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

Peter A. N. Bosman*[a] and Tanja Alderliesten[b]

[a] Centrum Wiskunde & Informatica (CWI), P.O. Box 94079, 1090 GB Amsterdam, The Netherlands;
[b] Dept. of Radiation Oncology, Academic Medical Center (AMC), P.O. Box 22660, 1100 DD Amsterdam, The Netherlands

ABSTRACT

We recently demonstrated the strong potential of using dual-dynamic transformation models when tackling deformable image registration problems involving large anatomical differences. Dual-dynamic transformation models employ two moving grids instead of the common single moving grid for the target image (and single fixed grid for the source image). We previously employed powerful optimization algorithms to make use of the additional flexibility offered by a dual-dynamic transformation model with good results, directly obtaining insight into the trade-off between important registration objectives as a result of taking a multi-objective approach to optimization. However, optimization has so far been initialized using two regular grids, which still leaves a great potential of dual-dynamic transformation models untapped: a-priori grid alignment with image structures/areas that are expected to deform more. This allows (far) less grid points to be used, compared to using a sufficiently refined regular grid, leading to (far) more efficient optimization, or, equivalently, more accurate results using the same number of grid points. We study the implications of exploiting this potential by experimenting with two new smart grid initialization procedures: one manual expert-based and one automated image-feature-based. We consider a CT test case with large differences in bladder volume with and without a multi-resolution scheme and find a substantial benefit of using smart grid initialization.

Keywords: Deformable image registration, grid initialization, multi-objective optimization, evolutionary algorithms, transformation model, large anatomical differences

1. INTRODUCTION

Deformable image registration typically becomes hard when large deformations are involved because they often impose many local optima for the underlying optimization method. Moreover, transformation models need to be sufficiently flexible. Many registration methods use *a single fixed* grid for the source image and *a single moving* grid for the target image.[1] One advantage of using fixed grids is that they enable many interpolation methods to define transformations, including, e.g., smooth b-splines. A disadvantage however is that a fine-grained grid may be required, e.g., when small areas in the source image need to be (substantially) enlarged. This increases the number of parameters (i.e., coordinates of grid points) that must be optimized over, decreasing efficiency. Dual-dynamic transformation models have the potential to counter this effect by using *two moving* grids: one for the source and one for the target image.[2] A mapping is then established either indirectly (using a virtual regular grid that allows again a variety of smooth transformations) or directly. The latter is possible when e.g., a simplex model is used; in 2D this means that the moving grids are triangulations and that triangles directly correspond to one another. Through optimization, these grids can be aligned more favorably with structures that need to be deformed, eventually potentially requiring less parameters to optimize over. Furthermore, both disappearing and appearing structures can be accommodated because two moving grids are used.

Fine-tuning registration methods to specific clinical applications is of great importance since one unique solution to all deformable image registration problems does not exist. One major reason for this is that often *single-objective* optimization is employed by optimizing a linearly weighted combination of all objectives that are of interest. Setting the weights anew for each registration task is non-trivial and often not very insightful, especially for complex registration tasks. Therefore, often this task is performed by trial and error.

* E-mail: Peter.Bosman@cwi.nl; Telephone: +31 (0)20 592 4265; Website: http://homepages.cwi.nl/~bosman/

Moreover, including another objective, e.g., exploiting additional guidance information such as contours or matching point pairs annotated by experts, only further complicates this task. By investigating deformable image registration from a *multi-objective* optimization perspective, the need to set a predetermined singular combination of objectives is removed.[3-5] In multi-objective optimization, objectives are kept separate and are never combined through a weighting. The solution to a multi-objective optimization problem is a *collection* of registration outcomes (a so-called Pareto front) that represents efficient trade-offs between all objectives.[6] Efficiency here means that no other registration outcomes have been found that are better in all objectives than any of the registration outcomes in the Pareto front. A multi-objective approach provides the potential to provide insight into the true possible outcomes of registration and what their trade-offs are (e.g., the extent of deformation vs. match quality). After studying the possible outcomes on the Pareto front, the preferred outcome for the application at hand can be selected in a transparent manner.

