4-8-2011

A Human Visual System-Driven Image Segmentation Algorithm

Renbin Peng
Syracuse University, pengrenbin@hotmail.com

Pramod Varshney
Syracuse University, varshney@syr.edu

Department of Electrical Engineering and Computer Science

Technical Report

SYR-EECS-2011-04 April 8, 2011

A Human Visual System-Driven Image Segmentation Algorithm

Renbin Peng, *Student Member, IEEE* rpeng@syr.edu

Pramod K. Varshney, *Fellow, IEEE* varshney@syr.edu

ABSTRACT: This paper presents a novel image segmentation algorithm driven by human visual system (HVS) properties. Quality metrics for evaluating the segmentation result, from both region-based and boundary-based perspectives, are integrated into an objective function. The objective function encodes the HVS properties into a Markov random fields (MRF) framework, where the just-noticeable difference (JND) model is employed when calculating the difference between the image contents. Experiments are carried out to compare the performances of three variations of the presented algorithm and several representative segmentation algorithms available in the literature. Results are very encouraging and show that the presented algorithms outperform the state-of-the-art image segmentation algorithms.

KEYWORDS: Image segmentation, human visual system, Markov random fields, just-noticeable difference

Syracuse University - Department of EECS,
4-206 CST, Syracuse, NY 13244
(P) 315.443.2652 (F) 315.443.2583
http://eecs.syr.edu

A Human Visual System-Driven Image Segmentation Algorithm[1]

Renbin Peng[*], *Student Member, IEEE,* and Pramod K. Varshney, *Fellow, IEEE*[2]

***Abstract*—This paper presents a novel image segmentation algorithm driven by human visual system (HVS) properties. Quality metrics for evaluating the segmentation result, from both region-based and boundary-based perspectives, are integrated into an objective function. The objective function encodes the HVS properties into a Markov random fields (MRF) framework, where the just-noticeable difference (JND) model is employed when calculating the difference between the image contents. Experiments are carried out to compare the performances of three variations of the presented algorithm and several representative segmentation algorithms available in the literature. Results are very encouraging and show that the presented algorithms outperform the state-of-the-art image segmentation algorithms.**

***Index Terms*—Image segmentation, human visual system, Markov random fields, just-noticeable difference**

***EDICS*—ARS-RBS**

I. INTRODUCTION

Image segmentation plays a critical role in image analysis. It subdivides an image into its constituent parts in order to extract information regarding objects of interest, and has an impact on all the subsequent image analysis tasks, such as object classification and scene interpretation [1]. Image segmentation is a challenging problem in computer vision, and a wide variety of solutions have been presented. These include thresholding techniques [2], Markov random fields (MRF)-based approaches [3][4], multi-resolution algorithms [5] and partial differential equations (PDE)-based methods [6]. Surveys of image segmentation techniques can be found in [1][7][8]. Based on the image information being employed for the segmentation task, image segmentation algorithms can be classified into three categories: region-based segmentation, boundary- or edge-based segmentation and the methods combining both region and

[1] This work was supported by AFOSR under grant FA9550-06-C-0036.

[2] Renbin Peng and Pramod K. Varshney are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, USA. (Emails: rpeng@syr.edu, varshney@syr.edu; phone: (315) 443-1060; fax: (315) 443-4745; mailing address: 4-206 Center for Science and Technology, Syracuse University, Syracuse, NY, 13244.)

[*]: Corresponding Author.

boundary (edge) information.

Region-based segmentation methods aim at exploiting the image contextual information, such as spatial dependency or spatial distribution. The segmented images are expected to consist of regions within which the image content is homogeneous, while the contrast between neighboring regions is high. Typical methods falling into this category include region growing, watershed, some MRF-based methods [3], mean-shift [9] and the recently presented lossy data compression-based approach [10]. Segmentation methods based on the boundary or edge information are designed to exploit the discontinuity of the image features, such as difference in texture or pixel intensity, on the two sides of the boundary. Typical methods in this group include gradient-based methods, such as the Canny edge detector [11], line detection methods, such as the Hough transform [12], those taking into account the interaction between boundaries (or edges) [13][14][15], and the methods derived from physics models [16][17]. There also exist algorithms that combine region-based and boundary-based segmentations in order to benefit from fusing these two complementary approaches. There are two types of algorithms that belong to this category. The first type of algorithms carries out region and boundary segmentations sequentially [18][19][20], where one segmentation method is employed as the preprocessing or initialization step of another. The second type performs segmentation by considering region and boundary information simultaneously [21][22].

