

# Discriminative Dimensionality Reduction for Patch-Based Label Fusion

Gerard Sanroma<sup>1</sup> (✉), Oualid M. Benkarim<sup>1</sup>, Gemma Piella<sup>1</sup>, Guorong Wu<sup>2</sup>,  
Xiaofeng Zhu<sup>2</sup>, Dinggang Shen<sup>2</sup>, and Miguel Ángel González Ballester<sup>1,3</sup>

<sup>1</sup> Department of Information Technologies and Communication,  
UPF, Barcelona, Spain  
gerard.sanroma@upf.edu

<sup>2</sup> Department of Radiology and BRIC, UNC-Chapel Hill, Chapel Hill, USA

<sup>3</sup> ICREA, Barcelona, Spain

**Abstract.** In this last decade, multiple-atlas segmentation (MAS) has emerged as a promising technique for medical image segmentation. In MAS, a novel target image is segmented by fusing the label maps of a set of annotated images (or atlases), after spatial normalization. Weighted voting is a well-known label fusion strategy consisting of computing each target label as a weighted average of the atlas labels in a local neighborhood. The weights, denoting the local anatomical similarity of the candidate atlases, are often approximated using image-patch similarity measurements. Such an approach, known as patch-based label fusion (PBLF), may fail to discriminate the anatomically relevant patches in challenging regions with high label variability. In order to overcome this limitation we propose a supervised method that embeds the original image patches onto a space that emphasizes the appearance characteristics that are critical for a correct labeling, while suppressing the irrelevant ones. We show that PBLF using the embedded patches compares favourably with state-of-the-art methods in brain MR image segmentation experiments.

## 1 Introduction

Medical image segmentation aims at estimating a dense label map of the anatomical structures in medical images, such as magnetic resonance images (MRI) of the human brain. Quantitative analysis of segmentation data is useful in many fields such as the neurosciences, where the morphometric analysis of brain structures helps characterizing the progression of diseases such as Alzheimer and Schizophrenia [2]. Manual annotation is a tedious time-consuming process which has to be done by trained experts and thus, automatic methods are highly valuable.

Partly enhanced by the success of image registration, multiple-atlas segmentation (MAS) has recently gained attention for segmenting medical images. Three main steps are involved in MAS: (i) the *image registration* step registers each individual atlas onto the target image [7], (ii) the *atlas selection* step selects the best atlases for segmenting a particular target image [1, 13, 14], and

(ii) the *label fusion* step fuses the registered label maps onto a consensus segmentation [3, 4, 10–12, 15–17, 21, 22].

By combining the labels from multiple atlases, the label fusion step can compensate for the registration errors by the individual atlases. Even a simple label fusion strategy such as majority voting [11] (which assigns each target voxel to the label appearing most frequently among the corresponding atlas labels) yields better segmentation performance than any of the single atlases used individually [10].

Another commonly used label fusion strategy is weighted voting, in which the label on each target point is computed as a weighted average of the atlas labels, where the weights reflect the estimated anatomical similarity between the target and each atlas. A critical issue here, is how to set the weights that accurately reflect the anatomical correspondence. One common approach, adopted in patch-based label fusion (PBLF), consists in estimating the weights based on the local similarity between the atlas and the target image patches [3, 4, 12, 17, 22].

