Minimizing Joint Risk of Mislabeling for Iterative Patch-Based Label Fusion

Guorong Wu[1], Qian Wang[1], Shu Liao[1], Daoqiang Zhang[2], Feiping Nie[3], and Dinggang Shen[1]

[1] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, USA
[2] Department of Computer Science, Nanjing University of Aeronautics and Astronautics, China
[3] Department of Computer Science, University of Texas Arlington, USA

Abstract. Automated labeling of anatomical structures in medical images is very important in many neuroscience studies. Recently, patch-based labeling in the non-local manner has been widely investigated to alleviate the possible misalignment when registering atlases to the target image. However, the weights used for label fusion from the registered atlases in conventional methods are generally computed independently and thus lack the capability of preventing the ambiguous atlas patches from contributing to the label fusion. More critically, these weights are often calculated based only on the simple patch similarity, thus not necessarily providing optimal solution for label fusion. To address these issues, we present a novel patch-based label fusion method in multi-atlas scenario, for the goal of labeling each voxel in the target image by the best representative atlas patches that also have the lowest joint risk of mislabeling. Specifically, sparse coding is used to select a small number of atlas patches which best represent the underlying patch at each point of the target image, thus minimizing the chance of including the misleading atlas patches for labeling. Furthermore, we examine the joint risk of any pair of atlas patches in making similar labeling error, by analyzing the correlation of their morphological error patterns and also the labeling consensus among atlases. This joint risk will be further recursively updated based on the latest labeling results to correct the possible labeling errors. To demonstrate the performance of our proposed method, we have evaluated it on both whole brain parcellation and hippocampus segmentation, and achieved promising labeling results, compared with the state-of-the-art methods.

1 Introduction

With the advent of modern imaging techniques, image analysis plays an important role in quantitatively measuring the structural differences between either individuals or groups. In many neuroscience and clinical studies, certain regions of interest (ROIs), e.g., hippocampus, in the human brain are widely investigated due to their close relation to brain diseases, such as Alzheimer's disease (AD). Consequently, automatic and accurate labeling of anatomical structures becomes critical in those studies that often have large datasets. However, automatic labeling still remains as a

K. Mori et al. (Eds.): MICCAI 2013, Part III, LNCS 8151, pp. 551–558, 2013.

challenging problem because of complicated brain anatomy and high inter-subject variability.

Recently, patch-based labeling methods have emerged as a popular direction for multi-atlas based segmentation [1-4]. The basic assumption of these methods is that, *if two image patches are similar in appearance, they should have the same anatomical label* [1]. Most patch-based labeling methods perform label fusion from registered atlases in a non-local manner [1, 2, 5]. Specifically, to label a patch in the target image, all possible candidate patches from different atlases are considered, with their contributions weighted by their similarities w.r.t. the target patch. In this way, these non-local based labeling methods can alleviate the labeling errors due to possible registration errors.

Although patch-based labeling methods are effective, they also have several limitations. First, all candidate patches from atlases contribute to label fusion, as long as they are similar to the target patch. However, even patches with high appearance similarities from atlases could bear the wrong labels, thus affecting the final labeling result. Second, weights assigned to candidate patches during label fusion are often estimated independently. Thus, if majority of candidate patches have wrong labels, they will dominate the label fusion process and lead to incorrect labeling result [4]. Third, the weights calculated from patch appearance similarity are often directly applied for label fusion. However, although these weights are optimal for patch representation, i.e., making the combination of candidate patches close to the target patch, they are not necessarily optimal for label fusion.

To address these limitations, we aim to label the target patch by the best representative candidate patches that have the lowest joint risk of mislabeling. Specifically, sparse coding [6] is used to select only a small number of candidate patches that can best represent the target patch. In this way, the chance of including the misleading candidate patches for label fusion can be minimized. Moreover, we examine the joint risk of mislabeling by atlas patches in two ways. First, we measure the pairwise correlation of morphological error patterns for any pair of candidate patches, in order to reject those candidate patches with repeated incorrect labels. Second, we further examine whether the labeling result for the target patch achieves the largest labeling consensus among the candidate patches. In this way, our label fusion method is able to iteratively improve the label fusion result by gradually refining the estimation of the joint risk of mislabeling among the candidate atlas patches. We have applied our proposed label fusion method to both whole brain parcellation and hippocampus segmentation. Promising labeling results have been achieved with comparison of the state-of-the-art labeling methods, which indicates the high applicability of our patch-based labeling method in neuroscience and clinical scenario.

