Computer Vision and Image Understanding 117 (2013) 145–157

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

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

A multiple object geometric deformable model for image segmentation ☆

John A. Bogovic [a,*], Jerry L. Prince [a], Pierre-Louis Bazin [b]

[a] Johns Hopkins University, Baltimore, MD, USA
[b] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

ARTICLE INFO

Article history:
Received 30 June 2011
Accepted 19 October 2012
Available online 29 October 2012

Keywords:
Multiple object segmentation
Geometric deformable model
Level sets
Topology preservation

ABSTRACT

Deformable models are widely used for image segmentation, most commonly to find single objects within an image. Although several methods have been proposed to segment multiple objects using deformable models, substantial limitations in their utility remain. This paper presents a multiple object segmentation method using a novel and efficient object representation for both two and three dimensions. The new framework guarantees object relationships and topology, prevents overlaps and gaps, enables boundary-specific speeds, and has a computationally efficient evolution scheme that is largely independent of the number of objects. Maintaining object relationships and straightforward use of object-specific and boundary-specific smoothing and advection forces enables the segmentation of objects with multiple compartments, a critical capability in the parcellation of organs in medical imaging. Comparing the new framework with previous approaches shows its superior performance and scalability.

1. Introduction

Image segmentation is one of the most fundamental problems in computer vision, with applications in scene reconstruction, motion tracking, content-based image retrieval, aerial imaging, etc. (see Fig. 1). Medical image analysis in particular has a growing need for the automatic segmentation of multiple organs and complex sub-structures from increasingly large data sets of multi-dimensional images [1–3]. The segmentation step often directly affects subsequent processing tasks such as quantification, registration, and visualization. In many cases, the segmentation of multiple objects should provide a complete parcellation of the image into components without overlaps and gaps between the segmented regions. When multiple regions meet, the boundary around a region often becomes heterogeneous in nature, e.g., where the image data provides little information, boundaries may need to be inferred based on a prior model, while image information can be relied upon elsewhere [2,4]. Prior knowledge may also extend to the overall organization of the structures of interest and their spatial or topological relationships [1,5,6]. Finally, segmentation problems can easily involve large numbers of components or objects [7,8], and the complexity of the applied methods must be taken into account for practical reasons.

Parametric deformable models (PDMs) – i.e., active contours implemented by explicitly tracking points – have been widely used in computer vision to perform image segmentation [9]. An important property of this representation is its capability to represent boundaries at a sub-grid resolution as it is essential in the segmentation of thin structures (e.g., cortical sulci). Image-based "external forces" drive the contour toward desired features while contour-dependent "internal forces" regularize and smooth the boundary. Geometric deformable models (GDMs) – i.e., active contours implemented with level sets [10,11] – permit flexible topological changes and yield contours with no self-intersections. In the GDM framework, "speed" functions describe the local movement of the contour and are analogs to the forces used in PDMs. With a single level set function, GDMs permit the segmentation of multiple isolated regions; but in their most basic implementation, they do not control the number of objects or their topology. Topology-preserving extensions [12–14] permit control of single object topology, but do not address topological relationships between objects or permit one to model boundaries between multiple objects at once.

A number of multiple object segmentation methods based on the level set framework have been proposed [15–24]. Most of these approaches use N level set functions to segment N objects and rely on coupling terms to avoid overlaps and gaps [18–20,23,25]. These methods have the advantage that each object can be independently specified in both its own topology and its internal and external speeds. However, coupling terms do not forbid certain object interactions, so these approaches can still produce overlaps and gaps in practice. As well, most are not formulated to consider the relationships between objects, and memory requirements become daunting as the number of objects to be segmented grows.

Vese and Chan [15] introduced the *multiphase* (MP) segmentation framework that represents N objects with $\log_2(N)$ level sets

☆ This paper has been recommended for acceptance by Nikos Paragios.
* Corresponding author.
E-mail address: bogovic@jhu.edu (J.A. Bogovic).

1077-3142/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cviu.2012.10.006

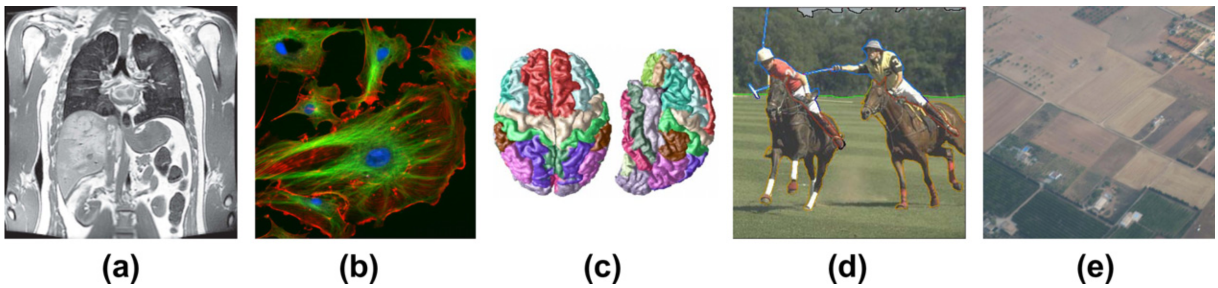

Fig. 1. Various problems where images are to be segmented into multiple interacting objects: (a) an MRI of the abdomen, showing many organs; (b) fluorescent microscopy imaging involving complex interactions of multiple cells; (c) a parcellation of the cortex into 78 gyral regions; (c) images and videos of sporting events where the different players interact; (d) aerial images of crops and farmlands. These examples were obtained from computer vision and medical imaging databases [3,26,27] or our own work (c) [28].