Multi-objective optimization however also requires more computation time. It is therefore even more important in this case to ensure that the grid complexity remains as small as possible and is used as efficiently as possible. It may take considerable (global) optimization effort to make the best use of the flexibility offered by dual-dynamic transformation models. Moreover, the topology of the grid is also a determining factor in the extent to which large deformations *can* be efficiently modeled. Allowing a different topology than a regular grid may result in far more suitable grid alignments for specific registration problems using the same, or less, grid points. Changing the topology is possible with a simplex transformation model. In this paper, we will therefore study the implications in terms of optimization efficiency of initially aligning the grid points and topologies of a simplex transformation model better with the structures that are likely to deform rather than using a regular, uniformly-spaced grid as a starting point.

We recently demonstrated the strong potential of the use of a dual-dynamic transformation model within a multi-objective optimization framework to tackle hard registration problems involving large deformations and (dis)appearing structures.[2,5,7] Using a state-of-the-art, powerful model-based evolutionary optimization algorithm, excellent registration results were obtained using regularly-initialized grids. However, in order to obtain these results, the algorithm must be run sufficiently long and with a sufficiently fine-grained regular grid. In this work, we study new, smart grid initialization procedures to obtain similar or better results much more efficiently using less grid points.

2. MATERIALS AND METHODS

Image registration can be naturally posed as a multi-objective optimization problem, since there are different objectives of interest. The advantage of treating image registration as a multi-objective optimization problem is that direct insight may be obtained into the nature of the interplay between the objectives of interest. For this reason, we developed a multi-objective deformable image registration framework.[2-5,7] In the following, we provide several details on the current state of this framework and the choices made for the work presented in this paper.

2.1 Objectives

In this work we considered three objectives:

- *Match quality.* We use the well-known cross-correlation similarity metric[8] to compute a notion of quality of a match between a deformed source image and its target. Since all objectives in our implementation are to be minimized, we actually use the negation of this value. Values for this objective lie in the range [-1;1].

- *Amount of deformation.* We employ Hooke's law[9] on the basis of the deformation vector field (*dvf*) to quantify the amount of deformation associated with a transformation. For every pixel, we use the average length of the difference vector that is obtained by comparing the deformation of that pixel with each of its neighboring pixels. Let the set of considered pixels be denoted by P, the set of neighbors of pixel p be denoted by $n(p)$, and let the deformation vector in the *dvf* associated with pixel p be denoted by dvf_p. This objective can then be defined as:

$$F_{deform} = \sum_{p \in P} \frac{1}{|n(p)|} \left[\sum_{np \in n(p)} \left\| dvf_p - dvf_{np} \right\| \right]^2$$

- *Guidance error.* The third objective is related to guidance information. Guidance information is stored in a set of tuples. Each tuple consists of two sets, each containing one or more points associated with the source and the target image, respectively. The guidance objective captures landmark/contour correspondence in terms of distances between pairs of guidance-point sets. For each pair of guidance-point sets, the points defined in the source image are transformed to the target image and vice versa. For each point, after transformation, the

smallest distance is computed to the points that define the contour in the other image. The distance between two guidance-point sets is then computed in a symmetric fashion as the sum of these distances over all points in the guidance sets, normalized by dividing by the total number of points. The sum over all pairs of guidance-point sets is then to be minimized. In other words, let G be the set of tuples (Gs,Gt) of guidance-point sets, let T be the transformation function as defined by the solution and let d denote Euclidean distance. Then, the guidance objective can be formulated as follows:

$$
F_{guidance} = \sum_{(G_s, G_t) \in G} \frac{1}{|G_s| + |G_t|} \left[\sum_{g_s \in G_s} \min_{g_t \in G_t} \{ d(T(g_s), g_t) \} + \sum_{g_t \in G_t} \min_{g_s \in G_s} \{ d(T^{-1}(g_t), g_s) \} \right]
$$