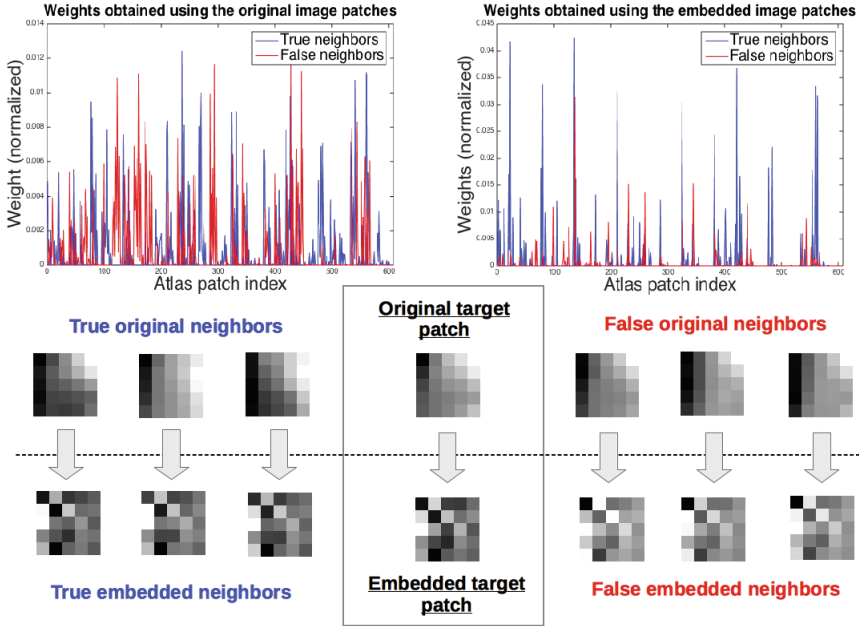
However, there are uncertain regions, such as the interfaces between two anatomical structures, where very similar atlas patches may bear different labels. In such cases, the appearance cues responsible for a correct discrimination may be too weak to be correctly captured by simple image similarity measurements. In order to overcome this limitation we propose a method that learns an embedding of the image patches so that the relevant variations for a correct discrimination are emphasized while the misleading ones are suppressed. We pose this problem in a supervised learning setting, where we seek the linear mapping of the image patches that simultaneously (i) maximizes the similarity of the target patch with its true neighbors (i.e., similar atlas patches with the same label) and (ii) minimizes the similarity with its false neighbors (i.e., similar atlas patches with different label).

The proposed method bears some similarity with manifold learning methods such as neighborhood preserving embeddings (NPE) [9] and locality preserving projections (LPP) [8] in the sense that it aims at preserving the true neighborhood but it simultaneously enforces an additional discriminative component aimed at simultaneously maximizing the separation between false neighbors.

The weights obtained by the proposed method using an example target patch are illustrated in Fig. 1. As we can see in the top-left plot, using the similarity in the original image space, a fair amount of atlas patches with the wrong label (denoted in red) accumulate a considerable amount of weight. On the other hand, using the similarity of the projected patches with the proposed method, the weights of the wrong atlas labels are considerably reduced, while still maintaining a significant amount of weight for the atlas patches with the correct labels (denoted in blue), as shown in the top-right plot.

The contributions of the proposed method are three-fold:

- The embeddings are learned offline in a common space so that they can be readily applied to any new target image.
- The learned embeddings can be plugged into any existing PBLF method to enhance its performance.



**Fig. 1.** Weights obtained using the example target patch in the middle box. **Top row:** estimated weights for the neighboring atlas patches using the original patches (left) and the embedded patches (right). Vertical axis represent the weights and horizontal axis represent the atlas patch index. Using the embedded patches, false neighbors (in red) accumulate less weight than true neighbors (in blue). **Middle row:** Original target patch and some true and false neighbors with high weights (note that we only show the center slice of a cubic  $5 \times 5 \times 5$  patch). **Bottom row:** Most discriminative 25 coordinates of the embedded patches, arranged in a  $5 \times 5$  patch (we show the first 25 coordinates for the convenience of displaying them in a  $5 \times 5$  patch). (Color figure online)

- The proposed method provides a compact representation with a much lower dimensionality than the original patch.

The remainder of the paper is organised as follows: in Sect. 2 we describe the method. In Sect. 3 we present the experiments and results and, finally in Sect. 4 we outline the conclusions.

## 2 Method

### 2.1 Patch-Based Label Fusion

Consider a set of  $n$  atlas images and label maps, denoted as  $\{A^i, L^i\}_{i=1}^n$ , that have been previously registered to a common space, denoted as  $\Omega_C$  (for instance by groupwise non-rigid registration [20]). Therefore,  $A_x^i$  denotes the image intensity value at voxel  $x \in \Omega_C$  of the  $i$ -th atlas, and  $L_x^i \in \{0, 1\}$  denotes whether  $x$

belongs (1) or not (0) to the area of interest to be segmented. We denote as  $T$  the to-be-segmented target image, after being registered to the common space.

