Pattern Recognition 46 (2013) 2408–2419

Contents lists available at SciVerse ScienceDirect

Pattern Recognition

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

Hierarchical MRF of globally consistent localized classifiers for 3D medical image segmentation

Sang Hyun Park [a], Soochahn Lee [b], Il Dong Yun [c,*], Sang Uk Lee [a]

[a] Department of Electrical Engineering, ASRI, INMC, Seoul National University, 151-744 Seoul, Republic of Korea
[b] Samsung Electronics, Maetan-3-Dong, Gyunggi-Do, Republic of Korea
[c] Department of Digital Information Engineering, Hankuk University of Foreign Studies, 449-791 Yongin, Republic of Korea

ARTICLE INFO

Article history:
Received 10 July 2012
Received in revised form
29 January 2013
Accepted 20 February 2013
Available online 5 March 2013

Keywords:
Segmentation
Hierarchical Markov random field
Medical image analysis
Discrete optimization
Local and global prior

ABSTRACT

A suitable object model is crucial in guiding an object segmentation method of three-dimensional medical images to avoid difficulties such as complex object structures, inter-subject variability and ambiguous boundaries between organs. The main challenge is to make the model sufficiently complex to represent a wide range of variations effectively, while maintaining compatibility with the segmentation methodology. To address this problem, we propose a new segmentation method based on a hierarchical Markov random field (H-MRF). The H-MRF is composed of local-level MRFs based on adaptive local priors which model local variations of shape and appearance and a global-level MRF enforcing consistency of the local-level MRFs. The proposed method can successfully model large object variations and weak boundaries and is readily combined with well-established MRF optimization techniques. Furthermore, it works well with limited training data and does not require a complex training model or non-rigid registration. The performance of the proposed method is evaluated for bone and cartilage from knee magnetic resonance (MR) images, the liver from body computed tomography images, and the hippocampus from brain MR images. Both qualitative and quantitative evaluations demonstrate that the proposed method provides robust and accurate segmentation results.

1. Introduction

Object segmentation is a fundamental problem in computer vision. Over the years, several different methods based on low level statistics such as color or intensity have been established [1–4]. Although these methods work well on many applications, the models are often insufficient to reflect the characteristics of medical images, which include deformable object structures, inter-subject variability, and ambiguous boundaries between organs. Since small errors may critically affect analysis of disease or pathology, significant user interaction or post-processing are required for practical use. For example, cartilage segmentation from a knee MR image takes from 30 to 50 min due to repetitive modification, even through a semi-automatic method based on graph cut operated by experts [5]. Recently, many methods [6–26] incorporate a high level of prior knowledge learned from training data into their framework. Compared with the earlier methods, these methods demonstrate enhanced segmentation accuracy, despite requiring less user interaction. Especially, these methods have been proven to be useful in medical image analysis, since high level priors are much better in modeling the deformation of organs in medical images compared to low level statistics. Here, specific applications of medical image segmentation methods include longitudinal studies based on a baseline segmentation label, diagnosis of degenerative organs, and registration of images with different modalities.

Heckemann et al. [6] and Aljabar et al. [7] proposed methods based on warping templates. In those methods, voxels are classified into foreground (FG) or background (BG) by majority voting of the warped training templates to the target object. Although the methods are simple and instinctive, the performances are highly influenced by non-rigid registration algorithms that usually have high complexity. Coupe et al. [8] and Rousseau et al. [9] proposed patch-based label fusion methods in order to reduce the registration dependency. They align the training templates to the target object by rigid registration algorithms, and fuse the labels with a weight in local search regions. The weight, determined by the appearance similarity of local patches, emphasizes the prior knowledge of the locally similar training region. Although the methods reduce the relative importance of registration, the locally fused labels do not cover the variations of highly deformable objects. Most of the label fusion methods have been only applied to the segmentation of brain structures with

* Corresponding author. Tel.: +82 31 330 4260.
E-mail addresses: shpark@cvl.snu.ac.kr (S.H. Park), sc.lee@samsung.com (S. Lee), yun@hufs.ac.kr (I.D. Yun), sanguk@ipl.snu.ac.kr (S.U. Lee).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.02.014

small shape variations. Furthermore, it is slow to segment an object of large size because the process has to be done at every voxel inside an initial mask.

To learn a more systematic model, Cootes et al. proposed the active shape model [10]. Here, landmark points are extracted, and a statistical shape model is constructed by the covariance of landmark positions. Since the model preserves the object shape, even on vague boundaries, the method has been applied to segment various organs including the liver [11], bone [12] and cartilage [13]. However, the training of the good statistical model that flexibly covers the object variations is difficult with a limited training set, and the extraction of landmarks is laborious during the training stage.

Unlike the active shape model that uses an explicit shape representation based on the landmarks, variational methods [14–17] impose an implicit shape prior into a level set framework. Leventon et al. [14] and Rousson et al. [15,16] represented the training shapes by a signed distance function and mean level set function with the Gaussian fluctuations model, respectively. More recently, Cremers et al. [17] used a dynamical shape prior to selectively apply the shape knowledge according to regions. Although these methods have a flexible framework with easy modification and noise robustness, they are sensitive to initialization, and prone to converge at the local minima.

