Pattern Recognition 48 (2015) 2513–2529

Contents lists available at ScienceDirect

Pattern Recognition

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

Inhomogeneity-embedded active contour for natural image segmentation

Lingzheng Dai, Jundi Ding*, Jian Yang

Department of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, PR China

ARTICLE INFO

Article history:
Received 21 July 2014
Received in revised form
28 December 2014
Accepted 1 March 2015
Available online 11 March 2015

Keywords:
Active contour
Level set
Natural image segmentation
Intensity inhomogeneity
Saliency-inspired initial contour

ABSTRACT

Active contour model (ACM) is one of the popular methodologies for image segmentation. However, the ACMs developed so far have not shown powerful performance on natural images. The reason is that natural images are rich in color, intensity or texture. The object pixels are often not artifact inhomogeneous, but inherently inhomogeneous. In this paper, we propose an inhomogeneity-embedded active contour (InH_ACM) for natural image segmentation. InH_ACM describes the inhomogeneity in natural images by a pixel inhomogeneity factor and utilizes it for segmentation, unlike most of existing methods that use some averaging convolution to reduce or remove the inhomogeneity in images. Moreover, we build a saliency-inspired framework that can automatically locate the initial contour for InH_ACM to start the evolution. Experimental results on Alpert's 100 gray images, MSRA's 1000 color images and our collected 300 images where the contained objects are mostly intrinsic inhomogeneous indicate that our proposed InH_ACM can produce reliably satisfactory segmentation in many situations, outperforming most of current popular ACMs.

1. Introduction

In recent decades, active contour models (ACMs) have been one of the most widely studied methods for object segmentation [6,15,17,18,21]. The basic idea is to minimize an energy function so as to allow a contour curve to deform under the theory of surface evolution. Ideally, the desired result is that the deformed curve right stops on the true boundary of the object.

However, existing ACMs whether edge-based [5,8,9,20] or region-based [10,12–14,16,22] seem not able to yield the desired object boundary curve in many situations. For example, edge-based ACMs that utilize the image gradient to construct some edge oriented constraint into the energy are highly sensitive to image gradient information. The evolving contour usually stops at the pixels with the maximal gradients. So, they often work poorly for the segmentation of high-textured objects (like zebras) with many edges and the objects low-contrasted to the background.

Region-based ACMs that use the statistical information of regions inside and outside the evolving contour are free of image gradients. As reported in [14], they can extract the objects with weak edges in some situations. Even, some models, e.g., the piecewise constant (PC) models (also known by Chan–Vese models), are sometimes less sensitive to the initial placement of evolving contour. One problem is that both the regions inside and outside the contour must be as homogeneous as possible. That is, these regions should be homogeneous enough with few variations in intensity, color or any other low-level attribute, i.e., being constant. This is a globally restricted assumption. However this assumption does not often hold true. The region-based models like PC [14] often fail in the cases where the regions are with inhomogeneity.

To address the issue of inhomogeneity (abbreviated as InH), Vese and Chan [15] and Tsai et al. [16] independently suggest two similar variants of the known PC model. They – known as piecewise smooth (PS) models – are to find an optimal approximation for each region by some piecewise smooth function, instead of a global constant in PC models. However, their computation of smooth function for each region is time-consuming. For this, Li et al. [18,1] put forward an energy of local binary fitting (LBF) that exploits the spatial relation between pixels by involving a Gaussian kernel function. By minimizing this energy, the fitting value of each pixel is efficiently computed as the average intensity of its local neighboring pixels, not all pixels in the whole image. The idea is that the faraway neighbors of one pixel in general have few or almost no influence on the fitting value of that pixel. With this idea, Zhang et al. [2] go further on, to come up with a local image fitting (LIF) energy that minimizes a difference between the original image and its 'LBF'-image. That is, LIF aims to look for a LBF-image that best fits the original image. Besides, Wang et al. in [19] are inspired to propose a local Chan–Vese (LCV) model that goes to utilize the Gaussian convolution of the original image to describe the local information.

These models have indeed exhibited a certain capability of handing InH. The local convolutions embedded in these models

---

* Corresponding author.

E-mail addresses: lingzhengdai@163.com (L. Dai), dingjundi2010@njust.edu.cn (J. Ding), csjyang@njust.edu.cn (J. Yang).

http://dx.doi.org/10.1016/j.patcog.2015.03.001
0031-3203/

acting like some operations of smoothing can 'correct' the 'artificially' blurred features of pixels to some extent, and thereby could clear the intensity (or color) contrast between neighboring pixels (if they have). However, they seem to produce good results only in such cases where the InH is slowly varying. When the InH is randomly or sharply varying, e.g., from white to black like zebra stripes, they still tend to yield erroneous segmentation.

Recently, a number of texture-embedded ACMs have been built to tackle such inherent InH problems [37,38,40,41]. Their idea is that such InH mainly arises from the texture in our living natural world. The involved photographs describing our natural world are usually called natural images. However, natural textures are far richer, and far more random than one could imagine. Gabor filters [40,41], structure tensors [42], or other existing texture descriptors [37] seem all not powerful enough to describe so rich natural textures. Besides, quite often, there may be kinds of 'artificial' InH appearing in natural images due to some imperfect imagery factors. For these reasons, ACMs with texture descriptors have still been seen to report satisfactory results over many natural images.