Weighted voting label fusion estimates the label at each target point, denoted as  $\hat{F}_x$ , as a weighted average of the neighboring atlas labels<sup>1</sup>. That is,

$$\hat{F}_x = \sum_{i=1}^n \sum_{y \in \mathcal{N}_x} \omega_y^i L_y^i \quad (1)$$

where  $\omega_y^i$  denotes the weight of atlas label  $L_y^i$  at position  $y \in \mathcal{N}_x$  on the  $i$ -th atlas, with  $\mathcal{N}_x$  denoting the local neighborhood of point  $x$ . The local neighborhood  $\mathcal{N}_x$  consists of the patches in a cubic neighborhood of a certain radius from point  $x$ .

The eventual segmentation performance depends on the ability of the label fusion method to identify the true anatomical neighbors of the to-be-segmented target point among the atlas labels. In particular, PBLF assigns higher weights to the atlas locations with higher local image similarity to the target point [3, 4, 12, 17, 22]. As for the image similarity measures, the Gaussian kernel is widely used to estimate the weights [4, 12]. That is,

$$\omega_y^i = \exp(-\|\mathbf{t}_x - \mathbf{a}_y^i\|^2 / \gamma), \quad (2)$$

where  $\mathbf{t}_x, \mathbf{a}_y^i \in \mathbb{R}^p$  denote the vectors of the target and the ( $i$ -th) atlas image patch centered at  $x$  and  $y \in \mathcal{N}_x$ , respectively,  $\gamma$  is a normalization factor, which is set here as in [4] as  $\gamma = \min_{y,i} \|\mathbf{t}_x - \mathbf{a}_y^i\|^2$ , and  $\|\cdot\|$  is the Euclidean norm.

## 2.2 Learning Discriminative Embeddings

PBLF assumes that the higher the similarity between the target patch  $\mathbf{t}_x$  and an atlas patch  $\mathbf{a}_y^i$ , the higher the likelihood that they share the same label. This simplistic assumption, as expressed in Eq. (2), considers that all the features are equally relevant in capturing the anatomical similarity.

Our goal is to learn a transformation for each point  $x \in \Omega_C$ , denoted by the matrix  $\mathbf{P} \in \mathbb{R}^{p \times d}$ , to a lower-dimensional space so that the weights obtained using the transformed patches successfully identify the *anatomically* equivalent patches (rather than the *apparently* similar).

We use all the available atlases as training set, where the image patch from each of the atlases is used as target patch, denoted as  $\mathbf{a}_x^t$ , and the neighboring patches from the rest of the atlases are used as atlas patches, denoted as  $\mathbf{a}_y^i$  (with  $i \neq t$  and  $y \in \mathcal{N}_x$ ).

The training is performed in the common space. This means that all the training images are registered to a template image (built e.g., by groupwise registration [20]). We learn a different transformation (denoted as  $\mathbf{P}$  below) for each point in the common space.

We seek the transformation  $\mathbf{P}$  that simultaneously maximizes the distance with the false neighbors and minimizes the distance with the true neighbors,

<sup>1</sup> The estimated label map in the common space  $\hat{F}$  is finally transformed back to the original target space and thresholded to obtain the binary labels.

where the true (false) neighbors are the sub-set of positive (negative) samples with higher appearance similarity with the target patch. This can be expressed as follows:

$$\max_{\mathbf{P}} \frac{\sum_{t=1}^n \sum_{(i,y) \in \mathcal{F}_x^t} \|\mathbf{P}^\top \mathbf{a}_x^t - \mathbf{P}^\top \mathbf{a}_y^i\|^2 u_y^{t,i}}{\sum_{t=1}^n \sum_{(i',y') \in \mathcal{T}_x^t} \|\mathbf{P}^\top \mathbf{a}_x^t - \mathbf{P}^\top \mathbf{a}_{y'}^{i'}\|^2 v_{y'}^{t,i'}}, \quad (3)$$

where  $u_y^{t,i}$  and  $v_{y'}^{t,i'}$  are the weights identifying the false and true neighbors, respectively (i.e.,  $u_y^{t,i} > 0$  only for those negative samples with higher appearance similarity to the target patch) and  $\mathcal{F}_x^t = \{(i, y) | L_y^i \neq L_x^t, i \neq t, y \in \mathcal{N}_x\}$  is the set of negative samples (the set of positive samples  $\mathcal{T}_x^t$  is similarly defined). We compute the weights  $u_y^{t,i}$  and  $v_{y'}^{t,i'}$  for each target patch in our training set by restricting Eq. (2) to the set of its positive and negative samples, respectively.