2 Methods

Let S be the set of N atlas images $I = \{I_k | k = 1, \dots, N\}$ and their corresponding label maps $L = \{L_k | k = 1, \dots, N\}$, already registered to the target (to-be-labeled) image T. For each point $v \in \Omega_{I_k}$ at the domain of deformed atlas I_k, $\vec{L}_k(v)$ is its binary vector of $\{0,1\}^M$, representing the particular label, where M is the total number of labels.

Thus, the goal of label fusion for labeling the target image T is to propagate the labels from the registered atlases to the target image. For each point $u \in \Omega_T$ in the target image T, its label $\vec{L}_T(u)$ will be estimated through the interaction between the target patch $\mathcal{P}_T(u)$ centered at point u and all possible candidate atlas patches $\mathcal{P}_k(v)$ from the registered atlas images I_k. In order to reduce the computation burden, the search range is usually confined to a relatively small neighborhood $n(u) \subset \Omega_T$. Given the weight $w_k(u,v)$ associated with $\mathcal{P}_T(u)$ and $\mathcal{P}_k(v)$, we can calculate the assignment vector $\theta(u)$ for voxel u as:

$$\vec{\theta}(u) = \frac{\sum_{k=1}^{N}\sum_{v\in n(u)} w_k(u,v)\cdot \vec{L}_k(v)}{\sum_{k=1}^{N}\sum_{v\in n(u)} w_k(u,v)}. \quad (1)$$

It is worth noting that $\vec{\theta}(u) = [\theta^1(u), \dots, \theta^m(u), \dots, \theta^M(u)]$ is a vector of likelihood for each label at point u after label fusion. Then, the final label at point u can be determined by binarizing the vector $\vec{\theta}(u)$ to a binary label vector $\vec{L}_T(u) = [l^1(u), \dots, l^m(u), \dots, l^M(u)]$, where

$$l^m(u) = \begin{cases} 1, & if\ \theta^m(u)\ has\ the\ highest\ value \\ 0, & otherwise \end{cases} \quad (2)$$

2.1 Conventional Patch-Based Labeling Method by Non-local Voting

For each point $u \in \Omega_T$ in the target image, we vectorize the local patch $\mathcal{P}_T(u)$ (red box in Fig. 1(a)) into the column vector $\vec{y}$. In order to account for the registration uncertainty, a set of candidate patches are included in a search neighborhood $n(u)$ (blue boxes in Fig. 1(a)) from different atlas images. For clarity, we arrange each candidate patch $\mathcal{P}_k(v)$ (pink boxes in Fig. 1(a)) into a column vector $\vec{a}_j$ and then assemble them into a matrix $A = \left[\vec{a}_j\right]_{j=1,\dots,Q}$, where $j = (k,v)$ is a bivariate index of particular candidate patch $\mathcal{P}_k(v)$ and $Q = N \cdot |n(u)|$ denotes the total number of candidate patches. With the same order, we assemble the label vector $\vec{L}_k(v)$ of each candidate patch into the label matrix, denoted as $\mathbf{\Lambda} = \left[\vec{\lambda}_1, \vec{\lambda}_2, \dots, \vec{\lambda}_Q\right]$. In non-local voting, each candidate patch $\vec{a}_j$ contributes to label fusion, with the non-local weight w_j in the column vector $\vec{w} = \left[w_1, w_2, \dots, w_Q\right]'$ calculated by:

$$w_j = w_k(u,v) = e^{-\frac{\left\|\vec{y}-\vec{a}_j\right\|^2}{\sigma^2}} \quad (3)$$

where σ is named as the decay parameter in [1] to control the strictness on distance penalty. Given the weighting vector $\vec{w}$ for the point u, we are able to predict the label $\vec{\eta} = \vec{L}_T(u)$ by following Eqs. (1) and (2). The procedure of non-local voting is demonstrated in Fig. 1(b), where each one-end arrow indicates the independent calculation of weight w_j for each candidate patch $\vec{a}_j$.

