Computer Vision and Image Understanding 116 (2012) 146–165

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

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

Sequential model-based segmentation and recognition of image structures driven by visual features and spatial relations

Geoffroy Fouquier [a,*], Jamal Atif [b], Isabelle Bloch [a]

[a] Institut Telecom, Telecom ParisTech, CNRS LTCI, 46 rue Barrault, 75013 Paris, France
[b] TAO INRIA, CNRS, LRI – Paris-Sud University, 91405 Orsay Cedex, France

ARTICLE INFO

Article history:
Received 13 October 2010
Accepted 14 September 2011
Available online 2 October 2011

Keywords:
Segmentation
Knowledge-based system
Spatial relations
Graph representations
Fuzzy sets
Medical images
MRI

ABSTRACT

A sequential segmentation framework, where objects in an image are successively segmented, generally raises some questions about the "best" segmentation sequence to follow and/or how to avoid error propagation. In this work, we propose original approaches to answer these questions in the case where the objects to segment are represented by a model describing the spatial relations between objects. The process is guided by a criterion derived from visual attention, and more precisely from a saliency map, along with some spatial information to focus the attention. This criterion is used to optimize the segmentation sequence. Spatial knowledge is also used to ensure the consistency of the results and to allow backtracking on the segmentation order if needed. The proposed approach was applied for the segmentation of internal brain structures in magnetic resonance images. The results show the relevance of the optimization criteria and the interest of the backtracking procedure to guarantee good and consistent results.

1. Introduction

In this paper, we deal with segmentation and recognition of objects or structures in an image, based on a generic model of the scene. As a typical example, we focus on the recognition of internal brain structures in 3D magnetic resonance images (MRI), based on an anatomical model. More specifically, we address two important problems occurring in sequential approaches, as detailed below.

In Refs. [1,2], the authors introduced a new paradigm combining segmentation and recognition tasks. We will refer to this paradigm in the remainder of this paper as sequential segmentation and interpretation. It is defined as a knowledge-based object recognition approach where objects are segmented in a predefined order, starting from the simplest object to segment to the most difficult one. The segmentation and recognition of each object are then based on a generic model of the scene and rely on the previously recognized objects. This approach uses a graph which models the generic spatial information about the scene in an intuitive and explicit way (presented in [3]). This sequential segmentation framework allows decomposing the initial problem into several sub-problems easier to solve, using the generic knowledge about the scene. This approach differs from a regular divide-and-conquer approach since each sub-problem contributes to improve the resolution of the next sub-problems. It also avoids relying on an initial segmentation of the whole image.

This approach, as pointed out in Ref. [2], requires to define the order according to which the objects have to be recognized and the choice of the most appropriate order is one of the problems that remain open. It also lacks a step which could evaluate the quality of the segmentation of a particular object and detect errors to prevent their propagation.

In this paper, we propose original methods to answer these two open questions. Our contribution is twofold: first, we extend the sequential segmentation framework by introducing a pre-attentional mechanism, which is used, in combination with spatial relations, to derive a criterion for the optimization of the segmentation order. Secondly, we introduce criteria and a data structure which allow us to detect the potential errors and control the ordering strategy.

The pre-attentional mechanisms were defined in [4–6] to guide the focus of attention in modeling the visual system such as in feature integration theory. The sequential segmentation framework may be viewed as a way to focus attention on a small part of the scene and thus limit the search domain and the computational load. Among these mechanisms, we propose to use the notion of saliency to optimize the sequence of segmentation.

Our approach is applied to the segmentation and the recognition of internal brain structures in 3D magnetic resonance images. The intrinsic variability of these structures, the lack of clear boundaries and the insufficient radiometry make this

* Corresponding author.
E-mail addresses: geoffroy.fouquier@telecom-paristech.fr (G. Fouquier), atif@lri.fr (J. Atif), isabelle.bloch@telecom-paristech.fr (I. Bloch).

1077-3142/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2011.09.004

segmentation problem a difficult one. Some of the difficulties can be overcome by relying on generic knowledge about the human anatomy, that will be exploited to derive the model guiding the whole process.

This article is organized as follows. First we present in Section 2 a survey of knowledge based-approaches to the recognition of objects in a scene and provide an overview of the proposed approach. Section 3 presents the knowledge representation model. In Section 4 we propose to use some concepts of the visual attention to optimize the sequential segmentation framework. Then, the optimization of the sequential segmentation itself is described in Section 5 and the mechanisms for evaluating each structure segmentation in Section 6. Experiments on internal brain structure segmentation and results are presented in Section 7. Finally we draw some conclusions in Section 8.