The intuition of Eq. (3) is to seek the linear transformation that emphasizes the characteristic differences between false neighbors, so that they are less likely to mislead label fusion, while at the same time downscaling the characteristic differences between true neighbors, so that they end up accumulating more weight. This optimization is somewhat related to linear discriminant analysis (LDA) [5] in the sense that it distinguishes both positive and negative samples for learning the transformation. However, the objective function of LDA is different since it seeks to maximize the between-class scatter and minimize the within-class scatter. In this regard, our approach is more related to manifold learning methods such as locality preserving projections (LPP) [8] and neighborhood preserving embeddings (NPE) [9], but with the difference that we not only minimize the distance with the true neighbors but also jointly maximize the distance with the false neighbors.

Equation (3) can be expressed more compactly as follows:

$$\max_{\mathbf{P}} \frac{\text{Tr} [\mathbf{P}^\top \mathbf{E}_F \mathbf{U} \mathbf{E}_F^\top \mathbf{P}]}{\text{Tr} [\mathbf{P}^\top \mathbf{E}_T \mathbf{V} \mathbf{E}_T^\top \mathbf{P}]}, \quad (4)$$

where  $\mathbf{E}_F = [\dots, \mathbf{e}_y^{t,i}, \dots] \in \mathbb{R}^{p \times q}$  is a matrix with the columns containing vectors of differences between pairs of false neighbors, i.e.,  $\mathbf{e}_y^{t,i} = \mathbf{a}_x^t - \mathbf{a}_y^i$  (with  $t = 1, \dots, n$  and  $(i, y) \in \mathcal{F}_x^t$ ), and  $\mathbf{U} \in \mathbb{R}^{q \times q}$  is a diagonal matrix with the corresponding weights  $u_y^{t,i}$ . ( $\mathbf{E}_T$  and  $\mathbf{V}$  are similarly defined using the differences between true neighbors and their weights, respectively). The larger dimension of the matrices is  $q = nk$ , where  $n$  is the number of atlases in the training set and  $k$  is the expected number of false/true neighbors for each target patch. The solution of Eq. (4) can be found according to the following generalized eigenvalue problem:

$$(\mathbf{E}_T \mathbf{V} \mathbf{E}_T^\top)^{-1} \mathbf{E}_F \mathbf{U} \mathbf{E}_F^\top \mathbf{p} = \lambda \mathbf{p}, \quad (5)$$

where the desired embedding  $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_d]$  is composed of the  $d < p$  eigenvectors with the largest eigenvalues, where  $d$  is the desired number of dimensions.

To avoid the possible over-fitting problem, in Eq. (5) we substitute the matrices  $\mathbf{S}_F \equiv \mathbf{E}_F \mathbf{U} \mathbf{E}_F^\top$  and  $\mathbf{S}_T \equiv \mathbf{E}_T \mathbf{V} \mathbf{E}_T^\top$  by their regularized counterparts [6], as follows:

$$\mathbf{R}_F = (1 - \alpha) \mathbf{S}_F + \frac{\alpha}{p} \text{Tr}[\mathbf{S}_F] I \quad (6)$$

(and similarly for  $\mathbf{S}_T$ ), where  $0 \leq \alpha \leq 1$  is a parameter controlling the amount of regularization and  $I$  is the identity matrix.

In the testing stage, we estimate the label map  $\hat{F}$  of a new target image  $T$  according to the following steps:

1. We register the target image to the common space.
2. For each point  $x \in \Omega_C$ , we extract the surrounding target image patch  $\mathbf{t}_x$  and the set of neighboring atlas image patches from all the atlases, i.e.,  $\mathbf{a}_y^i$ , with  $i = 1, \dots, n$  and  $y \in \mathcal{N}_x$ , where  $\mathcal{N}_x$  is a cubic neighborhood of a certain radius from point  $x$ .
3. We estimate the weights  $\omega_y^i$  for each atlas patch according to Eq. (2) using the embedded target and atlas image patches,  $\mathbf{P}^\top \mathbf{t}_x$  and  $\mathbf{P}^\top \mathbf{a}_y^i$ , respectively.
4. We estimate the label  $\hat{F}_x$  on each target point by fusing the atlas labels according to Eq. (1) using the weights  $\omega_y^i$  estimated in the previous step.
5. We transform the estimated target label map  $\hat{F}$  back to the original target space by using the inverse spatial transformation to the common space.

