Interactive Image Segmentation Framework Based On Control Theory

Liangjia Zhu[a], **Ivan Kolesov**[a], **Peter Karasev**[b], and **Allen Tannenbaum**[a]

[a]Stony Brook University, Stony Brook, New York, USA

[b]Atlanta Agilent Technologies, Atalanta, Georgia, USA

Abstract

Segmentation of anatomical structures in medical imagery is a key step in a variety of clinical applications. Designing a generic, automated method that works for various structures and imaging modalities is a daunting task. Instead of proposing a new specific segmentation algorithm, in this paper, we present a general design principle on how to integrate user interactions from the perspective of control theory. In this formulation, Lyapunov stability analysis is employed to design and analyze an interactive segmentation system. The effectiveness and robustness of the proposed method are demonstrated.

Keywords

Interactive Image Segmentation; Control Theory; Lyapunov Stability; Active Contours

1. DESCRIPTION OF PURPOSE

The purpose of this research is to present a generic principle on how to design interactive segmentation system from the feedback control perspective.

2. METHODS

The overall of the proposed framework in shown in Figure 2. It can be regarded as a dual control system: in the top level, user adaptively apply inputs to guide the system; in the lower level: the dynamic system reacts accordingly.

Without the loss of generality, the level-sets formulation is used here. Let $I:\Omega \to \mathbb{R}^m$ be an image defined on $\Omega \in \mathbb{R}^n$, where m 2 and n 2. Suppose an image to be segmented consists of N regions. Each region $\Omega_i(x,t)$ is associated with a level set function $\varphi_i(x,t)$ and is moving from an initial state $\varphi_i(x,0) = \Omega_i(x,0)$. The Heaviside function, denoted by $H(\varphi(x))$, is used to indicate the exterior and interior regions and its derivative is denoted by $\delta(\varphi(x))$.

Suppose the user has an ideal segmentation $\{\phi_i^*(x)\}$, $i = 1, \cdots, N$ in mind. Then, our goal is to design a feedback control system

Further author information: (Send correspondence to Liangjia Zhu), Liangjia Zhu: liangjia.zhu@stonybrook.edu.

$$\begin{cases} \frac{\partial \phi_i(\boldsymbol{x},t)}{\partial t} = (G\left(\phi_i\left(\boldsymbol{x},t\right),I\right) + F\left(\phi_i\left(\boldsymbol{x},t\right)\phi_i^*\left(\boldsymbol{x}\right)\right)) \delta\left(\phi_i\left(\boldsymbol{x},t\right)\right) \\ \phi_i\left(\boldsymbol{x},0\right) = \phi_i^0\left(\boldsymbol{x}\right) \end{cases} \quad (1)$$

such that $\lim_{t\to\infty} \varphi(\boldsymbol{x},t) \to \varphi^*(\boldsymbol{x})$ for $i = 1, \cdots, N$, where $G(\varphi_i(\boldsymbol{x}),I)$ is a generic term representing an image force and the interactions between φ_i and other $\varphi_j, j \neq i$, and $F\left(\phi_i\left(\boldsymbol{x},t\right)\phi_i^*\left(\boldsymbol{x}\right)\right)$ is the control signal needs to be determined. Without the loss of generality, we can decompose $G(\varphi_i(\boldsymbol{x}), I)$ into two *competing* components as

$$G\left(\phi_i\left(\boldsymbol{x}\right),I\right) = -\left(g\left(\phi_i\left(\boldsymbol{x}\right),I\right) - g^c\left(\phi_i\left(\boldsymbol{x}\right),I\right)\right), \quad (2)$$

where $g(,) \geq 0$ represents an internal image energy of $\varphi_i(\boldsymbol{x})$ and $g^c(\varphi_i(\boldsymbol{x}), I)) \geq 0$ is the image energy from all other $\varphi_j(\boldsymbol{x}), j \neq i$. This way of decomposition has been used for modeling multiple active contours in different region-based algorithms.[1–3] On the other hand, since distance information from given points in an image can be implemented using the level-sets formulation,[4] segmentation algorithms that are based on clustering pixels according to the minimal distance to given seeds naturally fit into this formulation. That is, the presented framework works for: 1) *region-based active contour models* and 2) *distance-based clustering*.