2.2 Multi-objective optimization algorithm

To perform multi-objective optimization, we use a model-based evolutionary algorithm (EA) that exploits features of a problem's structure via probabilistic modeling.[10-13] The particular EA that we employ is iMAMaLGaM (incremental Multi-objective Adapted Maximum Likelihood Gaussian Model mixture) in which the underlying probabilistic model is a Gaussian mixture distribution.[4] In related work, iMAMaLGaM was shown to have excellent performance, converging to high-quality approximations of the optimal Pareto front on well-known benchmark problems.[4]

Following results from previous work, we include adaptive steering in the algorithm to zoom in on only a part of the Pareto front because transformations with very little deformation are typically not interesting.[5] Adaptive steering shrinks the space of feasible solutions during optimization adaptively, based on the best values found for individual objectives so far. Specifically, all solutions for which one or more objective values fall within a certain range are deemed to be infeasible. For the similarity objective we use $[s^{best}; 0.99 \times s^{best}]$ and for the guidance objective we use $[g^{best}; 1.5 \times g^{best}]$. The amount of deformation objective is not further constrained. These ranges were previously found to give good results.[5] To help the algorithm find and maintain solutions in the feasible space, during the first half of an optimization run, the feasible range is shrunk from a very large range to the aforementioned value. During the second half of an optimization run, the ranges as indicated above are used.

2.3 Dual-dynamic transformation model

We compute deformations by considering corresponding pairs of triangles (i.e., in the source and target grids) and by employing linear simplex interpolation. A transformation is defined by associating coordinates with each point in *two* grids. Therefore, the number of real-valued parameters to be optimized equals two times the number of grid points times the spatial dimensionality of the image (e.g., 2D or 3D). Unlike when using the traditional b-spline based transformation model, the use of such grids allows us to investigate the benefit of different initialization procedures because an irregular set of points can be used. We study and compare the following three different grid initialization procedures.

- *Regular grid initialization.* This case considers a regular grid of $n \times n$ points in which diagonal edges are also included in order to ensure that the grid is a triangulation.

- *Manual smart grid initialization.* A triangulation can be built from any set of points by iteratively adding edges that do not intersect previously added edges. We consider all possible edges between all pairs of points in order of ascending edge-length. Points can be used to delineate structures that are expected to play a crucial role in deformation. For the purpose of this work, an expert manually annotated grid points in the source image.

- *Automatic smart grid initialization.* Although accurate and likely useful, manually annotating grid points can be tedious and laborious. We therefore further investigated an automatic approach to smart grid initialization. To this end, we first apply edge detection (Sobel filter[14]) to the source image after which we use a greedy subset selection approach to select grid points that are as well-spread as possible in the image space, but are only located on the detected edges.

2.4 Experiments

We use two CT scans acquired from the same patient with an empty and a full bladder. A pair of 2D slices is selected after rigid registration on the bony anatomy. Guidance information was annotated by a clinical expert. To solve this registration problem, large deformations are required (see Fig.2). There are furthermore small structures that need to be enlarged (e.g., between the bone and the bladder).

To study the impact of smart initialization, we consider two different experiment configurations:

1. In this configuration, grid resolutions of 25 and 81 points, resulting in 100 and 324 parameters to be optimized respectively, are used to initialize the grids either regularly (i.e., 5×5 and 9×9 grids), or using smart initialization. For smart initialization, 4 points are automatically used to ensure that the entire image content is covered. We let the EA run for a longer period of time and for a shorter period of time. The length of the long run is based on the time that is expected to be required to obtain high-quality results. Specifically, the length of the long run is based on a previously published scalability analysis[15] and allows $20\times93.6\times\ell^{1.81}$ evaluations where ℓ is the number of parameters to optimize (7,803,795 evaluations for 25 grid points; 65,522,780 evaluations for 81 grid points). The short run is allowed only 1% of these evaluation budgets, implying a 100-fold speedup. These runs are done for a single grid scale to study the impact of smart initialization at individual grid scales. To still provide the benefit of a multi-scale scheme[16,17] on the image-resolution side, during the first half of the run, when adaptive steering is used, the images are iteratively refined, starting from 3×3 pixels. During the second half of the run, the image is used at full resolution.

