trade-offs that lie in concave parts of the optimal Pareto front (Figure 1).[4] Furthermore, a uniform sampling of weights may not necessarily lead to a uniform sampling of solutions on the optimal Pareto front.[4] Finally, although this paper focuses on bi-objective problems, it should be noted that these issues with linear scalarization are exacerbated if more objectives are introduced (which is not uncommon in deformable image registration).

In this work, we consider deformable image registration as a problem with two objectives as formulated in (1). The specific objectives we consider are image dissimilarity and deformation magnitude that both need to be minimized. For this problem, we have limited knowledge of the shape of the optimal Pareto front. Moreover, the gradient-based optimization algorithms that are widely used to solve it, albeit powerful, especially when combined with multi-resolution schemes, can still get stuck in local optima. Therefore, it is not guaranteed that solutions on the optimal Pareto front will be obtained by solving the problem for different linear scalarizations. For this reason we do not know if a change in weights leads to an expected change in outcome, in terms of objective values and target registration error (i.e., a clinically relevant difference in location of corresponding anatomy). This is however crucial to determine the true utility of tuning weights by trial and error for existing single-objective optimization-based registration methods and to what extent multi-objective optimization is desirable. We will investigate how weight distribution translates in solution distribution in objective space as well as in registration error distribution for a typical single-objective registration method in various ways. We will use an open-source, single-objective registration method, to perform registration for multiple *randomly sampled* linear combinations. We will also use a multi-objective evolutionary algorithm (EA) to directly search for the best possible linear combinations to use with `elastix`. Further, we will use a true multi-objective, i.e., weight-free, image registration method that directly approximates the entire optimal Pareto front, i.e., not using linear combinations at all.

Table 1. Overview of test cases and methods.

Test datasets		elastix			Weight-free multi-objective method
		Random sampling	Optimized linear combinations		
A	Prone-prone	3D	3D	2D	2D
B	Volunteer prone-supine	3D	3D	2D	2D
C	Patient prone-supine	3D	3D	2D	2D

2.2 `elastix`

As our primary single-objective-based registration tool we used `elastix`, a well-known intensity-based image registration toolbox.[5] Within `elastix`, we used a B-spline transformation model[6] and an adaptive stochastic gradient descent algorithm[7] as the optimizer. We varied the weight λ associated with the deformation magnitude objective. Each λ corresponds to a different linear combination. In principle, a low value for this weight should lead to solutions with low dissimilarity and relatively large deformation. The similarity metric we used is the normalized correlation coefficient (NCC), whereas the deformation magnitude objective is the bending energy penalty,[6] which depends on the second order derivatives of the B-spline transformation. Moreover, for the registration problems at hand we used a multi-level pyramid scheme. Here, 3 registration levels were used for an easy test case, and 5 levels were used for two more challenging test cases.

2.2.1 Random sampling of linear combinations

We first randomly sampled one thousand weights, which correspond to one thousand different linear combinations. In this way, we can observe how the value of the weights is related to the location of the solutions in objective space. The value of the randomly sampled weights ranged from 0 to 10^3.
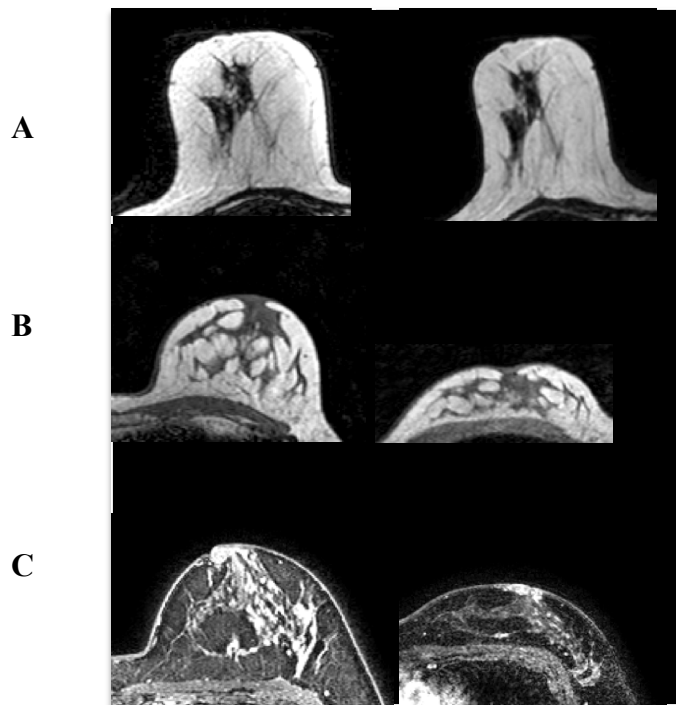