2.2 Joint Label Fusion with Sparse Patch Representation

In order to alleviate the issue of ambiguous patches in label fusion, sparsity constraint is imposed so that only a small number of candidate patches will contribute to label

fusion. Specifically, by regarding the set of candidate patches $\boldsymbol{A}$ as the local dictionary for the point u, the weights for label fusion can be obtained by solving the following LASSO problem [7]:

$$\hat{\vec{w}} = \arg\min_{\vec{w}} \|\vec{y} - \boldsymbol{A}\vec{w}\|_2^2 + \rho\|\vec{w}\|_1, \; s.t. w_j \geq 0, \forall j, \tag{4}$$

where ρ is the scalar which controls the strength of sparsity. Eventually, most elements in the weighting vector $\vec{w}$ approach to zero after imposing sparsity constraint. As demonstrated in Fig. 1(c), only a few candidate patches (connected to the red box with solid curves) will contribute to the final label fusion, which is different from the conventional non-local based methods (Fig. 1(b)).

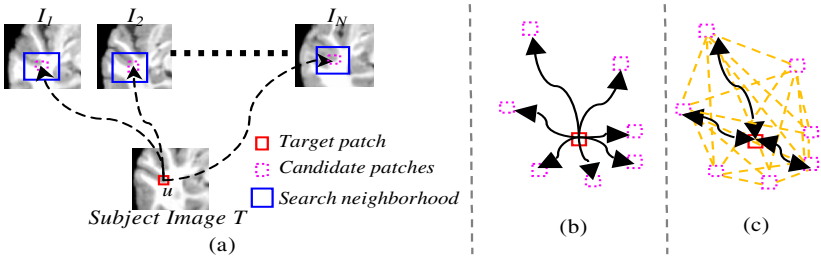

Fig. 1. The overview of patch-based labeling method in multi-atlas scenario. As shown in (a), the target patch (red box) seeks for the linear combination from all possible candidate atlas patches (pink boxes) in a search neighborhood (blue box). (b) and (c) demonstrate the labeling procedure on the particular point by non-local voting and our method, respectively.

It is apparent that Eq. 4 does not account for the dependency among candidate patches in label fusion. In order to make the estimation tractable, we define ϕ_{ij} as the joint risk of mislabeling that two candidates patches $\vec{a}_i$ and $\vec{a}_j$ simultaneously suffer from labeling error. Thus, besides seeking for the optimal patch representation by Eq. 4, the weight vector $\vec{w}$ is also required to minimize the joint risk of labeling error by any pair of patches as

$$\hat{\vec{w}} = \text{argmin}_{\vec{w}} \sum_{i=1}^{Q}\sum_{j=1}^{Q} w_i\phi_{ij}w_j = \arg\min_{\vec{w}} \vec{w}\boldsymbol{\Phi}\vec{w}', \tag{5}$$

where $\boldsymbol{\Phi} = \left[\phi_{ij}\right]_{Q\times Q}$ is a symmetric matrix.

In general, we can learn the joint risk ϕ_{ij} in two ways: (1) correlation of morphological error patterns (w.r.t. $\vec{y}$) between each pair of a_i and a_j in the dictionary $\boldsymbol{A}$, and (2) the labeling consensus between a pair of λ_i and λ_j in $\boldsymbol{\Lambda}$ and the tentative label fusion result $\vec{\eta}$. For the first criterion, we assume that two candidate patches $\vec{a}_i$ and $\vec{a}_j$ have high chance to produce similar labeling error only if their error patterns, i.e., $(\vec{a}_i - \vec{y})$ and $(\vec{a}_j - \vec{y})$, are highly correlated. Thus, we define ϕ_{ij}^A as the correlation of two patches in terms of error pattern as [1]:

$$\phi_{ij}^A = \frac{(\vec{a}_i - \vec{y})}{\|\vec{a}_i - \vec{y}\|} \cdot \frac{(\vec{a}_j - \vec{y})}{\|\vec{a}_j - \vec{y}\|}. \tag{6}$$

[1] We will further normalize ϕ_{ij}^A to the range from 0 to 1.

Given the tentatively estimated weighting vector $\hat{\vec{w}}$, we can determine the labeling result $\hat{\vec{\eta}}$ by Eq. 1~2. Here, we go one step further to examine whether the estimation $\hat{\vec{\eta}}$ achieves the largest labeling consensus between any pair of label λ_i on patch $\vec{a}_i$ and λ_j on patch $\vec{a}_j$. We use $\phi_{ij}^{\wedge}$ to measure the labeling consensus between λ_i and λ_j as:

$$\phi_{ij}^{\wedge} = 1 - \frac{[\delta(\lambda_i - \hat{\vec{\eta}}) + \delta(\lambda_j - \hat{\vec{\eta}})]}{2}, \tag{7}$$

