Image Segmentation with Active Contours based on Selective Visual Attention

E. MENDI¹, M. MILANOVA²
Department of Computer Science, University of Arkansas at Little Rock, Little Rock, AR, 72204 USA
¹esmendi@ualr.edu; ²mgmilanova@ualr.edu

Abstract: Telemedicine is growing and there is an increased demand for faster image processing and transmitting diagnostic medical images. Identifying and extracting the region of interest (ROI) accurately is an important step before coding and compressing the image data for efficient transmission or storage. The usual approach to extract ROI is to apply contour segmentation method. Chan-Vese active contour model [1] is a well-known image segmentation technique based on Mumford-Shah level set methods. Selective visual attention is a fundamental component of perceptual representation in a visual system. It influences the identification of a stimulus from those that operate after perception is complete. The SaliencyToolbox [2] is a collection of Matlab functions and scripts for computing the saliency map for an image, for determining the extent of a proto-object, and for serially scanning the image with the focus of visual attention. The implementation of the toolbox is extension of the saliency map-based model of bottom-up attention [3], by a process of inferring the extent of a proto-object at the attended location from the maps that are used to compute the saliency map. In this paper, we focus on extracting ROI by segmentation based on visual attended locations. Chan-Vese active contour model is used for image segmentation and attended locations are determined by SaliencyToolbox. Finally, we successfully segmented two attended locations of a medical image.

Key-Words: Active Contours, Selective Visual Attention, Image Segmentation.

1 Introduction

Image segmentation is an important problem in image analysis which is to partition a given image into disjoint regions. There are a wide variety of approaches to the segmentation problem. One of the popular approaches is active contour models, also called snakes. The basic idea is to start with a curve around the object to be detected, the curve moves towards an “optimal” position and shape by minimizing its own energy. Based on the Mumford-Shah functional [4] [5] [6] for segmentation, Chan and Vese [1] proposed a new level set model for active contours to detect objects whose boundaries are not necessarily defined by a gradient.

Visual attention [7] is the process of selecting and gating visual information based on saliency in the image itself (bottom-up), and on prior knowledge about scenes, objects and their interrelations (top-down) Visual attention addresses both problems by selectively enhancing perception at the attended location, and by successively shifting the focus of attention to multiple locations. It is also important for selecting the interest things from the input information and [8] provides the brain with a mechanism of focusing computational resources on one object at a time, either driven by low-level image properties (bottom-up attention) or based on a specific task (top-down attention). Moving the focus of attention to locations one by one enables sequential recognition of objects at these locations.

In this paper, we extracted ROI with active contours based on selective visual attention. Chan-Vese active contour model is used for image segmentation and attended locations are determined by SaliencyToolbox which is extension of the saliency map-based model of bottom-up attention [3], by a process of inferring the extent of a proto-object at the attended location from the maps that are used to compute the saliency map. The paper is organized as follows: In Section 2, we provide an overview of the Chan-Vese model. Section 3 presents the bottom-up salient region selection model. In Section 4, we present application of our approach. Finally, in Section 5 the conclusions of this paper are summarized.
2 Chan-Vese Model

The Mumford-Shah model [4] [5] [6] is a variational problem for approximating a given image by a piecewise smooth image of minimal complexity. Let \( u \) is differentiable on \( R \) and allowed to be discontinuous across \( C \), Mumford-Shah energy functional is as follows:

\[
F(\phi, C) = \frac{1}{\sigma^2} \int_R (\phi - f)^2 \, dx + \int_{\partial C} [\nabla \phi]^2 \, dx + \lambda |C|
\]

(1)

where \( R \) is the image domain, \( f \) is the feature intensity, \( C \) is the curve, \( \phi \) is the smoothed image, \(|C|\) is the arc length of \( C \) and \( \sigma, \lambda \) are positive parameters. Segmentation problem is restated as finding optimal approximations of \( g \) by piece-wise smooth functions \( u \), whose restrictions to the regions are differentiable.

The Chan-Vese model [1] is a special case of the Mumford Shah model by restricting (1) to piece-wise constant functions \( \phi \) and looking for the best approximation \( \phi \) of \( f \) taking only two values. Then the energy functional in (1) is expressed in terms of the level set function by replacing the \( \frac{1}{\sigma^2} \) by \( \frac{1}{\sigma} \) and the corresponding Euler-Lagrange equation for \( \phi \), using gradient descent in artificial time leads to:

\[
\frac{\partial \phi}{\partial t} = \nabla \phi \left[ \kappa(\phi) - \lambda \left( (c_1 - f)^2 - (c_2 - f)^2 \right) \right]
\]

(5)

where \( \kappa(\phi) \) is the curvature of the level sets and \( \lambda = \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \). A multigrid scheme on the discretized Euler-Lagrange equation (5) is used for the minimization of Chan-Vese energy functional.

\[
\inf_{\phi, c_1, c_2} F(\phi, c_1, c_2)
\]

(6)

where

\[
\min_{\phi} F(\phi, c_1, c_2) = \int_D [\nabla \phi + \lambda [(c_1 - f)^2 + (c_2 - f)^2]]^2 \, dx
\]

(7)

The explicit formula provided by (5) is solved by using gradient descent procedure as described in [9].