Following the derivation in,[5] we have the following theorem for the dynamical system defined in equation (1),

Theorem 1

The control law

$$F\left(\phi_i\left(\boldsymbol{x},t\right)\right) = -\alpha_i^2\left(\boldsymbol{x},t\right)\xi_i\left(\boldsymbol{x},t\right), \quad (3)$$

where $\alpha_i^2\left(\boldsymbol{x}\right) \geq \frac{g_M(\boldsymbol{x})}{\rho}$, stabilizes the system (1) from $\{\varphi_i(\boldsymbol{x},t)\}$ to $\{\phi_i^*\left(\boldsymbol{x}\right)\}$, $i = 1, \cdots, N$, provided that

$$\rho\int_\Omega \delta^2\left(\phi_i\left(\boldsymbol{x},t\right)\right)\left|\xi_i\left(\boldsymbol{x},t\right)\right|d\boldsymbol{x} \leq \int_\Omega \delta^2\left(\phi_i\left(\boldsymbol{x},t\right)\right)\xi_i^2\left(\boldsymbol{x},t\right)d\boldsymbol{x}, \quad (4)$$

in which ρ is a scale parameter. Here, $\xi_i(\boldsymbol{x},t)$ is the point-wise total error for the ith region, and $g_M(\boldsymbol{x})$ is the bounds of the image depend term $g(,)$ defined in the previous section.

3. RESULTS

Two orthopedic images were used to quantitatively compare the presented methods with the popular GrabCut algorithm.[6] The structures being segmented, the epiphysis and physis, are shown in Figure 2. User input via mouse click-and-drag was implemented and measured identically for all algorithms. A location through which the cursor was dragged is defined as an "actuated voxel"; the extents around the cursor that mark seed regions in GrabCut are not counted towards this total. Locations in the image whose assigned label changes between background and foreground are tracked over time and are referred to as "reclassified" voxels.

The total number of actuated voxels needed to complete the segmentation is presented in Figure 3. It shows that both the region- and distance-based interactive segmentation methods require less user input than the Grabcut in segmenting these two structures. Segmenting the physis is more difficult with GrabCut due to the elongated shape, the nearly identically-looking fluid around the bone and the bimodal appearance of cortical bone above and spongy bone below the physis. A GrabCut iteration can change the segmentation dramatically; when this change is erroneous, significant corrective effort becomes required. In Figure 3, we see this manifested by the large increases in actuated voxels during the first few rounds of GrabCut user input. In contrast, the proposed algorithms provide rapid continuous visual feedback for the user; small corrections are made before a large error can develop.

Predictability of how the segmentation changes in response to mouse strokes is a criterion for practical ease of use. Two scatterplots quantify the predictability in Figure 4; dynamic response is characterized in terms of the number of reclassified voxels (Y -axis) and the number of newly actuated voxels (X- axis). Each mark corresponds to one iteration when new user input was applied. Linear regression lines are overlaid on the data. All algorithms have a similar dynamic response in the epiphysis segmentation in Figure 4(a). Two issues become apparent for the juvenile physis segmentation. First, the distribution of GrabCut data points is quite broad; Second, some of the GrabCut data points are below the dashed green line, indicating a waste of user effort since there are more voxels actuated than reclassified. The dynamic response of GrabCut makes it hard for a user to predict how much change new mouse strokes will cause.

4. NEW OR BREAKTHROUGH WORK TO BE PRESENTED

The image segmentation has been an active research field over the past several decades and remains to be a very challenging task. It is even frustrating in certain segmentations that human being can recognize and extract target objects instantly, while it is still hard for computers to accomplish satisfactory results automatically. How to effectively integrate human's prior knowledge into a segmentation design has become a basic principle underlying numerous types of existing stat-of-the-art segmentation methods.

However, to the best of our knowledge, there are only very few attempts that model interactive segmentation process in a systematical way. In our previous work,[5] the authors formulate the interactive image segmentation in a feedback control framework based on single-object region-based active contour models. In this work, we present the generalization of the work to more generic cases, which seamlessly handles both region- and distance-based criteria for multi-object image segmentation design.

5. CONCLUSIONS

This paper has presented a systematical way of applying control theory to analyze and design an interactive medical image segmentation system. Preliminary results show the effectiveness and robustness of the proposed method. Though the examples used in this paper are based on level-sets formulation, the design principle is generalizable to other

interactive segmentation systems that can be described by dynamical systems. It is extensible to discrete systems as well.

REFERENCES

1. Vázquez C, Mitiche A, Laganière R. Joint multiregion segmentation and parametric estimation of image motion by basis function representation and level set evolution. IEEE Trans. Pattern Anal. Mach. Intell. 2006; 28(5):782–793. [PubMed: 16640263]

2. Vázquez C, Mitiche A, Ayed IB. Image segmentation as regularized clustering: a fully global curve evolution method. ICIP. 2004:3467–3470.

