FLOOR: Fusing Locally Optimal Registrations

Dong Hye Ye[1], Jihun Hamm[2], Benoit Desjardins[1], and Kilian M. Pohl[1]

[1] Department of Radiology, University of Pennsylvania, Philadelphia, PA, 19104
[2] Department of Computer Science, Ohio State University, Columbus, OH, 43210
Dong.Ye@uphs.upenn.edu

Abstract. Most registration algorithms, such as Demons [1], align two scans by iteratively finding the deformation minimizing the image dissimilarity at each location and smoothing this minimum across the image domain. These methods generally get stuck in local minima, are negatively impacted by missing correspondences between the images, and require careful tuning of the smoothing parameters to achieve optimal results. In this paper, we propose to improve on those issues by choosing the minimum from a set of candidates. Our method generates such candidates by running the registration algorithm multiple times varying the setting of the smoothing and the image domain. We iteratively refine those candidates by fusing them with the outcome of alternative approaches and locally adapting the smoothing parameters. We implement our algorithm based on Demons [1] and find alternative minima via manifold learning [2]. Compared to those two methods, our 600 pairwise registrations of cardiac MRIs significantly better handle the large shape variations of the heart and the different field of views captured by scans.

Keywords: Image Registration, Manifold Learning, Cardiac MRI.

1 Introduction

General-purpose registrations, such as [1], model deformable image alignment as a problem of finding the (local) minimum of an energy function. The accuracy of the search is sensitive to the setting of the optimization method, which commonly assumes one-to-one correspondence between moving and target scan. Scans, such as cardiac MRI (see Fig. 1), often vary with respect to the anatomy captured in their field of view as well as the shape of the anatomy, such as the heart. It is thus difficult to find the optimal setting for methods off-line. In this paper, we address this shortcoming of existing tools by refining their solution to the minimization problem through iteratively fusing the results of multiple searches.

Applications relying on non-rigid registration, such as atlas-based segmentation, contain various coping mechanisms for these issues. The most common technique is to crop the scan to a specific region, such as in skull stripping, before refining registration. Accurate cropping is only reliable for certain areas, such as brain MRI, but even then generally requires manual editing. Alternatively, researchers register scans multiple times or from multiple subjects and then fuse the aligned label maps [3].
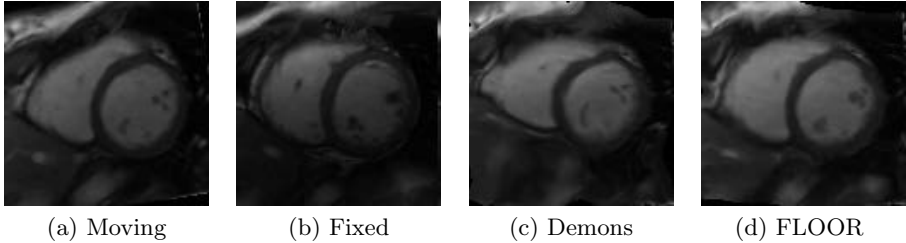
(a) Moving   (b) Fixed   (c) Demons   (d) FLOOR

Fig. 1. (c) shows the result of Demon aligning (a) to (b). The artifacts shown in (c) are caused by missing correspondences. FLOOR's output (d) much better matches (b).

These mechanisms are less beneficial when directly analyzing deformations, such as in [4]. Here, the optimal registration setting is often determined via cross validation [5] hoping the results during training translate to new scans. This assumption generally does not hold for heterogeneous populations. Alternatively, graphical models find the global instead of a local minimum [6]. The accuracy of these methods suffers from the discretization of the deformation map. Finally, manifold learning [2,7] improves registration by decomposing the problems into a set of simpler registration tasks. The accuracy of these methods depends on the size of the training data as well as the accuracy of the solutions to those simpler tasks. To address the first issue, manifolds are created for image parcels instead of the entire domain[8,9]. Our proposed method, called Fusing Locally Optimal Registrations (FLOOR), is a first attempt to advance these learning techniques with respect to the second issue by iterating between improving the tasks' solutions and refining the corresponding manifolds. To the best of our knowledge, FLOOR is also the first method to improve deformation maps (and not aligned segmentations as in [3]) by fusing the outcome of several registrations.