Design of a suitable objective function is crucial to the performance of image segmentation approaches. Good segmentation algorithms require an efficient scheme for parameter adjustment and an appropriate description of the desired properties of the segmentation result, which, of course, are all very challenging tasks. In real-world applications, the performance of some segmentation algorithms is influenced by their dependence on the parameters of these algorithms. But the optimum parameters and, therefore, satisfactory segmentation results are not easy to obtain. Some segmentation algorithms only partially incorporate the feature information from region and boundary perspectives, and fail to fully take advantage of fusing the two types of information. For example, Markov Chain Monte Carlo (MCMC) has been employed [23] to solve the maximum *a posteriori* (MAP)-MRF estimation problem for generative

image segmentation. Due to many constraints involved in this generative approach, the selection of suitable parameters for satisfactory segmentation becomes difficult. It is also not easy for the objective function in [13] to yield a satisfactory balance between connecting the boundary and labeling the pixels, since there are many parameters which need to be chosen carefully. In [15], the proposed objective function does not exploit fully the connectivity property of the neighboring edge components. The normalized cut methods [24][25] can capture salient parts of an image. However, due to the ad hoc approximations introduced when relaxing this NP-hard computational problem, these methods do not exploit well the image content information which is useful for segmentation. As a result, the algorithms often perform unsatisfactorily.

Another weakness of many existing segmentation algorithms is that they are developed based on the information provided only by the image data and neglect the fact that the human is the best and usually the ultimate evaluator of the segmentation result. That is, these algorithms do not consider the impact of the human visual system (HVS) on object interpretation and information extraction. As a result, many algorithms are inconsistent with the preferences of human vision. There do exist efforts to incorporate HVS information into image segmentation, e.g., [26][27][28], but their performances were constrained by the simplistic computational models as well as an insufficient consideration of the HVS properties when designing the objective functions.

Our work aims at designing an image segmentation algorithm based on HVS properties, with the segmentation performance robust to the variations in the parameter values of the algorithm. More specifically, we integrate region label estimation for each pixel with boundary localization for each region, according to the quality metrics for region-based and boundary-based segmentation evaluations. These metrics attempt to mimic the preferences of human vision to good segmentation and thus make the segmentation HVS-driven. Under a Bayesian framework, the HVS-driven quality metrics are encoded in the MRF as the priors of the *a posteriori* distribution which is the objective function for segmentation. Segmentation is carried out by optimizing the objective function which reflects the desired properties of

segmentation from both global and local perspectives. Three variations of the algorithm are developed. The first one integrates the region and boundary information simultaneously during segmentation. The second one carries out region-based segmentation and boundary-based segmentation iteratively. The third one takes advantage of only the boundary information for segmentation. To the best of our knowledge, the presented algorithms incorporate a more comprehensive consideration of HVS properties as compared to existing segmentation algorithms, from both region-based and boundary-based perspectives.

The performance of the presented algorithms is compared against several representative segmentation and clustering methods available in the literature. These include Gaussian assumption-based dynamic clustering algorithm [29] (GADCA), iterative mode separation algorithm [29] (IMSA), higher-order statistics method based on local maxima detection and adaptive wavelet transform [30] (HOSLW), the conventional MRF-based algorithm, Otsu thresholding [31][32], the level set evolution-based method without reinitialization (LSEWRI) [33], the region-based active contour model (RACM) [34], and the multi-scale normalized cuts-based segmentation (MNCut) [35]. Experiments and analyses are carried out on medical and natural images, where the medical images are two typical mammogram and MRI brain images, and the Cameraman image is used as the representative image of a natural scene.

This paper is structured as follows. Section II introduces the HVS-driven image segmentation model under the MAP-MRF framework. The criteria for evaluating region-based segmentation and the resulting energy function are discussed in Section III. In Section IV, the boundary-based evaluation criteria are discussed and encoded into the energy function via the development of a novel concept, called boundary element in this paper, which describes the interaction between pixel labels, boundary configuration and the image content. The integrated objective function that includes both region and boundary information is described in Section V, where the optimization method and the three variations of the HVS-driven segmentation algorithm are discussed. Experimental results and performance comparisons between the presented algorithms and other representative segmentation and clustering algorithms are presented in Section VI. Concluding remarks are provided in Section VII.