Figure 2. From left to right: moving image, fixed image. From top to bottom: test case A, test case B, and test case C.
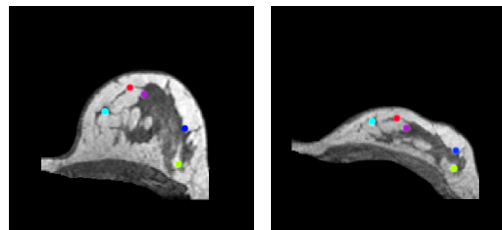

Figure 3. Example of annotated slices from the volunteer prone-supine breast MRI case (test case B).

2.2.2 Optimized linear combinations

Given that a random weight λ does not necessarily lead to a solution on the optimal Pareto front, it makes sense to consider only weights that lead to the optimal solutions in objective space. This does however require an additional optimization loop. To do so, we employed a multi-objective EA. EAs are state-of-the-art in multi-objective optimization, being global, population-based optimization methods.[8] These properties make them very suitable for solving multi-objective optimization problems, since they are able to evolve populations of solutions altogether to approximate the optimal Pareto front in one run. The EA we used in particular is an Estimation-of-Distribution algorithm[9], meaning that it builds a Gaussian distribution from the population. It subsequently samples the distribution, selects the best solutions and then adapts its estimation based on the sampled solutions. In this case, the EA was used to find weights that lead to registration outcomes with the best trade-offs (i.e., a Pareto front) between dissimilarity and deformation magnitude. The values for the objectives are obtained by running `elastix`, which performs the registration, during the run of the EA (as part of a multi-objective function evaluation).

2.3 Weight-free multi-objective image registration method

Multi-objective optimization can also be used *directly* to solve the registration problem, i.e., to serve as the optimization component of an image registration method. The advantage of this is that by nature of the iterative search and refinement process underlying the EA, an approximation of the optimal Pareto front of solutions is found that represents relatively smooth transitions between solutions that are near each other in objective space. In our current early implementation, which is only available for 2D images,[10] the registration method consists of the EA as the optimizer, a dual-dynamic B-spline transformation model, and two objectives: the dissimilarity objective as expressed by the NCC and a deformation magnitude objective again expressed in terms of the second order derivatives of the transformation. The dual-dynamic transformation model handles large deformations well.[11] Here, we do not use a multi-resolution scheme, as it is not yet available in this implementation.

2.4 Test datasets

In this work, we registered breast MR images. We investigated three test cases of increasing problem difficulty. The easiest case (test case A) consists of registering two healthy breast MR images, both in prone orientation, but acquired at different time points. Furthermore, we tested two prone-supine cases. Aligning images of the breast in prone and supine orientation is a very challenging problem, due to the large deformation that the breast undergoes.

The first prone-supine case (test case B) is a pair of images acquired from a healthy volunteer, while the second one (test case C) is acquired from a patient suffering from breast cancer. In both cases there is a very large deformation, however the volunteer images have structures that are more easily identifiable. The patient images were acquired by use of a different scan protocol that was optimized for tumor visualization. In this case, the images are noisier, leading to an even more challenging deformable image registration problem.