FLOOR improves the accuracy of a registration by iteratively reducing the corresponding minimum based on a set of local optima, called candidates. At each iteration, FLOOR generates candidates by applying the registration with different smoothing settings to various parcels of the image and including the outcome of alternative methods, such as manifold learning. It improves the minimum by fusing the candidates, i.e. selecting the deformation at each voxel that minimizes the energy function. Note, all candidates are local minima of the energy function but differ in the degree of flexibility, the parcel they are defined on, and the way they were generated. The approach converges to a smooth deformation map after the selection process does not improve candidates.

FLOOR addresses many of the previous issues as it draws from multiple searches, contains the influence of missing correspondences to parcels instead of the entire domain, and adapts the stiffness locally and online. If we include Manifold Learning based Deformation maps (MLD)[2] in the candidate set, the corresponding manifolds are also likely to be improved at each iteration as they are trained on refined pairwise registrations. Our implementation of FLOOR demonstrates this by fusing the results of Demons[1] with MLD. We register 600 pairs of cardiac MRIs. FLOOR not only significantly improves the result of Demons, such as in Fig. 1, but also of MLD trained on the entire domain.

2 Fusing Deformation Maps

Non-linear registration approaches generally define the best mapping between two images F and M as the spatial transformation $\boldsymbol{s}$ that minimizes $\text{Sim}(F, M \circ \boldsymbol{s})$, the dissimilarity between the aligned images, while being smooth according to the regularizer $\text{Reg}(\boldsymbol{s})$. In other words, $\boldsymbol{s}$ minimizes the energy function

$$E(\boldsymbol{s}) \triangleq \text{Sim}(F, M \circ \boldsymbol{s}) + \frac{1}{\sigma^2}\text{Reg}(\boldsymbol{s}), \tag{1}$$

where σ controls the flexibility of the deformation and $E(\cdot)$ is defined with respect to the image domain Ω. We now briefly review the Demons algorithm [1], which finds a local minimum for a specific instance of this equation. We then derive FLOOR, which iteratively refines those results by spatially varying σ, confining $E(\cdot)$ to parcels and fusing them with alternative deformation maps.

Demons: The method defines both terms of Eq. 1 with respect to L2-norm $||\cdot||$, where $\text{Sim}(F, M \circ \boldsymbol{s}) \triangleq ||(F - M \circ \boldsymbol{s})||^2$ and $\text{Reg}(\boldsymbol{s}) \triangleq ||\nabla \boldsymbol{s}||^2$. The algorithm determines the local minimum $\boldsymbol{s}$ of $E(\cdot)$ by iteratively minimizing for $\boldsymbol{s}$ and an approximation of $\boldsymbol{s}$, called $\boldsymbol{c}$, according to

$$E(\boldsymbol{s}, \boldsymbol{c}) \triangleq \sum_{x \in \Omega} \left(||F(x) - M(x + \boldsymbol{c}(x))||^2 + ||\boldsymbol{c}(x) - \boldsymbol{s}(x)||^2 + \frac{1}{\sigma^2}||\nabla_x \boldsymbol{s}||^2 \right). \tag{2}$$

$\nabla_x \boldsymbol{s}$ is the gradient of $\boldsymbol{s}$ at x. Given $\boldsymbol{s}$, it first updates $\boldsymbol{c}$ at each voxel $x \in \Omega$:

$$\boldsymbol{c}(x) \leftarrow \arg\min_{\boldsymbol{t}(x) \in \mathbb{R}^3} ||F(x) - M(x + \boldsymbol{t}(x))||^2 + ||\boldsymbol{t}(x) - \boldsymbol{s}(x)||^2, \tag{3}$$

which has a closed form solution. Keeping $\boldsymbol{c}$ fixed, the methods updates $\boldsymbol{s}$ by applying $\boldsymbol{c}$ to a Gaussian kernel $K(\sigma)$ with the width depending on σ, i.e.

$$\boldsymbol{s} \leftarrow K(\sigma) \star \boldsymbol{c}. \tag{4}$$

The algorithm iterates between Eq. 3 and Eq. 4 until convergence.

FLOOR: Demons assumes the degree of stiffness of the deformation controlled by σ and the error measured by $\text{Sim}(\cdot, \cdot)$ is uniformly defined across the domain Ω. The method also does not improve the results once it converges to a local optimum. These assumptions generally reduce Demons accuracy when, for example, missing correspondences in one region impact the deformation in the remaining image domain. FLOOR drops these assumptions by iteratively fusing and improving the results of Demons based methods (see Algorithm 1).

