Medical Image Analysis 24 (2015) 135–148

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

A transversal approach for patch-based label fusion via matrix completion

Gerard Sanroma[a], Guorong Wu[a], Yaozong Gao[a], Kim-Han Thung[a], Yanrong Guo[a], Dinggang Shen[a,b,*]

[a] Department of Radiology and BRIC, University of North Carolina, Chapel Hill, USA
[b] Department of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 22 October 2014
Revised 10 April 2015
Accepted 11 June 2015
Available online 20 June 2015

Keywords:
Label fusion
Matrix completion
Multiple-atlas segmentation

ABSTRACT

Recently, multi-atlas patch-based label fusion has received an increasing interest in the medical image segmentation field. After warping the anatomical labels from the atlas images to the target image by registration, label fusion is the key step to determine the latent label for each target image point. Two popular types of patch-based label fusion approaches are (1) *reconstruction-based approaches* that compute the target labels as a weighted average of atlas labels, where the weights are derived by reconstructing the target image patch using the atlas image patches; and (2) *classification-based approaches* that determine the target label as a mapping of the target image patch, where the mapping function is often learned using the atlas image patches and their corresponding labels. Both approaches have their advantages and limitations. In this paper, we propose a novel patch-based label fusion method to combine the above two types of approaches via matrix completion (and hence, we call it transversal). As we will show, our method overcomes the individual limitations of both reconstruction-based and classification-based approaches. Since the labeling confidences may vary across the target image points, we further propose a sequential labeling framework that first labels the highly confident points and then gradually labels more challenging points in an iterative manner, guided by the label information determined in the previous iterations. We demonstrate the performance of our novel label fusion method in segmenting the hippocampus in the ADNI dataset, subcortical and limbic structures in the LONI dataset, and mid-brain structures in the SATA dataset. We achieve more accurate segmentation results than both reconstruction-based and classification-based approaches. Our label fusion method is also ranked 1st in the online SATA Multi-Atlas Segmentation Challenge.

1. Introduction

Parcellation of the human brain structures is a key image processing step in many medical imaging studies related to computational anatomy and computer aided diagnosis (Li et al., 2014; Li et al., 2010; Nie et al., 2013; Nie et al., 2011). Manual annotation of anatomical structures is tedious and very time consuming, which makes it impractical in most of the current medical studies involving large amounts of imaging data. Therefore, high-throughput and accurate automated segmentation methods are highly desirable.

In the last two decades, multi-atlas segmentation (MAS) has emerged as a promising automated segmentation technique for segmenting a target image by propagating the labels from a set of annotated atlases. The use of multiple atlases makes MAS more capable of accommodating higher anatomical variability than using a single atlas. Moreover, as demonstrated in (Collins and Pruessner, 2009; Isgum et al., 2009; Rohlfing et al., 2004b), segmentation errors made by each individual atlas tend to be corrected when using multiple atlases. Generally, MAS consists of the following three steps: (1) the *atlas selection* step, where a subset of best atlases is first selected for a given target image based on a certain pre-defined measurement of anatomical similarity (Aljabar et al., 2009; Collins and Pruessner, 2009; Isgum et al., 2009; Rohlfing et al., 2004b; Sanroma et al., 2014a; Wu et al., 2007); (2) the *registration* step, where all selected atlases and their corresponding label maps are aligned to the target image (Klein et al., 2009; Shen and Davatzikos, 2002; Vercauteren et al., 2009; Wu et al., 2011); and finally (3) the *label fusion* step, where the registered label maps from the selected atlases are fused into a consensus label map for the target image (Artaechevarria et al., 2009; Cardoso et al., 2013; Coupe et al., 2011; Hao et al., 2013; Jia et al., 2012; Kim et al., 2013; Rousseau et al., 2011; Wang et al., 2011b; Warfield et al., 2004; Zikic et al., 2013). A great deal of attention has been put into the label fusion step, which is also the focus of the present paper, since it has a great influence on the final segmentation performance.

* Corresponding author.
E-mail address: dgshen@med.unc.edu (D. Shen).

http://dx.doi.org/10.1016/j.media.2015.06.002
1361-8415/

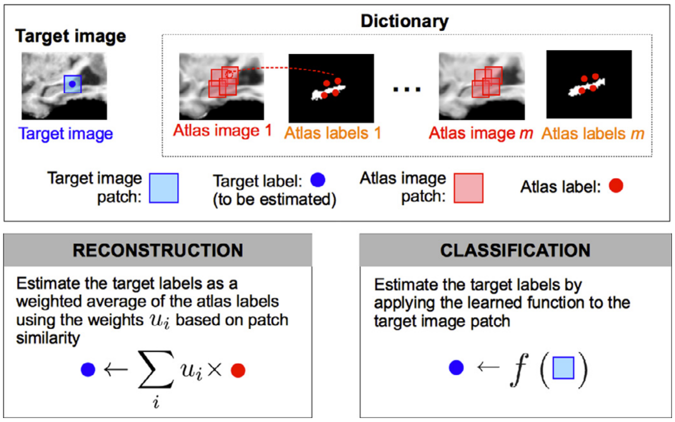

Fig. 1. Illustration of reconstruction-based and classification-based label fusions. **Top**: a dictionary of atlas image patches (red squares) and their center labels (red circles) are used to estimate the target label (blue circle) in the center of the target image patch (blue square). **Bottom-left**: reconstruction-based approaches estimate the target label as a weighted average of the atlas labels, where atlas patches with higher similarity are assigned higher weights. **Bottom-right**: classification-based approaches estimate the target label by applying the relationships learned using the dictionary of atlas patches and labels. (See Sec. 3.1 for details about how the reconstruction and classification functions are computed.)