II. HVS-Driven Image Segmentation and MAP-MRF Framework

In this section, we first introduce the motivation for HVS-driven image segmentation, and then discuss our segmentation model under the MAP-MRF framework, which incorporates the information from both region-based and boundary-based segmentation perspectives. The components of the model will be discussed in detail in Sections III and IV.

A. HVS-Driven Image Segmentation

HVS-driven segmentation, considered in this paper, is motivated by the fact that, in most circumstances, humans are the ultimate judge of the quality of a segmentation result. So a segmentation algorithm is likely to yield satisfactory results if the objective function is designed by including HVS preferences within the context of segmentation. In this work, the MAP-MRF framework, to be discussed next, is employed for the segmentation problem. More specifically, several HVS-based segmentation quality evaluation metrics are incorporated into the objective function as prior information, which are encoded in the MRF model to obtain the *a posteriori* probability distribution of the segmentation result given the observed image data. The segmentation is carried out in an iterative manner, which aims at finding the MAP solution to the optimization problem. In the segmentation procedure, the "better" intermediate segmentation result, as evaluated by the metrics to be discussed in the next two sections for region and boundary, is assigned a higher survival probability.

B. MAP-MRF Framework

Under the MAP framework, image segmentation can be obtained by solving the following optimization problem

$$\left(\hat{L},\hat{B}\right)=\arg\max_{L\in\Omega_L,B\in\Omega_B} P(L,B\mid Y)=\arg\max_{L\in\Omega_L,B\in\Omega_B} P(Y\mid L,B)P(L,B) \tag{1}$$

where $P(L,B\mid Y)$ is the *a posteriori* distribution of the label field, *L*, and the boundary field, *B*, given the observed image, *Y*. *L* and *B* are assumed to have the MRF property, and they consist of pixel labels and

boundary elements, respectively. The boundary elements will be defined in Section IV. Ω_L and Ω_B are the configuration spaces of L and B. $\Omega_L = \{l_1, l_2, \cdots, l_M\}$, where l_i is the label of the pixel with the index i. $l_i = \{0,1,\cdots, A-1\}$, and A is the number of possible region types. For example, for binary segmentation, A=2. $\Omega_B = \{b_1, b_2, \cdots, b_M\}$, where b_i is the boundary element of the pixel with the index i. M is the total number of pixels in an image. A segmented image region is composed of the pixels with the same label. In this paper label-based segmentation is equivalent to region-based segmentation, and the two terms will be used interchangeably.

Thus, we obtain our segmentation model under the MAP-MRF framework with the region label MRF L and the boundary MRF B. In our work, the label field L and the boundary field B are defined as functions of the image data Y, that is, $L = L(Y)$ and $B = B(Y)$. $L(Y)$ and $B(Y)$ will be precisely defined in the next two sections. Therefore, the likelihood term in (1) has the form

$$P(Y \mid L, B) = P(Y \mid L(Y), B(Y)) = 1 \tag{2}$$

So (1) is reduced to

$$\left(\hat{L}, \hat{B}\right) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\max} \; P(L(Y), B(Y)) \tag{3}$$

Since both L and B have been assumed to exhibit MRF properties, according to the Hammersley-Clifford theorem [36], they can be represented in terms of the Gibbs distribution and the optimization problem of (3) can be written as

$$\left(\hat{L}, \hat{B}\right) = \underset{L \in \Omega_L, B \in \Omega_B}{\arg\max} \; \frac{1}{Z} \exp\{-U(L(Y), B(Y))\} \tag{4}$$

where $U(L(Y), B(Y))$ is the energy function, denoting the interaction between label and boundary configurations as well as the observation. Z is included for normalization and is a function of the MRF parameters. For given MRF parameters, (4) is equivalent to

$$\left(\hat{L},\hat{B}\right)=\underset{L\in\Omega_L,B\in\Omega_B}{\arg\min}\ U\left(L(Y),B(Y)\right) \tag{5}$$

The energy function $U\left(L(Y),B(Y)\right)$ consists of two factors, corresponding to region-based and boundary-based segmentations. Here, as in some prior work [1][21], we express the energy function in the following additive form

$$U\left(L(Y),B(Y)\right)=U_L\left(L(Y)\right)+U_B\left(B(Y)\right) \tag{6}$$

where the energy functions $U_L\left(L(Y)\right)$ and $U_B\left(B(Y)\right)$ can be considered as the quality metrics corresponding to region- and boundary-based evaluations, respectively. We want to emphasize that *L* and *B* here are two different aspects of the same segmentation result, and the corresponding metrics are the complementary evaluations for the same segmentation result from region and boundary perspectives, respectively.