The combinatorial optimization approach underlying a Markov random field (MRF) has been extensively researched. Like the variational methods, prior knowledge is implicitly imposed into the energy function, but the energy is optimized by discrete optimization methods, such as graph cut [2,27], belief propagation [28], and tree-reweighted message passing (TRW) [29]. Freedman et al. [18] used the distance transform of shape templates, while Lijn et al. [19] and Lotjonen et al. [20] used the fused labels of warped templates to define the energy function. Corso et al. [21] incorporated a local image context term based on learned features into the energy function. In these methods, the appearance prior helps to find the local variations, while the shape prior helps to preserve the object shape. One of their limitations is that a weight, controlling the shape prior and the appearance prior, is equally set in the whole image without considering the local region characteristics. Therefore, the adaptability to a specific test image may be weakened if the shape prior is over emphasized, while segmentations with improbable shapes due to weak image gradients along the true boundary may often occur if the appearance prior is over emphasized.

To incorporate the local region properties, Pohl et al. [22] separately estimated the appearance and spatial priors for each of the multiple objects by a hierarchical tree structure. Scherrer et al. [23] partitioned an entire image into regular cubics, and constructed local MRFs. The model parameters of each local MRF are estimated by local constraints based on neighboring MRF models. Although the methods model the properties of each multiple object, the construction of local priors according to the multiple parts of the single object is not considered. Kumar et al. [24] and Kohli et al. [25] suggested methods based on local shape prior. They divide the structured object into multiple parts, and train the variations of parts as a set of shape parameters by a layered pictorial structure model, and skeleton model. The local shape parameters restrict inappropriate variation, even on very ambiguous regions. However, the methods are not applicable to medical images, because many organs have an unstructured variation, and high computation is required to train all parts of the model in three-dimensional space. For use of the local prior in medical images, Lee et al. [26] proposed a method based on localized classifiers. Each localized classifier adaptively emphasizes a suitable prior, according to local region characteristics. Specifically, the appearance prior is emphasized in regions where different tissue can easily be distinguished by voxel intensities, while the shape prior is emphasized in regions where they cannot. To adaptively enforce the weight on each local region, the object is divided into multiple patches and segmentation is individually conducted on each patch. A drawback of this method is that local variations are not effectively considered because multiple priors of training images are averaged into one localized classifier on each local region. Furthermore, since segmentation consistency between neighboring local regions is not considered in a global context, the patch-wise segmentations are not smoothly aggregated.

In this paper, we propose a new hierarchical Markov random field (H-MRF) composed of local-level MRFs based on adaptive local priors which model local variations of shape and appearance and a global-level MRF enforcing consistency of the local-level MRFs. The proposed H-MRF can effectively deal with both local and global variations of shape and appearance unlike previous MRF based methods which only utilize the local prior [24–26] or the global prior [18–21]. Furthermore, the H-MRF framework is readily combined with well established MRF optimization techniques. Overall, the proposed method mitigates the adverse effect of large object variations and weak boundaries between objects and other tissues on segmentation accuracy to provide robust results. In addition to these strengths, the proposed H-MRF framework has several advantages. First, the H-MRF covers large local variations, such as pathological deformities, even though the training set may not contain subjects with similar deformities in similar positions. Since the multiple localized classifiers based on different characteristics create various candidates for the local variations, some of them are able to find the abnormal deformities. Second, construction of a complex training model is not required, because training set images are essentially just templates which are searched to find local patches with similar position and appearance that are directly applied as local cues. Therefore, the H-MRF can be applied given various data sets even with a small training set. Third, complex non-rigid registration techniques [6,7] with extensive computations are not required. Finally, since identical processes are repeated for local regions in our algorithm, the method can easily be optimized by various parallelization techniques, in order to reduce the computation time. Details of the H-MRF are described in Section 2.

To demonstrate its effectiveness, we applied the H-MRF to segment various organs from different types of medical images: bone and cartilage in knee magnetic resonance (MR) images, the liver in body computed tomography (CT) images, and the hippocampus in brain MR images. Details of the experiments are described in Section 3.

2. The proposed hierarchical MRF segmentation method

Although the same organ in different individuals may have a similar global shape, it is likely that the shape and appearance of each organ will differ significantly in several localized volumes. The main idea of the proposed method is to combine the model for local deformations and the model for global smoothness in a hierarchical structure. The local-level is composed of local interest volumes defined as three-dimensional volume patches centered on roughly initialized interest points. Prior information of local deformation is modeled by a set of similar local patches from the training data based on position and appearance. We refer to these patches as reference patches. For each local patch, multiple segmentation candidates are computed based on each reference patch as the segmentation prior. Our assumption is that at least one of these candidates will represent the correct local segmentation. The optimal ones among the segmentation