2. Knowledge-based systems and spatial reasoning

The sequential segmentation framework of Colliot et al. [2] relies on a priori knowledge about the scene and uses intensively this knowledge at each step of the process. Thus, this framework may be described as a knowledge-based system using spatial relations. One can find a review of these systems in Refs. [7,8]. In this section, we focus on knowledge-based systems using spatial relations to describe the structure of the scene that have been applied to the recognition of brain structures in medical images.

Spatial relations play a crucial role in model-based image recognition and interpretation due to their stability compared to many other image appearance characteristics. They constitute structural information, which is particularly relevant when the intrinsic features of the objects are not sufficient to discriminate them.

2.1. Knowledge-based approaches for internal brain structures recognition

The difficulty of segmenting internal brain structures is due to the similarity between their gray levels, the lack of clear boundaries at some places and the partial volume effect. Their intrinsic features present a natural variability (in size and shape for example) between individuals, which is further increased in pathological cases. On the contrary, the spatial arrangement of these structures, i.e., their relative positions, is stable in healthy cases and even quite stable in pathological cases. For all these reasons, structural models of the internal brain structures have been used to segment and recognize the internal structures.

2.1.1. Structural model of the brain structures

One can find several anatomical descriptions of the brain, as atlas [9], nomenclature [10] or ontology [11]. These descriptions are often organized as a hierarchy of structures and provide descriptions of structures and relations between them. In Ref. [3], in collaboration with a neuro-anatomist, the internal brain structures are represented as a hierarchical graph where each vertex corresponds to an anatomical structure and each edge carries spatial relations between anatomical structures. This representation has been extended as the GRAFIP[1] [12] to include information about the structures composition, functional knowledge and about the pathologies.

2.1.2. Segmentation and recognition

Several classes of approaches for internal brain structures segmentation have been proposed in the literature. The first class of approaches uses a model graph and the image to segment is represented as a graph too. The segmentation and recognition process is then formalized as a graph matching problem [13]. The authors in Refs. [14,15] proposed to find a fuzzy morphism between a model graph built from a manual segmentation and an over-segmented image represented as a graph. Several optimization techniques have been proposed for this task [16,17]. Another approach was proposed in Ref. [18] and used an over-segmentation. The matching is viewed as a constraint satisfaction problem, with two levels of constraints and an ad-hoc algorithm. The authors recently extended their approach to cope with unexpected structures, such as tumors [19]. For this class of approaches, the initial graph is usually built from an over-segmentation of the image to segment, and the complexity of the method increases as the number of regions obtained from the over-segmentation grows.

In the second class of approaches, a sequential segmentation of the internal brain structures is performed, as proposed in Refs. [1,2]. In these approaches, the segmentation and the recognition are achieved at the same time. Each segmentation uses the spatial information encoded in the model, and more specifically the spatial relations to the already segmented structures. This information allows restricting the search domain around the structure. In these approaches, there is no initial segmentation of the image, but it raises some questions like the order of segmentation of the different objects or how to avoid the propagation of potential errors. Our approach belongs to this class and our contribution is an original answer to both questions.

The authors in Refs. [20,21] proposed a different type of approach, which is global and uses a constraint network. They proposed to link each anatomical structure with a region of space which satisfies all constraints in the network. Since it is hard to solve this problem directly, only the bounds of the domain of each variable (i.e. structure to be segmented) are modified by the process and sequentially reduced using specifically designed propagators derived from the spatial constraints. Finally, a segmentation is extracted using a minimal surface algorithm. This approach provides good results and does not need an initial segmentation either. However, due to the number of constraints, it is quite complex and the computation time is high, especially in 3D.

2.2. Proposed framework

We propose to extend the sequential segmentation framework proposed in Ref. [2], where structures are sequentially segmented from the easiest to segment to the most difficult ones. Each structure segmentation uses the information provided by the previous segmentations. Our extension aims at answering the following questions raised by this framework: "in which order should the objects of the scene be segmented?" and "how to assess the segmentation result in order to detect potential errors and avoid their propagation?".

The proposed framework has two levels, as depicted in Fig. 1. The first level is a generic bottom-up module which allows selecting the next structure to segment. This level does not rely on an initial segmentation or classification, but instead on a focus of attention and a map of generic features described in Section 4. The sequential approach allows this level to use two types of knowledge: generic and domain independent features in unexplored area of the image to segment, and high-level knowledge such as spatial relations linked to the already recognized structures. We propose to answer the first question by deriving a selection criterion from a pre-attentional mechanism: a saliency map. This criterion is used to optimize the segmentation order and to select the next structure to segment at each step.

The second level achieves recognition and segmentation of the selected structure, as well as the evaluation of the segmentation. The recognition of the structure is achieved at the same time as

[1] For "Graph of Representation of Anatomical and Functional data for Individual patients including Pathologies".