In fact, whichever InH the image encounters, it is just because neighboring pixels are not homogeneous in low-level cues. That is, there is spatial variation between neighboring pixels. The difference is that given a certain threshold, the variation between pixels may be above that threshold (i.e., with a large contrast) or below that threshold (i.e., with a small contrast). Intuitively, for one pixel p, if the majority of its neighbors have a large (small) contrast with itself – we could say – p exhibits a high (low) degree of InH with its neighbors. For example, pixels from the black-and-white stripe of one zebra must show a higher degree of InH with respect to each other, while pixels on the piece-wise region show a low degree of InH. By this motivation, in this paper we incorporate a pixel inhomogeneity factor (PIF) proposed in [33,43] to depict the InH of each pixel with respect to its local neighbors.

Further, we observe that pixels from one object tend to have a consistent degree of InH ('Local Information' in Fig. 1). It is clear that pixels on the cheetah disclose a consistently higher degree of InH (in white), pixels in the grass background have a relatively low degree of InH (in black). Inspired by this, we propose an inhomogeneity-embedded active contour (InH_ACM) for segmenting the objects from a variety of natural images reliably. To our knowledge, we are the first to attempt this. No previous ACMs have tried to depict InH and embed InH into the curve evolution. Instead, just as analyzed above, they seem all to 'eliminate' or 'correct' the InH in images [39]. So they have been more seen to be tested on the synthetic or medical images.

Fig. 1 gives a general schematic framework of InH_ACM, which utilizes the global color (or intensity) information but also exploits the local InH information. The overall energy E is mainly composed of two parts: a coherence energy E_G over pixel global color (or intensity), and a consistency energy E_L on pixel local InH. Then, we represent the contour by the zero level set of a Lipschitz function (i.e., a level set function [4,11] used in many existing ACMs [14,18]), and we can easily arrive at one minimum solution to InH_ACM in a variational level set formulation. The solution is often just as expected to right enclose the object(s) to be segmented in images, as shown in Fig. 1.

One point is that like most existing ACMs [14,18,3,2], InH_ACM also needs a manually initialized contour to start the evolution and the energy minimizing. Different initial contours also usually do not lead to the same result, since the numerical implementations focus only on the narrow-band yielding a method sensitive to initialization. However, our model embedded with the InH always can yield an effective segmentation when the initial contour is placed near the true boundary of the object(s) in images, as shown in Fig. 2. Manually initializing the contour for each image is tedious, time-consuming and impractical especially in view of image databases of increasingly larger sizes. Observing that the objects to be segmented in images are always visually salient [23–31,34–36], we here suggest a saliency-inspired way to tackle this initialization problem. This way can automatically figure out an initial contour around the true object boundary (see Fig. 2d).

At last, to evaluate the proposed method, we conduct the experiments on a wide variety of natural images from Alpert's database of 100 gray images [32], MSRA database of 1000 color images [27], and the one compiled by us. Our model can produce the reliably satisfactory segmentation in many situations, outperforming most of current popular ACMs.

The main contributions of this paper are follows. (1) A novel ACM (InH_ACM) is proposed. InH_ACM describes the inhomogeneity in natural images by a pixel inhomogeneity factor and utilizes it for segmentation. It can produce the promising segmentations on natural images. (2) We present a saliency-inspired framework for initializing the contour for ACMs. Compared with the tradition marking the general initial contour by the user, the proposed framework can locate it free of any user manual effort. (3) We compiled a natural image database which is especially made up of 300 images where the contained objects are mostly intrinsic inhomogeneous. It exhibits a great difficulty for the segmentation methods developed so far including most ACMs.

The remainder of this paper is organized as follows: Section 2 briefly reviews the well-known PC and LBF model. Section 3 is to detail our InH_ACM as well as all related concepts and analysis, while the minimization of InH_ACM is described in Section 4. Section 5 presents the saliency-inspired framework to automatically locate the

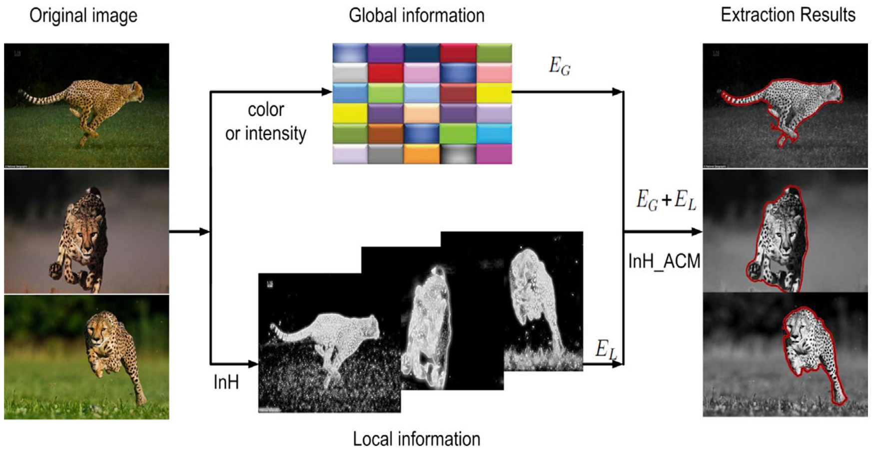

Fig. 1. The framework of our InH_ACM.