FLOOR drops the assumption of the uniform stiffness across Ω by replacing σ in Eq. 2 with $\sigma(\cdot)$, a smoothly varying map. At each voxel $x \in \Omega$, $\sigma(\cdot)$ is confined to a small range $\mathbb{S}$, such as $\mathbb{S} = [0.5, 4.5]$. Given that $\sigma(\cdot)$ should be specific to the corresponding image pair, we assume that our approach is initialized with a set of candidates $\mathbb{C} \triangleq \{C_1, \ldots, C_M\}$ where each candidate $C_i \triangleq (\boldsymbol{c}_i, \sigma_i)$ consists of a deformation map $\boldsymbol{c}_i$ that minimizes Eq. 2 for the specific smoothing map σ_i. We furthermore assume that all of those candidates are close to the true

deformation s and smoothing map $\sigma(\cdot)$. Thus, $||\boldsymbol{c}_i(x) - \boldsymbol{s}(x)||^2 \approx 0$ in Eq. 2 so that we omit the term in Eq. 3 and improve the estimate of $\boldsymbol{c}(\cdot)$ at $x \in \Omega$ via

$$(\boldsymbol{c}'(x), \sigma'(x)) \leftarrow \arg\min_{C_i \in \mathbb{C}} ||F(x) - M(x + \boldsymbol{c}_i(x))||^2. \tag{5}$$

If we assume that $\sigma'(x)$ is constant in close proximity of x, i.e. is locally smooth, then $\boldsymbol{c}'(x)$ is equal to $\boldsymbol{s}(x)$. To enforce smoothness, we first compute $\sigma'(\cdot)$ over Ω and then smooth $\sigma'(\cdot)$ similar to Eq. 4 via $\widehat{\sigma} \leftarrow K(\gamma) \star \sigma'$ with γ being the smoothing parameter specific to the smoothing map. Next, we replace the Gaussian kernel in Eq. 4 with one whose width at $x \in \Omega$ is defined by $\widehat{\sigma}(x)$:

$$\boldsymbol{s}(x) \leftarrow K(\widehat{\sigma}(x)) \star_x \boldsymbol{c}'. \tag{6}$$

where $\star_x$ is the convolution at x. We call this process *fusing* deformation maps.

Next, we drop the assumption of misalignment being uniformly accounted for across the domain Ω by fusing candidates, who are optimal solutions with respect to specific image parcels. Specifically, we first parcellate Ω into overlapping regions $\mathbb{O} \triangleq \{\Omega_1, \ldots, \Omega_N\}$ so that $\Omega_i \subset \Omega$, $\Omega_i \cap (\mathbb{O} \setminus \{\Omega_i\}) \neq \emptyset$, and $\cup_{i=1}^N \Omega_i = \Omega$. If we now replace Ω in Eq. 2 with Ω_i then the deformation map $\boldsymbol{s}$ of Eq. 4 is also only defined for Ω_i and not impacted by the mismatches between the images outside of Ω_i. We then create the candidates of $\mathbb{C}$ by applying Demons to Eq. 2 confined by samples taken from $\mathbb{O}$ and the smoothing interval $\mathbb{S}$. Now fusing those candidates according to Eq. 5 and Eq. 6 results in a deformation map over Ω that is the minimum solution with respect to maps defined on overlapping parcels instead of the entire image domain. Unlike the original outcome of Eq. 4, the impact of missing correspondences on $\boldsymbol{s}$ are thus locally confined.

Finally, we incorporate in FLOOR deformation maps from other registration algorithms, which are initialized by the current set $\mathbb{C}$. An example of such an algorithm is MLD[2]. We apply MLD to each collection of candidates, which register a group of images using the same smoothing setting and parcel. For each set, MLD encodes the manifold as a kNN graph with each node representing an image and the weight of the edges between nodes defined by the similarity measure of Eq. 5. MLD's solution to Eq. 2 is then based on the manifold's geodesic path from M to F. The "geodesic" deformations are added to $\mathbb{C}$ and $\boldsymbol{s}$ is updated according to Eq. 5 and Eq. 6. We then learn the manifold based on the new set of $\boldsymbol{s}$ defined over the group of images. We repeat the process of updating $\boldsymbol{s}$ and the manifold until convergence. Note, the algorithm has to converge as each iteration reduces the minimum of Eq. 5.