Lastly, after per test case rigidly registering the MR scans on the bony anatomy, we chose a pair of 2D slices from each test case with which we tested both `elastix`, and the weight-free multi-objective image registration method, since the latter is currently available only for 2D image registration. The slices were chosen so that a good spatial correspondence via registration was obtainable, with out-of-plane motion as small as possible. An overview of all test cases and comparisons can be found in Table 1 and Figure 2.

2.5 Registration accuracy

The registration accuracy was quantified by the mean target registration error (TRE) of expert-defined landmarks. For this purpose, an experienced breast radiologist annotated eight pairs of landmarks in each of the 3D images. For the 2D images, fewer pairs (5) were used (see, e.g., Figure 3), since fewer structures were visible. The mean TRE was then calculated by averaging the Euclidean distances between the locations of the landmarks in the fixed image and the transformed locations of the landmarks in the deformed moving image. The annotation of the landmarks was highly accurate for the first two cases, due to the images being taken with the patient in the same positioning (test case A) and the presence of identifiable structures in both test cases A and B. Annotation was much more challenging for test case C due to the large difference in positioning in combination with the noisiness of the image.

3. RESULTS

3.1 Weight value and solution distribution in objective space

For the simple prone-prone test case, the outcomes of the random sampling of weights already form a Pareto front, indicative of the absence of clear local minima (Figure 4, test case A). The solutions found by the optimized linear combinations approach are therefore in this case similar. Further, we can see that for all cases, the values of the weights and their associated solutions are mostly logically correlated in objective space, i.e., the lower the value for the weight, the smaller the dissimilarity and the larger the deformation (Figure 4). Therefore, iterative adjustment of the weights with a trial-and-error approach in such a situation leads to expected changes in registration outcomes.

However, in truly hard cases, such as test case C, even if the weight value follows approximately the same pattern, one cannot be sure at all that an optimal solution is obtained by random sampling. Indeed, the majority of the weights with small values (which should lead to solutions in the most interesting part of the Pareto front, due to the large deformation needed to find a good match) do not lead to solutions forming a Pareto front (Figure 4, test case C). This solution distribution indicates that there are multiple local minima that are not overcome by the optimization algorithm used. This could partially be due to the inability of the transformation model to solve such a hard case, but is also due to the local-search nature of the stochastic gradient-descent optimizer used. For the solutions on the Pareto front found by the optimized linear combinations approach, there is a much smoother transition from larger weights to smaller weights (Figure 4) and furthermore, superior solutions. In this case, the assessment of the final results is much easier.

3.2 Error distribution in objective space

We observed before that a lower weight in principle leads to solutions with larger deformation and lower dissimilarity value. However, one should be careful, since solutions associated with very low weights are not the best ones in terms of mean TRE, as seen in the 2D patient prone-supine case (Figure 5, test case C, see tail of the Pareto front), due to implausible deformations taking place, likely as a result of overfitting. This pattern in error distribution can also be observed in the results found by the weight-free multi-objective registration method (most clearly visualized in Figure 5 test case A). The final desirable registration result will typically be where the registration error is low, which is a unique identifiable part of the Pareto front for the simple problems. Hence, for simple cases trial-and-error is clearly viable.

Regarding registration outcomes, the random sampling of weights approach with `elastix` obtained good results for 3D test case A (minimum mean TRE <1.0 mm), and acceptable results for 3D test case B (minimum mean TRE 4.2 mm). These results are also found by the optimized linear combinations approach. The results for the 3D test case C are worse; the optimized linear combination approach finds overall better solutions, including the best ones for this problem, with a minimum mean TRE of 9.1 mm. The random sampling approach for test case C resulted in large variability of the mean TRE (which fluctuated as much as 5.0 mm), making a trial-and-error approach here inefficient (Figure 5, test case C, see zoomed part). Among the 2D images, the direct multi-objective image registration method performed best for test case B (minimum mean TRE 2.0 mm), even without a multi-resolution strategy, possibly due to the use of the dual-dynamic transformation model; the difference in mean TRE between its solutions and the ones found by the optimized linear combinations approach is in the order of 0.7-0.9 mm. For case C in 2D, the optimized linear combinations approach as well as the weight-free method obtained solutions of comparable quality (mean TRE: ~7.1 mm), however the optimized linear combinations approach, aided by the multi-resolution strategy, obtained overall better solutions.