### 3 Experiments and Results

We compare our proposed approach to the following methods: (i) majority voting (**MV**) [11], which assigns each target label as the label appearing most frequently among the corresponding atlas labels, and (ii) non-local weighted voting (**NLWV**) [4, 12], which uses Eqs. (1) and (2) to estimate the labels and weights, respectively. We apply the proposed discriminative dimensionality reduction on the NLWV pipeline (**DDRNL**), hence we can clearly evaluate the effect of embedding the patches by comparing with the baseline NLWV.

In all the methods, we perform 5-fold cross-validation experiments, where one of the folds is considered as the target images and the rest of the folds as the atlas images. In each fold, the projection matrices  $\mathbf{P}$  learned from the atlas images (one projection for each point in the common space) are used to segment the target images. Target images are segmented in the common space and evaluated in the target space by using the Dice similarity coefficient (DSC) with the ground-truth label maps. We use the group-wise non-rigid registration method in [20] to create the template image defining the common space, and diffeomorphic demons [19] to register the target images to the common space. In both NLWV and DDRNL, we use a patch size of  $5 \times 5 \times 5$  and a cubic search neighborhood  $\mathcal{N}_x$  of  $3 \times 3 \times 3$ . We evaluate the proposed method on brain MR image segmentation experiments on the ADNI<sup>2</sup> and SATA<sup>3</sup> datasets.

<sup>2</sup> <http://www.adni-info.org/>.

<sup>3</sup> [http://masi.vuse.vanderbilt.edu/workshop2013/index.php/Main\\_Page](http://masi.vuse.vanderbilt.edu/workshop2013/index.php/Main_Page).

### 3.1 ADNI Dataset

The ADNI dataset is provided by the Alzheimer’s Disease Neuroimaging Initiative and contains the segmentations of the left and right hippocampi. We use images from 40 randomly selected subjects, where the size of each image is  $256 \times 256 \times 256$ .

We first conduct a sensitivity analysis on a sub-set of 10 randomly selected images. Figure 2(a) shows the sensitivity to the regularization parameter  $\alpha$  and the number of dimensions  $d$ . Based on these results we choose the values of the regularization parameter  $\alpha = 0.9$  and the number of dimensions  $d = 7$  (which is considerably lower than the 125 dimensions of the original  $5 \times 5 \times 5$  patches). In Fig. 2(b) we show the average DSC (%) ( $\pm$ std) in segmenting the left and right hippocampus across the 40 images. As we can see, our proposed method obtains a considerable improvement of  $>1.4\%$  with respect to the NLWV baseline. Results of MV provide a reference of what can be obtained by the only means of non-rigid registration without using any confidence estimates to weigh the atlases.

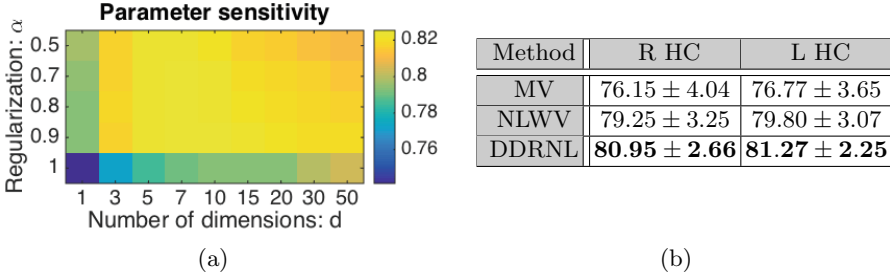


Fig. 2. (a)Parameter sensitivity analysis and, (b) quantitative segmentation results.