During the label fusion step, each target point is often independently labeled by using its own *dictionary* composed of the atlas patches and their labels selected from a neighborhood of the to-be-labeled target point (Coupe et al., 2011; Hao et al., 2013; Rousseau et al., 2011) (see the top panel in Fig. 1). Two recently popular label fusion approaches are the following: (1) reconstruction-based approaches, and (2) classification-based approaches. Reconstruction-based approaches are a particular type of weighted voting methods. As such, the target label is computed as a weighted average of the atlas labels (see the bottom-left panel in Fig. 1). Specifically, reconstruction-based approaches assign the weights based on the coefficients obtained by the linear reconstruction of the target patch using the dictionary of atlas patches (Tong et al., 2012; Zhang et al., 2012). This follows the idea of the image-similarity approaches, which assign higher weights to the atlas patches with more similarity to the target patch (Artaechevarria et al., 2009; Coupe et al., 2011; Rousseau et al., 2011). On the other hand, classification-based approaches use the dictionary of atlas image patches and their corresponding labels as the training set to learn the relationships between image appearance and anatomical labels (Hao et al., 2013) (Wang et al., 2011b). Then, in the labeling stage, the target label is estimated by directly applying the learned relationships to the target image patch (see the bottom-right panel in Fig. 1).

However, both reconstruction-based and classification-based approaches have their own limitations. Reconstruction-based approaches assume that the weights optimized based on patch-wise similarity are also optimal to fuse the labels. Unfortunately, as demonstrated in (Sanroma et al., 2014a), there is not always a clear relationship between appearance similarity and label consensus, and therefore similar atlas image patches could bear different labels. On the other hand, classification-based approaches overcome this limitation by specifically learning a mapping function from the image appearance domain to the label domain. However, all the atlas patches in the dictionary are given the same importance during the learning procedure, which may not be optimal since not all patches in the dictionary are equally representative for the target patch. Reconstruction-based approaches overcome this issue by adaptively weighting each atlas patch according to their estimated relevance in predicting the label of a particular target image point. In light of this, we present a novel label fusion method with the following contributions:

- We combine the advantages of both reconstruction-based and classification-based approaches by formulating label fusion as a matrix completion problem (but our method restricts to the *linear* sub-type of approaches). *First*, we build an *incomplete* matrix containing the target image patch as well as the atlas patches and their labels, where all the to-be-estimated target labels are missing. Based on the observation that there are high correlations among image patches and labels, we employ a low-rank constraint to estimate the missing elements in the above matrix. This entails taking full advantage of both row-wise and column-wise correlations (Candès and Recht, 2009), corresponding to the correlations in the vertical and horizontal directions of the matrix, respectively. As we will show, both reconstruction-based and classification-based approaches are particular cases where only row-wise (i.e., vertical) or column-wise (i.e., horizontal) correlations are exploited, respectively. By exploiting both types of correlations, our transversal method inherits the properties of both reconstruction-based and classification-based approaches, namely, (1) the property of the reconstruction-based approaches of representing the target patch as a weighted combination of the atlas patches, and (2) the discriminative ability of the classification-based methods in modeling the dependence of anatomical labels on the image appearance.
- We note that the labels at some parts of the image (e.g., deep inside the structures) can be determined more reliably than other parts (e.g., at boundaries of the structures), due to their anatomical characteristics and also due to their robustness to registration errors. However, most patch-based label fusion methods do not acknowledge this fact and label each target point independently. In this paper, we argue that it is more reasonable to let the high-confident points guide the labeling procedure of nearby less-confident points. Specifically, we embed our label fusion method into a sequential labeling framework that first labels the most confident target points and gradually labels those less-confident points iteratively. In this way, the anatomical labels estimated from the previous iterations can be used to help select more anatomically similar atlas patches to build the dictionary for improving the labeling.

We evaluate the label fusion performance of our proposed method on the ADNI, LONI, and SATA segmentation challenge datasets. We show that our proposed matrix completion based label fusion method outperforms both reconstruction-based and classification-based approaches. Moreover, we show that the sequential confidence-guided labeling scheme further improves our proposed method. Most importantly, our proposed method is ranked 1st in the online SATA Segmentation Challenge.

Note that a preliminary version of this work was presented in Sanroma et al., (2014b). The current paper (1) extends our previous work with the sequential confidence-guided labeling approach as described in Sec. 3.2, and (2) provides more exhaustive descriptions as well as illustrative examples of our extended method. We extensively (3) evaluate each component of our extended method by using additional datasets, and (4) compare it with the state-of-the-art methods.

2. Related work

With the advent of MAS, label fusion has become an increasingly active area of research. Label fusion is the key step that aims to segment the target image by finding a consensus among a set of registered atlas labels. The way in which the atlas information is used to derive the consensus segmentation has given rise to many different label fusion flavors. The simplest way, known as majority voting (MV), simply assigns each target point the label that appears most frequently among all corresponding atlas points (Heckemann et al., 2006; Rohlfing et al., 2005).