where $\delta(\cdot)$ is the Dirac pulse function at zero, i.e., $\delta(0) = 1$. It is apparent that $\phi_{ij}^{\wedge} = 1$ only if both candidate patches $\vec{a}_i$ and $\vec{a}_j$ bear different labels against $\hat{\vec{\eta}}$, for which our method will panelize the weights w_i and w_j to have high values. In this way, our method is able to not only prevent the repeated labeling error from two correlated patches, but also provide our label fusion method a chance to iteratively correct the possible mislabeling by refining the labeling consensus $\phi_{ij}^{\wedge}$. Finally, the joint risk of labeling error ϕ_{ij} is defined as:

$$\phi_{ij} = (1 - r) \cdot \phi_{ij}^{A} + r \cdot \phi_{ij}^{\wedge} \tag{8}$$

where $0 \le r \le 1$ is the scalar balancing the two measurements. $r = 0$ in the beginning since there is no estimation of $\vec{\eta}$ at that moment. We increase r from 0 to 0.5 linearly and iteratively. As we will show later, the labeling performance keeps improving as iteration progresses.

By integrating Eq. 8 into Eq. 4, the new energy function can be formulated as:

$$\left(\hat{\vec{w}}, \hat{\vec{\eta}}\right) = \arg\min_{(\vec{w},\vec{\eta})} \|\vec{y} - \boldsymbol{A}\vec{w}\|_2^2 + \beta \cdot \vec{w}\boldsymbol{\Phi}\vec{w}' + \rho \cdot \|\vec{w}\|_1, \; s.t. \, w_j \ge 0, \forall j, \tag{9}$$

where β controls the strength of joint labeling risk. Fig. 1(c) demonstrates the improvements of our method: (1) a small number of candidate patches are eventually involved in label fusion (i.e., few dash boxes are connected with red box by solid curves in Fig. 1(c)); (2) the dependencies of candidate patches are clearly described (as indicated by the dash lines in Fig. 1(c)); (3) the label fusion result $\vec{\eta}$ is recursively used to refine the labeling result by updating the joint labeling risk $\boldsymbol{\Phi}$ (as indicated by the two-end arrows in Fig. 1(c)). It is worth noting that our method will degrade to the sparse patch-based labeling method [8] if $\hat{a} = 0$.

To solve this problem, we alternatively repeat two sub-steps, i.e., (1) optimizing $\vec{w}$ by fixing $\vec{\eta}$, and (2) updating $\boldsymbol{\Phi}$ with the latest $\vec{\eta}$ by Eq. 1~2. For step (1), we use coordinate descent method [9] to efficiently find $\vec{w}$. The idea is to go through each w_j in $\vec{w}$ and minimize the energy function in Eq. 9 along one w_j at a time. Specifically, for each w_j, we discard all terms in Eq. 9 that are not related with w_j and turn Eq. 9 into:

$$\hat{w}_j = \arg\min_{w_j} \left\|\vec{\xi} - w_j\vec{a}_j\right\|_2^2 + \beta\phi_{jj}w_j^2 + 2\beta b w_j + \rho|w_j|, \tag{10}$$

where $\vec{\xi} = \vec{y} - \sum_{i=1,i\neq j}^{Q} w_i\vec{a}_i$ and $b = \sum_{i=1,i\neq j}^{Q} \phi_{ij}w_i$.

By letting $\grave{\imath} = \left(\vec{\xi}^T\vec{a}_j - \beta b\right)/\left(\vec{a}_j^T\vec{a}_j + \beta\phi_{jj}\right)$ and $\tau = \rho/\left(\vec{a}_j^T\vec{a}_j + \beta\phi_{jj}\right)$, we further rewrite Eq. 10 as:

$$\hat{w}_j = \arg\min_{w_j} \left(w_j - \mu\right)^2 + \tau|w_j|, s.t. \, w_j \ge 0, \tag{11}$$

which turns to the classic l_2 regression problem [9]. Thus, by requiring the directional derivative along the coordinate direction of w_j coincide with the ordinary partial derivative $\partial(w_j - \mu)^2/\partial w_j$, the optimal solution to particular $\hat{w}_j$ is given as:

$$\hat{w}_j = \begin{cases} \mu - \frac{\tau}{2} & \mu > \frac{\tau}{2} \\ 0 & \mu \leq \frac{\tau}{2} \end{cases} \tag{12}$$

In our implementation, we randomly set the order to visit each element w_j and fix the iteration number as 200 in optimizing $\vec{w}$ for each point u in the target image.