3. Gao Y, Kikinis R, Bouix S, Shenton ME, Tannenbaum A. A 3d interactive multi-object segmentation tool using local robust statistics driven active contours. Medical Image Analysis. 2012; 16(6):1216–1227. [PubMed: 22831773]

4. Cohen LD, Kimmel R. Global minimum for active contour models: A minimal path approach. Int. J. Comput. Vision. Aug.1997 24:57–78.

5. Karasev P, Kolesov I, Fritscher KD, Vela PA, Mitchell P, Tannenbaum A. Interactive medical image segmentation using pde control of active contours. IEEE Trans. Med. Imaging. 2013; 32(11): 2127–2139. [PubMed: 23893712]

6. Rother C, Kolmogorov V, Blake A. "grabcut": Interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. Aug.2004 23:309–314.

7. Bradski, G.; Kaehler, A. Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly; Cambridge, MA: 2008.
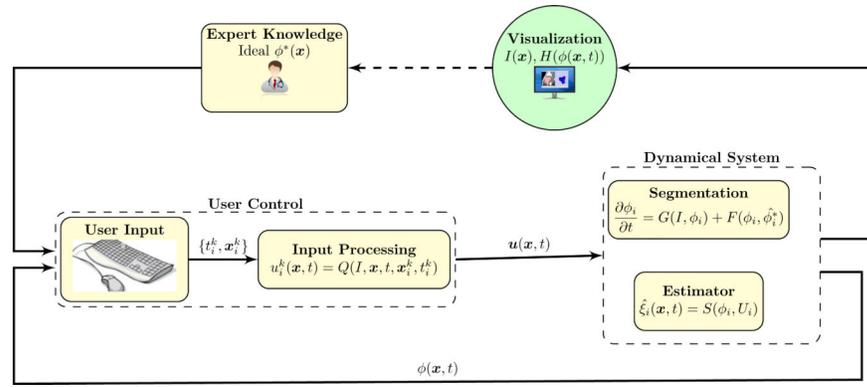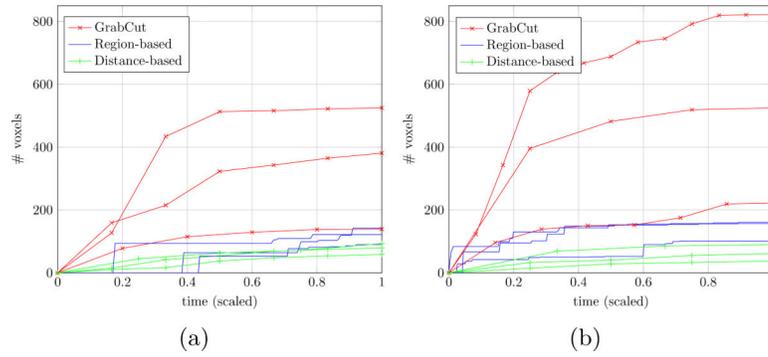
Figure 1.
Diagram of the control-based segmentation framework. The feedback compensates for deficiencies in automatic segmentation by utilizing the expert's knowledge.
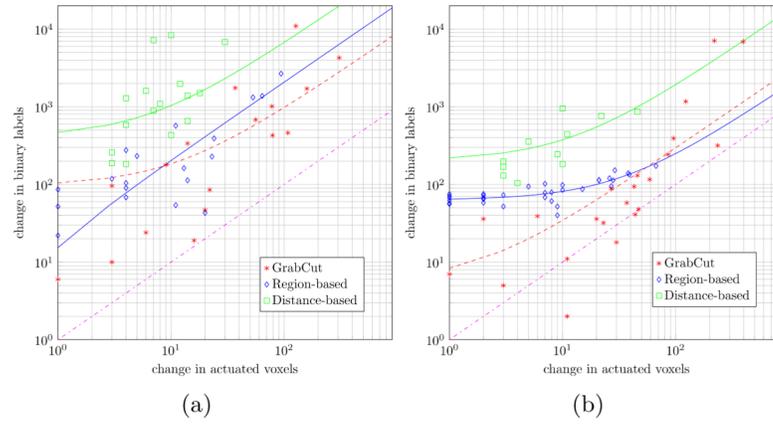
Figure 2.
Two test images are used in a quantitative comparison of GrabCut and the proposed algorithms. Manual segmentations are marked in yellow for the epiphysis (second from the left) and physis (the last).

Figure 3.
Comparison of actuated voxels over time, after initialization for (a) epiphysis (b) physis. The proposed algorithm has both a lower mean actuated count and tighter clustering across repeated segmentations.

Figure 4.
Comparison of dynamic response to user input; data points and linear fit lines for (a) epiphysis (b) physis. Points below the dashed green line indicate wasted user effort since more additional voxels were actuated than reclassified.