3 Bottom-Up Salient Region Selection Model

The model of bottom-up salient region selection presented by [2] [8] based on the model of saliency-based bottom-up attention by Itti-Koch [10] [11] is implemented as part of the SaliencyToolbox for Matlab (http://www.saliencytoolbox.net). This model introduces a process of inferring the extent of a proto-object at the attended location from the maps that are used to compute the saliency map.

Itti-Koch model [10] [11] which is for bottom-up selective visual attention based on serially scanning a saliency map that is computed from local feature contrasts, for salient locations in the order of decreasing saliency (Fig. 1). Presented with a manually preprocessed input image, their model replicates human viewing behavior for artificial and natural scenes.
Fig. 1: General architecture of Itti-Koch model [9]

Visual input [9] is first decomposed into a set of topographic feature maps. Different spatial locations then compete for saliency within each map, such that only locations which locally stand out from their surround can persist. All feature maps feed, in a purely bottom-up manner, into a master saliency map. The purpose of the saliency map is to represent the saliency at every location in the visual field by a scalar quantity and to guide the selection of attended locations, based on the spatial distribution of saliency. However this model’s usefulness [12] as a front-end for object recognition is limited by the fact that its output is merely a pair of coordinates in the image corresponding to the most salient location.

This model is extended [2] [8] by a process of inferring the extent of a proto-object, contiguous region of high activity in feature map, at the attended location from the maps that are used to compute the saliency map. This is achieved by introducing feedback connections in the saliency computation hierarchy in order to estimate the proto-object region based on the maps and salient locations computed in Itti-Koch model [10] [11].

Different visual features that contribute to attentive selection are combined into one single topographically oriented saliency map which integrates the normalized information from the individual feature maps into one global measure of conspicuity. The locations [2] in the saliency map compete for the highest saliency value by means of a winner take-all (WTA) networks of integrate-and-fire neurons. The winning of this process is attended to, and the saliency map is inhibited. Continuing WTA competition produces the second most salient location, which is attended to subsequently and then inhibited, thus allowing the model to simulate a scan path over the image in the order of decreasing saliency of the attended locations.

4 Experimental Results

In this section, we present image segmentation of two attended location of a medical image. In Fig. 2, conspicuity and saliency maps are depicted. Saliency map is summed by conspicuity maps with information of color, intensity and orientation. Fig. 3 shows WTAs, corresponding attended locations and active contours. Conspicuity maps, saliency map, WTAs and attended locations are operated by SaliencyToolbox [2]. Attended locations are set as initial contours to be segmented by using Chan-Vese Model [1]. First attended location took approximately 90 ms and second attended location took 172 ms simulated time.

For the telemedicine application we have integrated image segmentation with adaptive compression technique. The proposed compression technique is based on the hypothesis that image resolution exponentially decreases from the fovea to the retina periphery. This hypothesis can be represented computationally with different resolutions. The visual attention points may be considered as the most highlighted areas of the Visual Attention model. These points are the most salient regions in the image. When going further from these points of attention, the resolution of the other areas dramatically decrease. Different authors work with different filters and different kernel size to mimic this perceptual behavior [13]. These models ignore contextual information representation.

When the set of regions of interest is selected, these regions need to be represented with the highest quality while the remaining parts of the processed image could be represented with a lower quality. In result, higher compression is obtained. The adaptive compression technique proposed is based on new image decomposition called Inverse Difference Pyramid (IDP) [14].

5 Conclusions

We have proposed a novel markerless approach for medical image segmentation by combining saliency attention maps with active contours. We have implemented the Chan-Vese active contour model [1] by setting attended locations as initial contours. Attended locations are extracted by
SaliencyToolbox [2]. It is anticipated that this process will be useful for identifying and extracting the ROI accurately. The combination of the two techniques minimizes user interaction and speeds up the entire segmentation process. The method has been successfully tested on medical images and the intimal medial thickness of the arteries is extracted as a region of interest.

Acknowledgement:
This paper was supported by NIH Grant # P20 RR-16460 from the IDeA Networks of Biomedical Research Excellence (INBRE) Program of the National Center for Research Resources.

References:
Fig. 2: a) Input image, b) conspicuity map for color contrast, c) conspicuity map for intensity contrast, d) conspicuity map for orientation contrast, e) saliency map combined by conspicuity maps
Fig. 3: a) WTA map for the first attended location, b) WTA map for the second attended location, c) first attended location, d) active contours based on first attended location, e) second attended location, f) active contours based first two attended locations