3 Experiments

We apply our joint patch-based labeling method (Joint-PBM) on both ADNI and NIREP-NA40 [10] datasets to evaluate the labeling performance. For comparison, we also deploy the conventional patch-based method by non-local weighting (Nonlocal-PBM) and recently proposed sparse patch-based method [8] (Sparse-PBM) on the same datasets. For each subject, we first use FLIRT in FSL package (http://fsl.fmrib.ox.ac.uk/) and then diffeomorphic Demons [2][11] to deform all atlases to the underlying subject space.

In the following experiment, we fix the patch size as $7 \times 7 \times 7$ and the search range as $9 \times 9 \times 9$ for all label fusion methods. We follow the patch pre-selection strategy and local adaptive selection of decay parameter in [1] for Nonlocal-PBM. Also, according to [8], we set the parameter for l_1-norm strength as 0.1 for sparse-PBM. For our Joint-PBM method, we fix $\beta = 1.0$ and $\rho = 0.1$ for all experiments. Particularly, we repeat the steps of estimating $\vec{w}$ and updating labels $\vec{\eta}$ for 5 times.

3.1 Experiment Result on Hippocampus Labeling in ADNI Dataset

In this experiment, we randomly select 61 NC (normal control) subjects, 96 MCI (Mild Cognitive Impairment) subjects, and 41 AD (Alzheimer's disease) subjects from the ADNI dataset. For each subject image, the skull is stripped and the intensity range has been normalized by histogram matching. The segmentation ground truth of hippocampus for each image is also provided in the ADNI dataset.

The overall Dice ratio of left/right hippocampus by three label fusion methods are shown in Table 1, where our Joint-PBM method has achieved 4.7% and 2.2% improvements over Nonlocal-PBM and Sparse-PBM, respectively. Moreover, we show the average and standard deviation of Dice ratio in NC, MCI and AD groups by three label fusion methods in Table 2. Again, our Joint-PBM method has the best labeling performance in each group. It is worth noting that the highest Dice ratio of hippocampus is 0.893 in [4]. However, only 57 NC and 82 MCI subjects are included in [4].

[2] The parameters for diffeomorphic Demons are: 15, 15, and 15 iterations in low, middle, and high resolutions. The kernel size for deformation smoothing is set to 1.8. The option of histogram matching is turned on.

As shown in Table 2, labeling hippocampus of AD subjects are more challenging than MCI and NC groups. Discarding the AD subject, the overall Dice ratio of NC and MCI groups by our method is able to reach 0.896. As our labeling method iteratively refines the labeling result by increasing r (Eq. 8), the Dice ratio increases consistently from 0.882 to 0.887 in the end of labeling.

Table 1. The mean and standard deviation of Dice ratio on left/right hippocampus by Nonlocal-PBM, Sparse-PBM, and Joint-PBM

Method	Left Hippocampus	Right Hippocampus	Overall
Nonlocal-PBM	0.854 ± 0.040	0.849 ± 0.043	0.852 ± 0.042
Sparse-PBM	0.877 ± 0.032	0.869 ± 0.036	0.873 ± 0.034
Joint-PBM	$\mathbf{0.890 \pm 0.022}$	$\mathbf{0.884 \pm 0.023}$	$\mathbf{0.887 \pm 0.022}$

Table 2. The mean and standard devidation of Dice ratio in hippocampus labeling for three groups (NC, MCI, and AD) by Nonlocal-PBM, Sparse-PBM, and Joint-PBM

Method	NC	MCI	AD
Nonlocal-PBM	0.866 ± 0.034	0.859 ± 0.039	0.831 ± 0.046
Sparse-PBM	0.882 ± 0.030	0.873 ± 0.036	0.864 ± 0.041
Joint-PBM	$\mathbf{0.899 \pm 0.014}$	$\mathbf{0.893 \pm 0.019}$	$\mathbf{0.870 \pm 0.032}$

3.2 Experiment Result on Whole Brain Labeling in NIREP-NA40 Dataset

The NIREP-NA40 dataset consists of 16 MR images of 8 normal male adults and 8 normal female adults, each with 32 manually-delineated ROIs. The image size is $256 \times 300 \times 256$ and the voxel resolution is $0.7 \times 0.7 \times 0.7mm^3$.