Overall, weight tuning via the optimized linear combinations approach, along with the weight-free multi-objective registration method, yields the smoothest error distributions, thus facilitating the selection of a registration outcome.
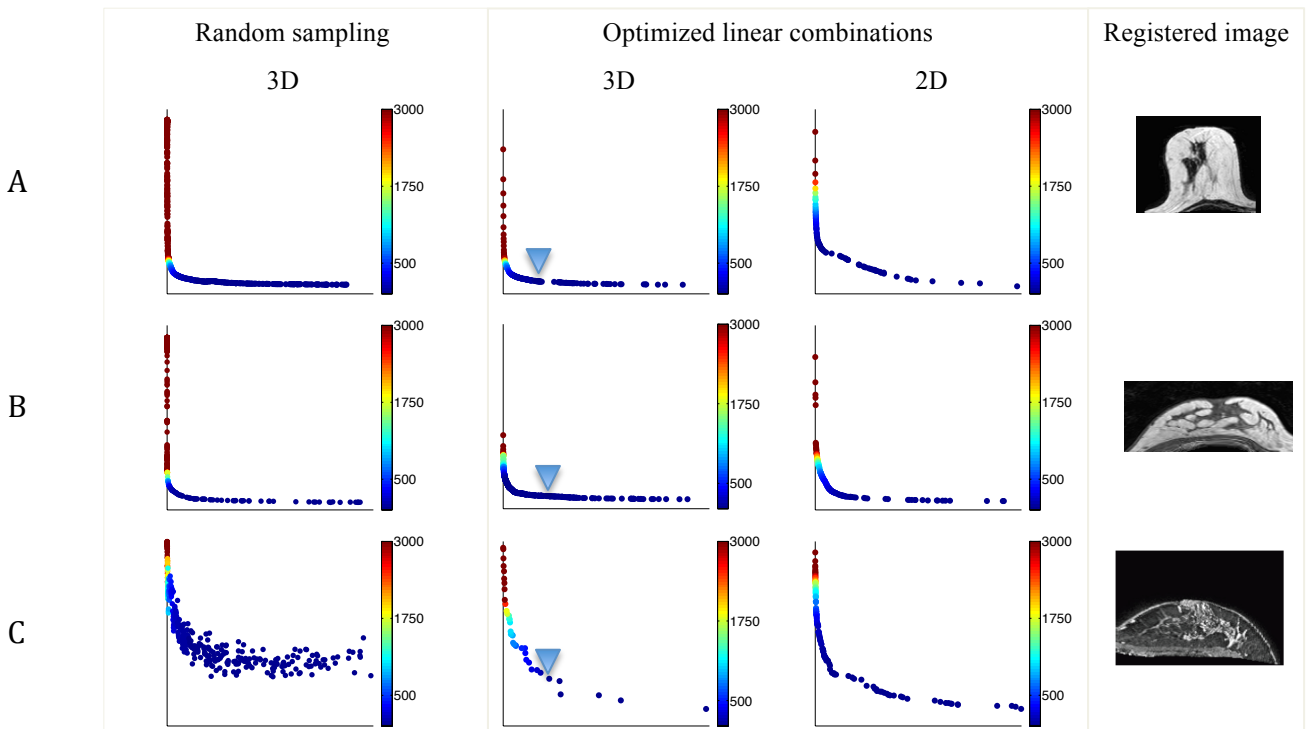
Figure 4: Color-coded distribution of the value of weight λ in objective space. Location of selected registration outcomes are indicated by blue arrow. Horizontal axis, deformation magnitude. Vertical axis, dissimilarity. Axis scale is the same per test case. Note: scales omitted because actual values are irrelevant.
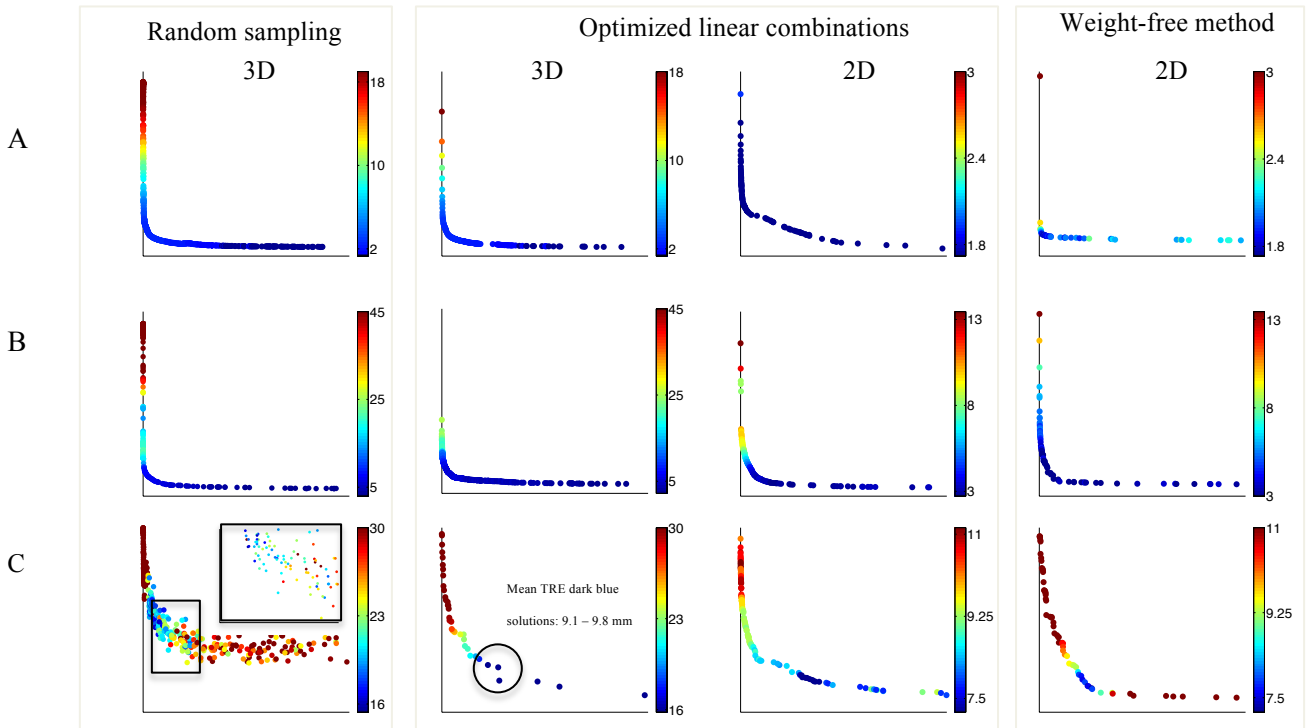