based on combination rules. This model permits multiple object boundaries, and guarantees no overlaps or gaps. As well, it substantially reduces the computational burden as the number of objects or compartments grows. However, this approach has three key limitations. First, its image-based external speed term is not easily generalized beyond the region-based model of Mumford and Shah [29]. This fact excludes a rich collection of external speeds that may be essential to solving many problems of practical importance. Second, the internal speeds in the multiphase framework – comprising a penalty on contour length – are applied to the level set functions rather than to the objects themselves. Thus, it is possible that while the lengths of the level set functions are minimized, it may not be so for the boundaries of the objects themselves. Third, the evolution/optimization can "get stuck" in situations where a pixel needs to acquire a label that can be reached only by changing two level set functions at the same time. The existing evolution strategy cannot resolve these situations, which are also more commonly found with an increasing number of objects. A remedy for this limitation involving a permutation of the level set combination rule used to represent a given label was proposed in [30]. As the number of level set functions increases, however, a greater number of permutations will need to be performed to ensure that transitions between all labels are possible.

Pohl et al. [21] proposed a probabilistic embedding to avoid overlaps and gaps by replacing the level set isocontours with a labeling of regions according to their maximum probability. In this case, however, the geometric properties of curvature associated with level set isocontours are no longer relevant, and the method still requires $N-1$ level sets or $N-1$ functions derived from level sets.

Brox and Weickert [24] presented a coupled curve evolution method where distinct objects are constrained through the coupling of the evolution equations rather than by changing the energy functional. Their coupled curve evolution shows that a pixel (or voxel) in one object competes with other objects, while additional terms help to discourage gaps. This approach deals with multiple objects, but it does not guarantee there are no overlaps or gaps, making relationships and other topology constraints difficult to enforce. The evolution requires N level set functions for N objects, increasing computational and storage burden as more objects are added.

Recent methods by Lie et al. [31] and Chung and Vese [32] formulate a "multilayer" method that represents multiple objects using a small number of nested level contours of a function. This approach is efficient with memory, requiring just two functions to represent triple junctions in 2D. However, it shares with the multi-phase approach a limitation in the types of speeds that may be applied, a lack of control of topology, and the interpretation of its regularization terms as minimizing level set length rather than object boundary length [15]. These representations also lose some of the computational advantages of using signed distance functions.

Much interesting work has gone into adapting the level set formalism for multiple object segmentation to account for prior shape information. Tsai et al. [22] developed a framework that constrains potential segmentation results using a parametric shape model based on principal component analysis. Uzunbas et al. [33] employed a similar framework, but built a statistical shape model using kernel techniques and furthermore modeled relative poses between objects. These methods used N level sets to represent N objects, and therefore, a heuristic approach was used in both to prevent overlaps. Vazquez-Reina et al. [34] used a shape model similar to [22], but used the MP level set representation of [15], rather than N level sets. This prevents overlap and gaps, and improves efficiency, but still suffers from some of the setbacks of the multiphase representation. Fussenegger et al. [35] extended [24] with a multi-object pose-invariant shape prior. These methods have been important contributions in constraining multiple object level set segmentation. However, none of these methods guarantee that single object topology, or topological relationships are preserved. The storage of N (or $\log_2(N)$) level sets along with the shape priors may become burdensome as the number of objects increases.

Markov random fields (MRFs) are graphical models that have achieved great success in image segmentation. Classical algorithms such as iterated conditional modes (ICM) [36] can provide approximate solutions in the multiple label problem. Other methods have been proposed that efficiently and accurately solve the Potts model segmentation problem using advanced optimization techniques. Zach et al. [37] presented a method that efficiently solves a continuous relaxation of the Potts model. Lellmann et al. [38] solved a similar formulation using an operator splitting optimization framework. Bae et al. [39] performed optimization using a dual formulation. While these methods all perform very well for a variety of segmentation tasks, as presented they lack certain important capabilities that aid in segmenting specific objects in images rather than identifying image regions with similar features, (e.g., intensity). In particular, these methods, as proposed, might segment "high intensity" rather than "femur" in computed tomography. This specific identification is important in medical imaging, for example, where quantitative measurements of anatomy are often used for diagnosis or research purposes. High level information such as object topology, statistical priors, as well as partial volume functions have been instrumental in advancing this area of image segmentation. Furthermore, graph-based methods generally lack the sub-grid resolution offered by deformable models. The presentation of the above