2. This configuration is targeted more at final practical use. The best final registration outcomes were previously obtained with a multi-scale approach where results from the 5×5 grid level (and a lower image resolution) are used to initialize a run at the 9×9 grid level.[7] We ran the EA using that scheme for the purpose of studying final outcomes, using the previously reported number of evaluations: $10,000\times\ell$ (1,000,000 evaluations for 25 grid points; 3,240,000 evaluations for 81 grid points). To further study the impact of smart initialization also in this configuration, we also ran the multi-scale approach with a budget of only 10% of the evaluations in the last stage. Since most of the runtime is used up in the 9×9 stage, this still implies a speedup factor of about 10.

3. RESULTS

Results of experiments ran according to configuration 1 are shown in Figure 1. What can clearly be observed is that inherently different results between regular and smart grid initialization are obtained in the case of grids with 25 points. With regular grid initialization, the matches were not nearly as good as with smart grid initialization. Grids with 81 points however allow far more detailed transformations, even if the grid is regular. This causes almost no difference to be observable between the results for the long run and grids with 81 points. It should be noted that this is furthermore indicative of the optimization capabilities of the EA, since a *regularly* initialized grid is initially not very well aligned with the parts of the image that need to be deformed, regardless of the grid resolution. This misalignment can also be seen from the short run in Figure 1: results obtained with the regularly initialized grid are nowhere near as good as results obtained with smart initialized grids. The EA needs much longer to obtain good results because the grid points need to be moved much more than when the grid points would have been (smartly) initialized to already (roughly) outline structures that need to be deformed. The EA does however ultimately achieve roughly the same type of deformations with a regular grid as with a smartly initialized grid. This can also be seen from the final registration results, shown in Figure 2, that were obtained with the multi-scale approach (i.e., experiment configuration 2). From the Pareto front of registration outcomes that resulted, a single solution was picked. Contrary to the meticulous parameter tuning that is required in common single-objective approaches to deformable image registration, picking the final registration outcome that makes the most sense to the user is fairly straightforward.[2] It is also an insightful process because all registration outcomes can be readily compared by traversing the Pareto front. By doing so, a transition of outcomes can be observed, going from outcomes that represent too little deformation to outcomes that represent too much deformation, making it immediately clear where the preferred registration outcome lies on the Pareto front. After 1,000,000 evaluations with a 5×5 regularly initialized grid, followed by 3,240,000 evaluations with a refined 9×9 grid, the EA has relocated the grid points so that the desired transformation can be achieved. The final results are even similar to when smart initialization is used, with virtually similar results in Figure 2 and differences in negative correlation coefficient that are only minor, similar to the long-run results in Figure 1, with smart initialization being slightly better.
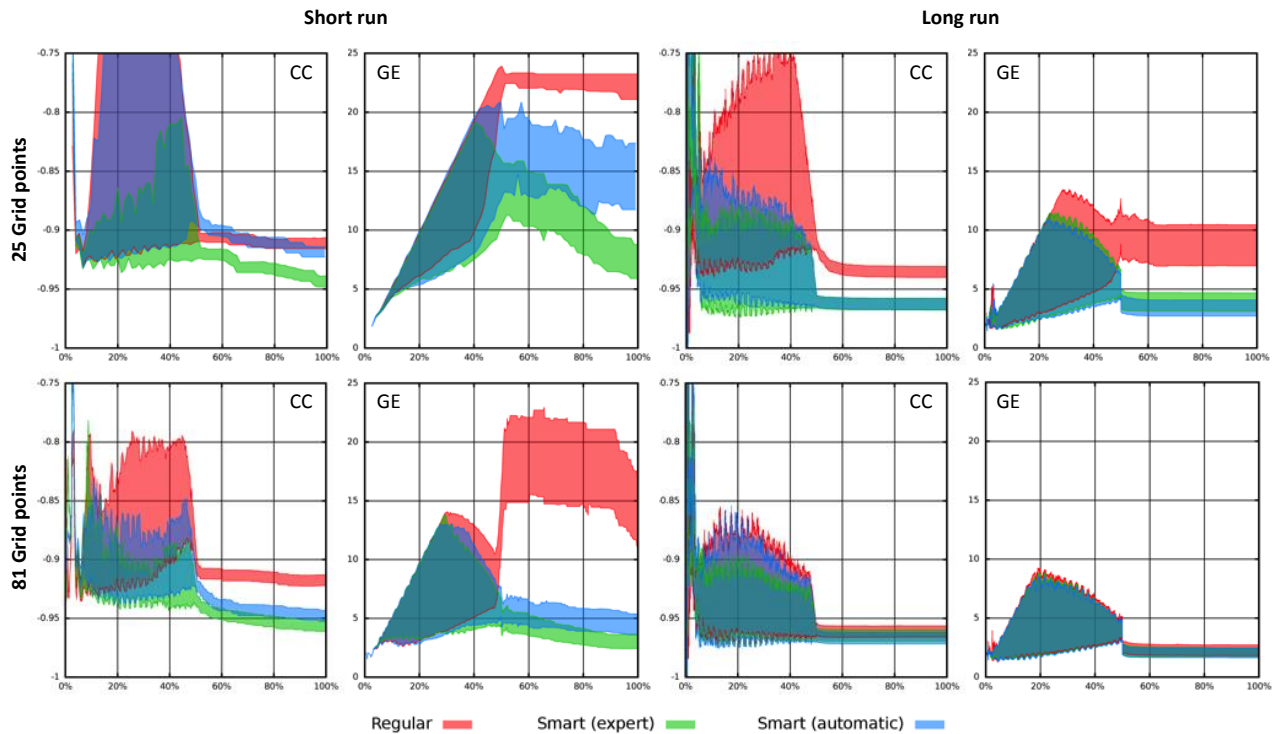
Figure 1. Results obtained using experiment configuration 1. *Top to bottom*: 25 to 81 grid points. *Left to right*: cross correlation (CC) and guidance error (GE, in mm) for the short and the long runs, respectively. Horizontal axes: percentage of run time. Shown are ranges of objective values along the so-far obtained Pareto front during the run for the three different grid initialization strategies.