Figure 5: Color-coded distribution of the mean TRE (mm) in objective space. Horizontal axis, deformation magnitude. Vertical axis, dissimilarity. Axis scale is the same per test case. Note: scales omitted because actual values are irrelevant.

4. DISCUSSION AND CONCLUSIONS

In this work, we explored the interplay between weights and objectives in deformable image registration in order to determine the extent to which it can affect the success of trial-and-error approaches for weight tuning, which are mainly employed in deformable image registration. We used multi-objective optimization to investigate the nature of the trade-off between objectives of interest, and how these are related to the choice of weights, on three problems of increasing difficulty. Further, we associated an indication of clinical relevance to the found solutions, by calculating the landmark-based mean target registration error. Results indicate that an iterative trial-and-error approach, albeit inherently less powerful than a purely multi-objective approach due to limitations related to the problem landscape, is viable for deformable image registration problems that require only small deformations, and for which a logical relation between solution quality and weight values exists. For harder cases, where changes in weights lead to unexpected and highly variable registration outcomes, trial-and-error approaches are not likely to succeed. However, we found that explicitly optimizing the weights in a multi-objective fashion or solving the registration problem with a direct multi-objective approach does facilitate choosing a final registration outcome in an insightful manner because multiple solutions that represent efficient trade-offs between the objectives of interest can be inspected at once. Moreover, our results indicate that by using multi-objective optimization better solutions can be found than when many weight combinations are tried (i.e., random sampling).