III. Energy Function for Region-based Image Segmentation

In this section, the energy function $U_L\left(L(Y)\right)$ in (6) corresponding to region-based segmentation is developed. The approach is based on human preference for good segmentation, from a region-based segmentation perspective.

We note that a human often evaluates the segmentation result in both global and local manners, that is, the fitness of a segmentation result to the entire image content and the local image region are considered simultaneously. Therefore, both region-based and boundary-based segmentation evaluations should be taken into account. In this section, we only consider region-based evaluation, and postpone the consideration of boundary-based evaluation to the next section. We summarize the desirable properties for good segmentation in terms of region-based evaluation as follows.

(i) The contrast of pixel intensities between two neighboring regions, i.e., inter-region contrast, should be large;

(ii) The contrast of pixel intensities within a region, i.e., intra-region contrast, should be small;

(iii) The pixel labels should correspond to homogeneous regions, that is, neighboring pixels prefer having the same label.

Criteria (i) and (ii) represent the global properties of a good segmentation, and criterion (iii) is a local property which indicates that the segmentation should yield large-sized regions.

Thus, the region-based segmentation evaluation metric $U_L(L(Y))$ should consist of two types of measures, namely, global inter- and intra-region contrast measures and a local label homogeneity measure. We express the composite measure also using an additive form as follows

$$U_L(L(Y)) = U_{Global}(L(Y)) + U_{Local}(L(Y)) \tag{7}$$

We can see from the three desirable properties that the contrasts between the neighboring regions and those between the neighboring pixels need to be calculated in order to obtain a quantitative metric for evaluating segmentation quality. In this work, we incorporate the HVS properties into the contrast measure via the JND model [37][38][39]. HVS is capable of only perceiving pixel intensity changes above a certain visibility threshold, which, in turn, is determined by the underlying physiological and psychophysical mechanisms. JND refers to the minimum visibility threshold above which visual contents can be distinguished. The JND model plays an important role in perceptual image and video processing, and has been successfully used in measuring the difference or distortion of the image contents [40][41]. In this paper, we use the spatial JND model, i.e., pixel-wise JND, presented in [42], which is defined as a nonlinear additive model,

$$\mathrm{JND}_p(i) = T^L(i) + T^t(i) - C^{L,t}(i) \cdot \min\left\{T^L(i), T^t(i)\right\} \tag{8}$$

where $\mathrm{JND}_p(i)$ is the JND threshold of the pixel indexed by i. $T^L(i)$ and $T^t(i)$ are the visibility thresholds due to luminance adaptation and texture masking, respectively, and $C^{L,t}(i)$ represents the overlapping effect in masking where $0 < C^{L,t}(i) \le 1$. Details on the definition of JND and its computation are available in [37][42].

A. Energy Function for Global Contrast

In order to incorporate the desired characteristic of global inter- and intra-region contrasts into the energy function, we define a global neighborhood system, 61 pixels by 61 pixels in this paper, with pixel of interest at the center. The segmentation quality metric based on criteria (i) and (ii) can be expressed as

$$Q_{Global} = a_1 \cdot C_{Inter} - a_2 \cdot C_{Intra} \tag{9}$$

where C_{Inter} and C_{Intra} denote the inter- and intra-region contrasts of the pixel intensity based on JND, respectively, and a_1 and a_2 are two non-negative weights that control the contributions of the two types of contrasts to the energy function. A higher value of Q_{Global} means a better segmentation quality.