### 3.2 SATA Dataset

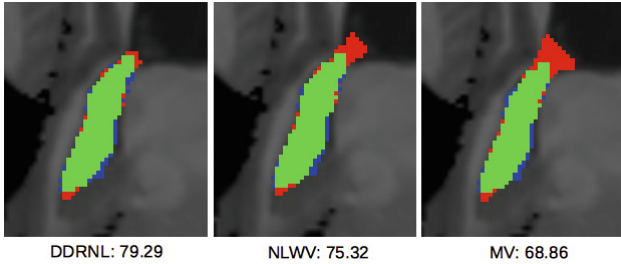
The SATA dataset is composed of 35 images with a resolution of  $1 \times 1 \times 1$  mm and contains the segmentation of 16 mid-brain structures. We will focus on the 10 smallest structures since they tend to be more sensitive to registration errors and hence, more challenging to segment. The segmented structures include the right and left parts of: *accumbens*, *amygdala*, *pallidum*, *caudate* and *hippocampus*.

Figure 3(a) shows the average segmentation performance by each method across the 35 images in each structure. Here, we have used the same parameter values  $d = 7$  and  $\alpha = 0.9$  as in the previous experiments. We have grouped the left and right parts of each structure, so each cell contains the average of 70 segmentations. As we can see, our proposed method obtains a consistent improvement of  $\sim 1\%$  and even  $>1\%$  in some structures with respect to the baseline NLWV. Again, MV provides a reference without resorting to image similarity measurements to weight the atlas contributions.

To gain further insight, in Fig. 3(b) we show the estimated hippocampus segmentations by each method on an example target image. As we can see by the MV results, the head of the hippocampus in this target image is consistently over-segmented by the majority of atlases. NLWV can partially correct this effect by using image-patch similarity to discard some misleading atlases. The proposed method can solve this over-segmentation by using only the discriminative image characteristics in the patch similarity comparisons.

Method	ACC	AMYG	PAL	CN	HC
MV	$67.43 \pm 10.84$	$71.35 \pm 8.36$	$71.16 \pm 15.43$	$79.42 \pm 8.50$	$79.06 \pm 6.22$
NLWV	$73.74 \pm 6.00$	$73.76 \pm 8.02$	$82.04 \pm 7.15$	$86.89 \pm 3.95$	$83.09 \pm 4.08$
DDRNL	<b><math>75.61 \pm 4.95</math></b>	<b><math>74.59 \pm 8.47</math></b>	<b><math>84.30 \pm 4.78</math></b>	<b><math>87.65 \pm 3.47</math></b>	<b><math>83.85 \pm 3.35</math></b>

(a)



(b)

**Fig. 3.** (a) Quantitative segmentation results and, (b) an example of qualitative segmentation results on the hippocampus, where green indicates coincidence with the ground-truth labels (i.e., true positive), red indicates excessive segmentation (i.e., false positives) and blue indicates insufficient segmentation (i.e., false negatives). (Color figure online)

## 4 Discussion and Conclusions

We have presented a dimensionality reduction method to learn optimal patch representations for label fusion that can be plugged into any existing PBLF method. Such representations are learned in the common space so that they can be readily applied to any target image that has been previously aligned to the common space.

Since the proposed method performs label fusion in the common space, the target image needs only to be registered once. This represents a computational advantage with respect to performing it in the target space, since the latter one requires each atlas to be independently registered to the target image space. However, there is some evidence pointing out to the superior performance of using the target space [18]. A possible reason is that pairwise registration accuracy through the common space might not be as accurate as directly warping



the atlas images onto the target image. As a future work, we plan to adapt our method to perform label fusion in the target space.

It is worth noting that the proposed method requires a fair amount of regularization ( $\alpha = 0.9$ ). We believe that this is due to the high complexity involved in learning a different model for each point. A possible solution would be to group the points into perceptually similar regions and learn a single classifier per each region instead of per each point.

We have shown the benefit of the proposed patch representations in the segmentation of several brain structures. We achieve considerable improvements using a compact representation of only 7 dimensions, compared to the 125 dimensions of the original patch.