For test case C some of the best solutions in terms of mean TRE for the random sampling approach were not located on the Pareto front. This signals the existence of a somewhat hidden issue, namely that the objectives used, when optimized, do not necessarily provide registration outcomes that are ultimately the most desirable (i.e., have a mean TRE of 0.0 mm as well as a good balance between all objectives of interest). This is however an inherent problem in deformable image registration and not one specific to our multi-objective approach.

The computational cost of the optimized linear combinations approach can be quite high, as each function evaluation can take several minutes, depending on the parameters of the registration such as number of resolution levels or maximum number of iterations given to the single-objective optimizer. On the other hand, the multi-objective EA that optimizes the weights in this setup, is well parallelizable. Given the long function evaluation time, the overhead of the EA itself is negligible, making the potential for parallel speedup excellent.

Overall, we find that with a multi-objective optimization approach a logical selection of a final registration outcome is facilitated, possibly allowing solving hard problems such as prone-supine registration with existing single-objective deformable image registration tools accurately, reliably, and in a repeatable manner. This needs to be however further validated with a larger number of registration cases.

REFERENCES

[1] Tang, L., Hamarneh, G. and Abugharbieh, R., "Reliability-driven, spatially-adaptive regularization for deformable registration," Proc. WBIR, LNCS 6204, 173-185 (2010).

[2] Yeo, B. T., Sabuncu, M. R., Vercauteren, T., Holt, D. J., Amunts, K., Zilles, K., Golland, P. and Fischl, B., "Learning task-optimal registration cost functions for localizing cytoarchitecture and function in the cerebral cortex," IEEE Trans Med Imaging 29(7), 1424-1441 (2010).

[3] Sotiras, A., Davatzikos, C. and Paragios, N., "Deformable medical image registration: a survey," IEEE Trans Med Imaging 32(7), 1153–1190 (2013).

[4] Das, I. and Dennis, J. E., "A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems," Struct Optimization 14(1), 63-69 (1997).

[5] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O. and Hawkes, D. J., "Nonrigid registration using free-form deformations: application to breast MR images," IEEE Trans Med Imaging 18(8), 712–721 (1999).

[6]   Klein, S., Staring, M., Murphy, K., Viergever, M. A. and Pluim, J. P. W., "elastix: a toolbox for intensity-based medical image registration," IEEE Trans Med Imaging 29(1), 196–205 (2010).

[7]   Klein, S., Pluim, J. P. W., Staring, M. and Viergever, M. A., "Adaptive stochastic gradient descent optimisation for image registration," Int J Comput Vis 81(3), 227–239 (2008).

[8]   Lozano, J. A., Larrañaga, P., Inza, I. and Bengoetxea, E., [Towards a new evolutionary computation. Advances in estimation of distribution algorithms], Springer-Verlag, Berlin (2006).

[9]   Bosman, P. A. N. and Alderliesten, T., "Incremental gaussian model-building in multi-objective EDAs with an application to deformable image registration," Proc. GECCO, 241-248, doi:10.1145/2330163.2330199 (2012).

[10] Pirpinia, K., Bosman, P. A. N., Sonke, J.-J., van Herk, M. and Alderliesten, T., "On the usefulness of gradient information in multi-objective deformable image registration using a B-spline-based dual-dynamic transformation model: comparison of three optimization algorithms," Proc. SPIE 9413, 941339, doi: 10.1117/12.2082172 (2015).

[11] Alderliesten, T., Sonke, J.-J. and Bosman, P. A. N., "Deformable image registration by multi-objective optimization using a dual-dynamic transformation model to account for large anatomical differences," Proc. SPIE 8669, 866910, doi:10.1117/12.2006783 (2013).