The advantages of smart grid initialization become truly clear when considering the outcomes obtained for the short runs: much better results could be obtained with smart grid initialization than with regular grid initialization, both for the case of 25 and 81 grid points. Moreover, with smart grid initialization good registration outcomes were already observed using only 25 grid points, which was impossible for regular grid initialization. Finally, especially in the case of short runs, *expert* smart initialization gave the best results. However, results obtained with automatic smart initialization were not far behind, indicating that most of the benefit of smart grid initialization may well be obtainable automatically.

These results indicate that using smart grid initialization can be greatly beneficial for studying deformable image registration from a multi-objective optimization perspective using transformation models that do not require a regular basis, such as simplex-based models. Such models are typically used in physical modeling (e.g., finite element modeling), making them good candidates for considering more detailed models within registration, e.g., by incorporating different tissue characteristics, so as to ensure physically accurate results. Moreover, combined with our EA, large deformations can be found directly, using a single registration method, which is promising for future extensions to 3D image registration. The need for an efficient use of grid points is then even greater, as moving from 5×5×5 grids to 9×9×9 grids then means going from 375 parameters to 2187 parameters to optimize over. Although our near-future planned work targets specific efficiency enhancements targeted specifically at improving the required computation time, covering a Pareto front of high-quality solutions in multi-objective optimization remains a computationally demanding task for which keeping the number of parameters to optimize over low is important, making the results in this paper even more relevant.
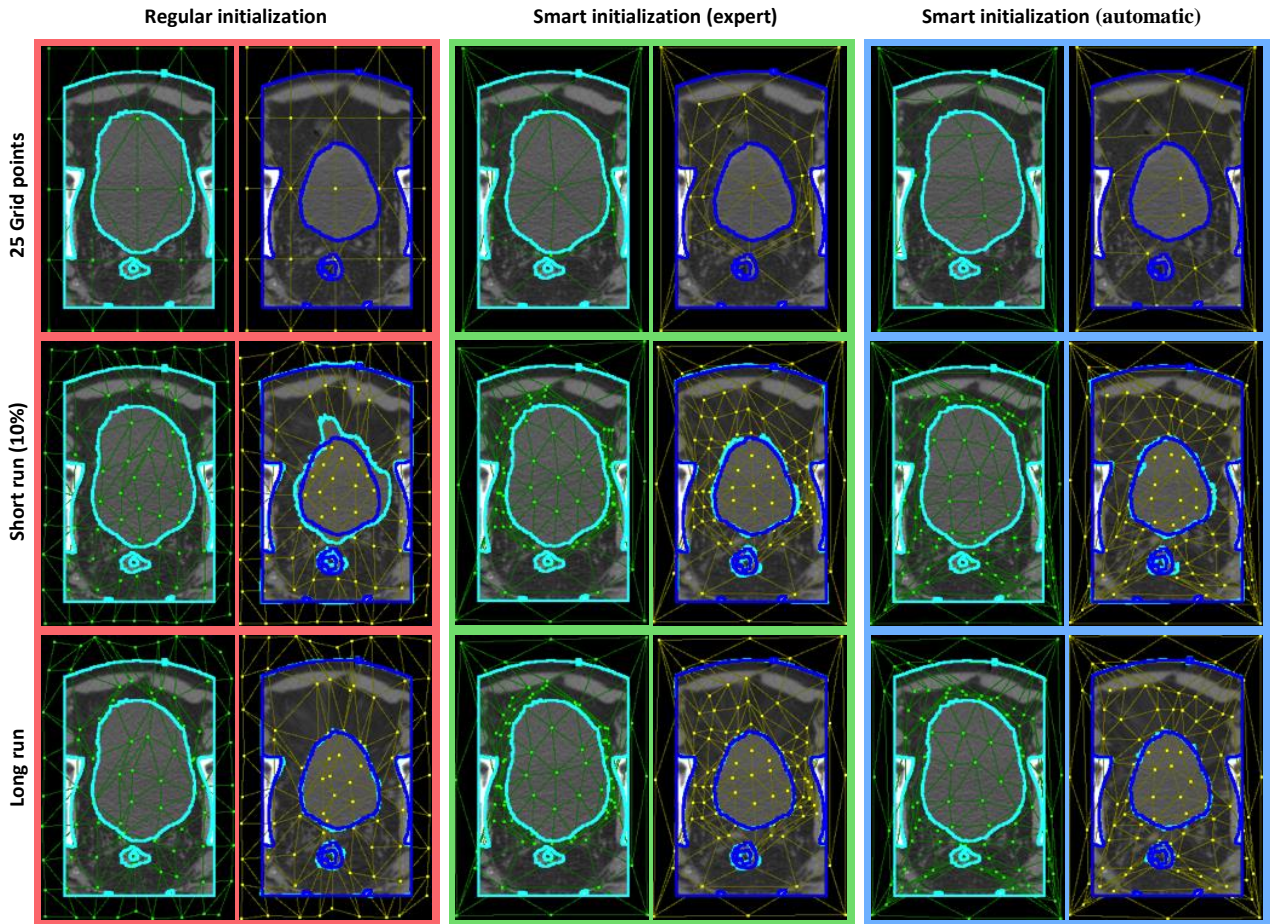
Figure 2. Results obtained using experiment configuration 2. *Top*: full and empty bladder CT images overlaid with triangulations using the three grid initialization strategies for 25 grid points. *Middle and bottom*: results after running a multi-scale approach short and long, respectively. Guidance information is illustrated in turquoise and blue. A single registration result, chosen from the final Pareto front obtained by the EA, is presented.