Table 3. The average Dice ratio in NIREP-NA40 dataset by three label fusion methods

ROI (left+right)	Nonlocal-PBM	Sparse-PBM	Joint-PBM
Occipital Lobe	0.801 / 0.813	0.815 / 0.820	**0.833 / 0.859**
Cingulate Gyrus	0.815 / 0.812	0.811 / 0.814	**0.819 / 0.852**
Insula Gyrus	0.851 / 0.873	0.855 / 0.878	**0.862 / 0.890**
Temporal Pole	0.837 / 0.829	0.838 / 0.841	**0.842 / 0.875**
Superior Temproal Gyrus	0.779 / 0.777	0.781 / 0.784	**0.801 / 0.811**
Infero Temporal Region	0.832 / 0.833	0.848 / 0.832	**0.867 / 0.871**
Parahippocampal Gyrus	0.829 / 0.843	0.831 / 0.851	**0.842 / 0.864**
Frontal Pole	0.820 / 0.804	0.824 / 0.820	**0.849 / 0.852**
Superior Frontal Gyrus	0.807 / 0.785	0.805 / 0.800	**0.822 / 0.837**
Middle Frontal Gyrus	0.791 / 0.753	0.809 / 0.763	**0.819 / 0.805**
Inferior Gyrus	0.755 / 0.751	0.758 / 0.775	**0.785 / 0.790**
Orbital Frontal Gyrus	0.833 / 0.831	0.841 / 0.835	**0.863 / 0.860**
Precentral Parietal Lobule	0.762 / 0.744	0.785 / 0.762	**0.801 / 0.789**
Superior Parietal Lobule	0.742 / 0.739	0.756 / 0.780	**0.802 / 0.805**
Inferior Parietal Lobule	0.759 / 0.738	0.771 / 0.802	**0.800 / 0.812**
Postcentral Gyrus	0.707 / 0.694	0.713 / 0.711	**0.756 / 0.738**

The leave-one-out strategy is used in this experiment by alternatively taking 16 images as the target image in each leave-one-out case. Table 3 shows the average Dice ratio in each brain structure (left and right combined) across 16 leave-one-out cases. The overall Dice ratios across 32 ROIs are 0.792 by Nonlocal-PBM, 0.803 by Sparse-PBM, and 0.827 by Joint-PBM. Our Joint-PBM method achieves the best labeling accuracy over the other two methods, by 3.5% and 2.4%, respectively.

4 Conclusion

In this paper, we present a novel patch-based label fusion method for multi-atlas segmentation. Sparsity constraint is introduced in our method to suppress the misguidance from ambiguous patches. Furthermore, our method explicitly describes the dependency of patches from different atlases or same atlas to guide the label fusion, which is iteratively updated based on the latest label fusion result. In this way, our method is able to gradually improve the labeling accuracy by reducing the chance of making the repeated labeling errors. Our experiments on hippocampus segmentation and whole brain parcellation show the promising labeling results, indicating its applicability in neuroscience and clinical studies.

References

1. Coupe, P., et al.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. NeuroImage 54(2), 940–954 (2011)
2. Rousseau, F., Habas, P.A., Studholme, C.: A Supervised Patch-Based Approach for Human Brain Labeling. IEEE Trans. Medical Imaging 30(10), 1852–1862 (2011)
3. Tong, T., et al.: Segmentation of Brain Images via Sparse Patch Representaion. In: STML, Nice, France (2012)
4. Wang, H., et al.: Multi-Atlas Segmentation with Joint Label Fusion. IEEE Trans. Pattern Anal. Mach. Intell. (2012)
5. Buades, A., Coll, B., Morel, J.-M.: A non-local algorithm for image denoising. In: Computer Vision and Pattern Recognition, Beijing, China (2005)
6. Vinje, W.E., Gallant, J.L.: Sparse coding and decorrelation in primary visual cortex during natural vision. Science 287(5456), 1273–1276 (2000)
7. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 58(1), 267–288 (1996)
8. 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)
9. Wu, T.T., Lange, K.: Coordinate descent algorithms for lasso penalized regression. The Annals of Applied Statistics 2(1), 224–244 (2008)
10. Christensen, G.E., Geng, X., Kuhl, J.G., Bruss, J., Grabowski, T.J., Pirwani, I.A., Vannier, M.W., Allen, J.S., Damasio, H.: Introduction to the non-rigid image registration evaluation project (NIREP). In: Pluim, J.P.W., Likar, B., Gerritsen, F.A. (eds.) WBIR 2006. LNCS, vol. 4057, pp. 128–135. Springer, Heidelberg (2006)
11. Vercauteren, T., et al.: Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 45(suppl. 1), S61–S72 (2009)