By improving the pairwise registrations, we also expect the accuracy of the manifolds to be improved. FLOOR thus address an issue of current manifold learning techniques, whose accuracy depend on the initial, pairwise registrations, by iteratively fusing pairwise with manifold-based deformation maps. This idea applies to many other registration algorithms. The fusing of deformation maps as proposed by FLOOR is simple as long as one can first find the minimum locally and then modify the corresponding deformation map to comply with global constraints.

Algorithm 1. FUSING LOCALLY OPTIMAL REGISTRATIONS($\mathbb{S}, \mathbb{O}$)

Create a set of candidates $\mathbb{C}$ by running Demons varying $\sigma \in \mathbb{S}$ and $\Omega \in \mathbb{O}$
repeat
  STEP 1: Train MLDs for a set of images based on $\mathbb{C}$. Add the results to $\mathbb{C}$
  STEP 2: Find the best candidate (c', σ') according to Eq. 5 and smooth σ'
  STEP 3: Update s via Eq. 6 and add to $\mathbb{C}$
until s converges

---

3 Experiments on Cardiac MRI Dataset

We analyze the accuracy of FLOOR on the pairwise registrations of 25 cardiac
MRIs at end-diastole. We measure the accuracy for the 600 registrations (25×24)
by the Dice score [10] between the aligned and corresponding manual segmenta-
tions of the right ventricle blood pool (RV), the left ventricle myocardium (MY)
and blood pool (LV) (see Fig. 2(a)). RV and LV have similar intensity patterns
and MY is relatively thin. Thus, the Dice score with respect to these structures
is a good indicator for the quality of deformations. For comparison, we measure
the accuracy of different implementations of FLOOR.

Data Set: The set consists of short-axis MRIs at end-diastole of 15 healthy and
10 Tetralogy of Fallot (TOF) subjects. Registering these scans is difficult due to
missing correspondences, low image quality, and the irregular ventricular shapes
of TOF hearts (see Fig. 1). The resolution of each scan is $1.25 \times 1.25 \times 8mm^3$.
To speed up registration, we crop scans to 12 slices showing the ventricles and
120×120 inplane voxels centered at the heart. Scans are bias corrected [11] and
linearly aligned to a template [12]. Finally, an expert segmented LV, RV and MY
in each scan via [13]. Note, these maps are only used for the Dice computation.

Parameters: Parameters of FLOOR are the parcellation $\mathbb{O}$, the set of smoothing
parameters $\mathbb{S}$, and the size K of the neighborhood of the kNN graph generated
by MLD. Out of convenience, we parcellate the scans into $3 \times 3 \times 2$ overlapping
parcels as shown in Fig. 2(b). Each parcel is of size $55 \times 55 \times 7$ voxels with
the overlapping regions shown in red and blue. The set of smoothing widths is
$\mathbb{S} = \{0.5, 1.5, 2.5, 3.5, 4.5\}$ as Demons achieves the highest overall accuracy for
$\sigma = 2.5$. Note, we now refer to σ as the actual width of the Gaussian kernel of
Eq. 4 and not the weight of Eq. 2, which indirectly defines σ. Thus, the larger σ
the less flexible the resulting deformation. Finally, we choose K = 10, which is
double the neighborhood size of the thinnest connected kNN graph.

Implementations: As a baseline, we apply Demons [1] to the entire image
with fixed $\sigma = 2.5$ and train MLD on the resulting 600 pairwise registrations
(called GMLD). Next, we run Algorithm 1 omitting Step 1, i.e. not generating
candidates via MLD. Specifically, we fuses the outcome of Demons applied to the
entire image domain with the different smoothness parameter σ of $\mathbb{S}$ (sDemon),
applied to the different parcels of $\mathbb{O}$ but with fixed $\sigma = 2.5$ (pDemon), and by

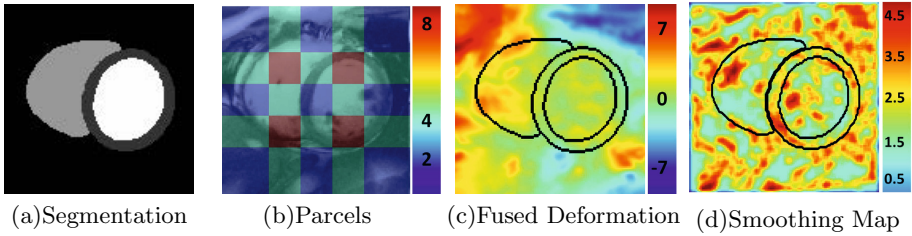(a)Segmentation   (b)Parcels   (c)Fused Deformation   (d)Smoothing Map