4. CONCLUSIONS

We studied, within a multi-objective framework that eliminates the need to a-priori define and tune weighting factors, the impact of using smartly initialized simplex grids (using either expert knowledge or image-processing techniques), rather than common regularly initialized grids, on optimization efficiency. This is a novel contribution since this has not been studied before, especially in combination with a dual-dynamic transformation model and our multi-objective optimization approach.

Results illustrate that similar or better results can be obtained much more efficiently with smartly initialized grids than with regularly initialized grids, within as little as 1%-10% of the total runtime. Moreover, inherently better results can be obtained with smart initialization using less grid points, making the road to studying registration with physically-enhanced 3D transformation models from a multi-objective perspective a much more computationally friendly one.

REFERENCES

[1] Crum, W. R., Hartkens, T. and Hill, D. L. "Non-rigid image registration: theory and practice," Br J Radiol 77(Spec No 2), S140-S153 (2004).

[2] 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).

[3] Alderliesten, T., Sonke, J.-J. and Bosman, P. A. N., "Multi-objective optimization for deformable image registration: proof of concept," Proc. SPIE 8314, 831420, doi:10.1117/12.911268 (2012).

[4] 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).

[5] Alderliesten, T., Bosman, P. A. N. and Bel, A., "Getting the most out of additional guidance information in deformable image registration by leveraging multi-objective optimization," Proc. SPIE 9413, 94131R, doi: 10.1117/12.2081438 (2015).

[6] Deb, K., [Multi-objective optimization using evolutionary algorithms], John Wiley & Sons Inc., New York (2001).

[7] Alderliesten, T., Bosman, P. A. N., Sonke, J.-J. and Bel, A.," A multi-resolution strategy for a multi-objective deformable image registration framework that accommodates large anatomical differences, " Proc. SPIE 9034, 90343G, doi:10.1117/12.2042856 (2014).

[8] Roche, A., Malandain, G., Pennec, X. and Ayache, N., "The correlation ratio as a new similarity measure for multimodal image registration," Proc. MICCAI, LNCS 1496, 1115-1124 (1998).

[9] Arfken, G., [Mathematical methods for physicists], Academic Press, Inc., San Diego (1985).

[10] 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).

[11] Pelikan, M., Sastry, K. and Cantú-Paz, E., [Scalable optimization via probabilistic modeling: from algorithms to applications], Springer-Verlag, Berlin (2006).

[12] Bosman, P. A. N. and Grahl, J., "Matching inductive search bias and problem structure in continuous estimation-of-distribution algorithms," EJOR 185(3), 1246-1264 (2008).

[13] Mühlenbein, H. and Höns, R., "The estimation of distributions and the minimum relative entropy principle," Evolutionary Computation 13(1), 1-27 (2005).

[14] Gonzalez, R. C. and Woods, R. E., [Digital Image Processing], Prentice Hall, Upper Saddle River, New Jersey (2002).

[15] Bosman, P. A. N., Grahl, J. and Thierens, D., "Benchmarking parameter-free AMaLGaM on functions with and without noise," Evolutionary Computation 21(3), 445-469 (2013).

[16] Bajcsy, R. and Kovačič, S., "Multiresolution elastic matching," Comput Gr Image Process 46(1), 1-21 (1989).

[17] Lester, H. and Arridge, S. R., "A survey of hierarchical non-linear medical image registration," Pattern Recognition 32, 129-149 (1999).