## References

1. Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, R.V., Rueckert, D.: Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. *NeuroImage* **46**, 726–38 (2009)
2. Amieva, H., Le Goff, M., Millet, X., Orgogozo, J., Prs, K., Barberger-Gateau, P., Jacqmin-Gadda, H., Dartigues, J.: Prodromal Alzheimer’s disease: successive emergence of the clinical symptoms. *Ann. Neurol.* **64**, 492–498 (2008)
3. Artaechevarria, X., Munoz-Barrutia, A., Ortiz-de Solorzano, C.: Combination strategies in multi-atlas image segmentation: application to brain MR data. *IEEE Trans. Med. Imaging* **28**(8), 1266–1277 (2009)
4. Coupé, P., Manjón, J., Fonov, V., Pruessner, J., Robles, M., Collins, D.L.: Patch-based segmentation using expert priors: application to hippocampus and ventricle segmentation. *NeuroImage* **54**(2), 940–954 (2011)
5. Duda, R.O., Hart, P.E., Stork, D.H.: *Pattern Classification*. Wiley Interscience, New York (2000)
6. Friedman, J.H.: Regularized discriminant analysis. *J. Am. Stat. Assoc.* **84**(405), 165–175 (1989)
7. Goshtasby, A.A.: *2-D and 3-D Image Registration: For Medical, Remote Sensing, and Industrial Applications*. Wiley-Interscience, Hoboken (2005)
8. He, X.F., Niyogi, P.: Locality preserving projections. In: *NIPS* (2003)
9. He, X., Cai, D., Yan, S., Zhang, H.: Neighborhood preserving embedding. In: *ICCV* (2005)
10. Heckemann, R.A., Hajnal, J.V., Aljabar, P., Rueckert, D., Hammers, A.: Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage* **33**(1), 115–126 (2006)
11. Rohlfing, T., Brandt, R., Menzel, R., Maurer Jr., C.R.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. *NeuroImage* **21**(4), 1428–1442 (2004)
12. Rousseau, F., Habas, P.A., Studholme, C.: A supervised patch-based approach for human brain labeling. *IEEE Trans. Med. Imaging* **30**(10), 1852–1862 (2011)
13. Sanroma, G., Wu, G., Gao, Y., Shen, D.: Learning-based atlas selection for multiple-atlas segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3111–3117. IEEE (2014)
14. Sanroma, G., Wu, G., Gao, Y., Shen, D.: Learning to rank atlases for multiple-atlas segmentation. *IEEE Trans. Med. Imaging* **33**(10), 1939–1953 (2014)

15. Sanroma, G., Wu, G., Gao, Y., Thung, K.H., Guo, Y., Shen, D.: A transversal approach for patch-based label fusion via matrix completion. *Med. Image Anal.* **24**(1), 135–148 (2015)
16. Sanroma, G., Wu, G., Thung, K., Guo, Y., Shen, D.: Novel multi-atlas segmentation by matrix completion. In: Wu, G., Zhang, D., Zhou, L. (eds.) *MLMI 2014*. LNCS, vol. 8679, pp. 207–214. Springer, Heidelberg (2014)
17. Tong, T., Wolz, R., Hajnal, J.V., Rueckert, D.: Segmentation of brain MR images via sparse patch representation. In: *STMI* (2012)
18. Tong, T., Wolz, R., Coupé, P., Hajnal, J.V., Rueckert, D.: Segmentation of MR images via discriminative dictionary learning and sparse coding: application to hippocampus labeling. *NeuroImage* **76**, 11–23 (2013)
19. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Diffeomorphic demons: efficient non-parametric image registration. *NeuroImage* **45**(1, Supplement 1), S61–S72 (2009)
20. Wu, G., Jia, H., Wang, Q., Shen, D.: Sharp mean: groupwise registration guided by sharp mean image and tree-based registration. *NeuroImage* **56**(4), 1968–1981 (2012)
21. Wu, G., Kim, M., Sanroma, G., Wang, Q., Munsell, B.C., Shen, D., Initiative, A.D.N., et al.: Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. *NeuroImage* **106**, 34–46 (2015)
22. Zhang, D., Guo, Q., Wu, G., Shen, D.: Sparse patch-based label fusion for multi-atlas segmentation. In: Yap, P.-T., Liu, T., Shen, D., Westin, C.-F., Shen, L. (eds.) *MBIA 2012*. LNCS, vol. 7509, pp. 94–102. Springer, Heidelberg (2012)