Fig. 2. (a) Expert segmentation (LV: white, RV: light grey, MY: dark grey) of Fig. 1(b). (b) The image is divided into $3 \times 3 \times 2$ parcels with the colors indicating the number of overlapping parcels. (c) shows the y-displacement of the fused deformation map and (d) the smoothing map ($\sigma(\cdot)$) generated by FLOOR. Both maps are smooth.

running Demons with different σ as well as on different parcels (fDemon). Finally, we run FLOOR as specified in Alg. 1. pFLOOR applies Demons and learns a MLD for each of the parcels of $\mathbb{O}$ with fixed $\sigma = 2.5$. fFLOOR applies Demons and learns a MLD for different smoothing setting as well as parcels. It also learns a new MLD based on the updated (600) deformations until convergence.

Comparison: We now review Table 1, which lists each implementation's average Dice score (Avg) and standard deviation (std), denoted by $Avg \pm std\%$, across all pairwise registrations for each structure and across structures, i.e. 600 $\times 3$ scores. Note, we call improvements significant if the p-value is below 1% for the unpaired t-test of the scores of two implementations across all image pairs.

Out of all implementations, GMLD received the lowest score (72.9%), which was slightly lower than Demons. The decrease in accuracy indicates that the neighborhood size K was too small. Increasing K did not greatly improve the score until we set it close to 25, which means that every image pair is directly connected, i.e. the results are equivalent to Demons. An alternative explanation for the low score of GMLD is the number of training samples (25), which was too small to properly learn the manifold over the entire image domain. This is a common problem in medical imaging [8,9] and a motivation behind FLOOR.

The score increases by simply fusing the results of Demons, when run with different smoothing parameters (sDemon). There is significant improvement for pDemon and fDemon, who both apply Demons to different parcels. These results indicate that for our data set the overall accuracy of Demons is mostly impacted by missing correspondences in the background and not the uniform smoothness setting of the deformation map. The way our approach addresses this problem is by running Demons on different parcels so that not all of the results are corrupted by the same missing correspondences during the fusion process (Step 2 in Algorithm 1). The corruption free maps most likely will be chosen over the corrupted ones resulting in a more accurate registration.

Finally, the scores further improve when including MLDs such as for pFLOOR vs. pDemon. Unlike for GMLD, whose manifold is based on the entire image domain, the training set is now large enough to properly learn the manifold for specific parcels. fFLOOR receives the highest average scores with 77.9 % after the first iteration and converges to 79.2 % after three iterations. This score is
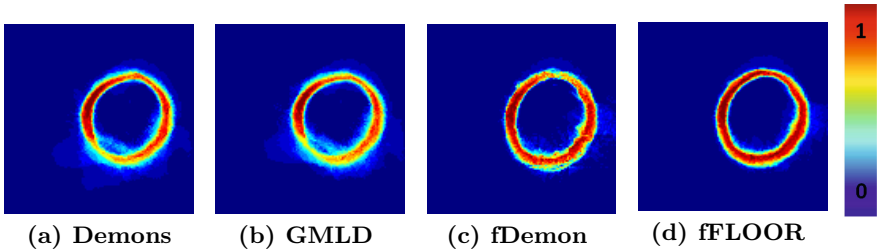
(a) Demons   **(b) GMLD**   **(c) fDemon**   **(d) fFLOOR**

Fig. 3. Average across 24 binary maps of MY aligned to Fig. 1(b) with (d) being best

Table 1. Average and standard deviation of Dice scores of 600 registrations for each individual structure and across all 3 structures. fFLOOR achieves the highest scores

	GMLD	Demon	sDemon	pDemon	fDemon	pFLOOR	fFLOOR
LV	84.9 ± 0.07	85.3 ± 0.06	86.5 ± 0.06	88.3 ± 0.04	88.6 ± 0.04	88.7 ± 0.04	89.7 ± 0.04
MY	58.6 ± 0.11	59.7 ± 0.10	58.0 ± 0.09	65.1 ± 0.08	64.4 ± 0.08	65.8 ± 0.08	68.1 ± 0.08
RV	75.2 ± 0.09	75.8 ± 0.09	76.5 ± 0.09	78.5 ± 0.08	78.9 ± 0.08	79.1 ± 0.08	79.7 ± 0.08
ALL	72.9 ± 0.14	73.6 ± 0.13	73.7 ± 0.13	77.3 ± 0.11	77.6 ± 0.11	77.9 ± 0.12	79.2 ± 0.11

significantly better than fDemon. Thus, manifold learning really improves the results of FLOOR as it provides alternative candidates that locally minimize Demons energy function. Furthermore, fFLOOR produces smooth deformation and smoothing maps as also shown in Fig. 2(c-d). The smooth deformation maps indicate that our method correctly fuses the different candidates produced by fFLOOR. The variation in the smoothing width map support our claim that optimal regularization parameter differ across the image. We end the comparison noting that fFLOOR was significantly better than the state of the art represented by 5.6 % improvement over Demons and 6.3 % over GMLD.

Our findings based on the average Dice score also translates to the scores of the individual structures. Most noticeable are the differences of the implementations with respect to MY, for which the average score of fFLOOR was 68.1% which is 3.7% higher than fDemon, 8.4% higher than Demons, and 9.5% higher than GMLD. The MY is smaller than RV and LV so that the Dice score is most sensitive to deformation errors. This is also reflected by the average maps shown in Fig. 3. The maps are the results of aligning 24 cases to the scan of Fig. 1 (b) and then applying the deformations to the corresponding MY segmentations. The map produced by fFLOOR shows the least amount of variation across the 24 aligned segmentations. The map of fDemon varies a little bit more but far less than the maps of Demons and GMLD. These results further confirm our intuition behind FLOOR that iteratively fusing the deformation maps of pairwise registration with parcel-specific manifolds further improves registration.

4   Conclusion

We derived FLOOR, an algorithm for improving registrations by fusing multiple solutions to the corresponding minimization problem. FLOOR generated

such solutions, called candidates, by applying known methods, such as Demons, varying the regularization and registration regions. FLOOR can also include candidates from other methods, such as MLD, in an iterative fashion. Compared to those original methods, the 600 pairwise registrations of cardiac MR scans of FLOOR were much more accurate in dealing with the large shape variations of the heart and the difference in the field of views across scans. We expect similar results when adapting FLOOR to other registration methods and scans.

References

1. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: Efficient non-parametric image registration. NI 45, S61–S72 (2009)
2. Hamm, J., Ye, D.H., Verma, R., Davatzikos, C.: GRAM: A framework for geodesic registration on anatomical manifolds. MedIA 14(5), 633–642 (2010)
3. Rohlfing, T., Russakoff, D., Maurer, C.R.J.: Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. TMI 23(8), 983–994 (2004)
4. Ashburner, J., Friston, K.J.: Voxel-based morphometry - the methods. NI 11, 805–821 (2000)
5. Yeo, B.T.T., Sabuncu, M.R., Desikan, R., Fischl, B., Golland, P.: Effects of registration regularization and atlas sharpness on segmentation accuracy. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 683–691. Springer, Heidelberg (2007)
6. Glocker, B., Komodakis, N., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through MRFs and efficient linear programming. MedIA 12(6) (2008)
7. Gerber, S., Tasdizen, T., Fletcher, P.T., Joshi, S.C., Whitaker, R.T.: Manifold modeling for brain population analysis. MedIA 14(5), 643–653 (2010)
8. Ye, D.H., Hamm, J., Kwon, D., Davatzikos, C., Pohl, K.M.: Regional manifold learning for deformable registration of brain MR images. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part III. LNCS, vol. 7512, pp. 131–138. Springer, Heidelberg (2012)
9. Bhatia, K.K., Rao, A., Price, A.N., Wolz, R., Hajnal, J., Rueckert, D.: Hierarchical manifold learning. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 512–519. Springer, Heidelberg (2012)
10. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology 26(3), 297–302 (1945)
11. Tustison, N.J., Avants, B.B., Cook, P.A., Zheng, Y., Egan, A., Yushkevich, P.A., Gee, J.C.: N4ITK: improved N3 bias correction. TMI 29(6), 1310–1320 (2010)
12. Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M.: Bayesian analysis of neuroimaging data in fsl. NI 45(suppl. 1), S173–S186 (2009)
13. Heiberg, E., Ugander, M., Carlsson, M., Engblom, H., Arheden, H.: Design and validation of segment–freely available software for cardiovascular image analysis. BMC Med. Imaging 10, 1 (2010)