Due to the Markov property assumed in this paper, the quality metric value corresponding to a single pixel, *s*, is independent on other pixels given the segmentation result in the global neighborhood system of *s*. Therefore, we have the energy function for the global contrast as

$$U_{Global}\left(L(Y)\right) = -\sum_{s=1}^{M} Q_{Global-s|\Gamma^s}(s) = -\sum_{s=1}^{M}\left[a_1 \cdot C_{Inter-s|\Gamma^s}(s) - a_2 \cdot C_{Intra-s|\Gamma^s}(s)\right] \tag{10}$$

where $s \mid \Gamma^s$ means that the contrast is a function of the label of the pixel *s* and is calculated given that the labels of the rest of the pixels in the global neighborhood system of *s*, Γ^s, are fixed. To speed up the computation, instead of calculating the contrasts in a pixel-wise manner, the following region-wise measure of the inter-region contrast is employed,

$$C_{Inter-s|\Gamma^s}(s) = \varphi\left\{\Xi_{s|\Gamma^s}(\mu_1, \mu_0)\right\} \tag{11}$$

where the calculation is carried out on $\Gamma^s \cup \{s\}$, the set consisting of the pixels included in Γ^s plus pixel *s*, and "$\cup$" means "union". Here, $\Xi_{s|\Gamma^s}(\mu_1, \mu_0) = |\mu_1 - \mu_0| / \min[\mathrm{JND}_R(1), \mathrm{JND}_R(0)]$. $\mu_u = \sum_{i=1, \text{ and } l_i=u}^{N_u} y_i / N_u(s)$, which is the mean value of the pixel intensity in the u^{th} type region, $u \in \{0,1\}$ for binary segmentation, and $N_u(s)$ is the number of pixels with the label *u*. y_i is the i^{th} pixel with the label *u*.

$\mathrm{JND}_R(u) = \sum_{i=1,\text{ and } l_i=u}^{N_u} \mathrm{JND}_p(i) / N_u(s)$, which is the average JND value of the regions with the label *u*. Thus, $\Xi_{s|\Gamma^s}(\mu_1, \mu_0)$ is the measure of the average intensity difference of the two types of regions weighted by the minimum of the average JND values of the two types of regions. $\varphi(D)$ is a robust function that reduces the impact of outliers and is defined in [15] as

$$\varphi(D) = \frac{D^4}{G + D^4} \tag{12}$$

where *G* is a small positive constant.

For the intra-region contrast, we define and employ the metric,

$$C_{Intra-s|\Gamma^s}(s) = \frac{1}{N_1(s) + N_2(s)} \left\{ \Phi_{s|\Gamma^s}(0) + \Phi_{s|\Gamma^s}(1) \right\} \tag{13}$$

Still, the calculation is carried out on $\Gamma^s \cup \{s\}$, where $\Phi_{s|\Gamma^s}(0) = \sum_{i=1 \text{ and } l_i=0}^{N_0} \varphi\left(\left| y_i - \mu_0 \right| / \mathrm{JND}_P(i) \right)$ and $\Phi_{s|\Gamma^s}(1) = \sum_{i=1 \text{ and } l_i=1}^{N_1} \varphi\left(\left| y_i - \mu_1 \right| / \mathrm{JND}_P(i) \right)$. Here, $\left| y_i - \mu_j \right| / \mathrm{JND}_P(i)$, *j*=0 and 1 for binary segmentation, is the weighted difference between the intensity of the i^{th} pixel and the average intensity of the region to which the pixel belongs. So $\Phi_{s|\Gamma^s}(0)$ and $\Phi_{s|\Gamma^s}(1)$ in (13) measure the "variation", or the inhomogeneity, of the two types of regions.

We can see that (11) and (13) measure how the segmentation result of a single pixel *s*, i.e., the label of *s*, affects the segmentation in a global manner, and take into account the inter- and intra-region contrasts at the same time. Therefore, the energy function defined in (10) reduces the risk of being too biased when segmenting an image [24]. For example, an algorithm may become very greedy in finding small but tight clusters in the image data if only intra-region contrast is considered [24]. Additionally, (11) and (13) do not require modeling of the probability distributions of the noise and the image data, which may improve the robustness of the presented algorithm.

B. *Energy Function for Local Homogeneity*

In this sub-section, we discuss the energy function based on criterion (iii), incorporating local homogeneity for good segmentation. A second-order neighborhood system, that includes 8 nearest neighboring pixels of the pixel of interest *s*, is employed for describing local homogeneity. The pixels, excluding *s*, included in the solid (yellow) rectangular shown in Fig. 1 (b) is a configuration of the second-order neighborhood system of *s*. Besides the label homogeneity, we also incorporate information about the dynamics of the pixel intensity in order to make the metric adaptive to non-stationary image contents. The energy function is defined as

$$U_{Local}\left(L(Y)\right) = -\sum_{s=1}^{M}\left\{\sum_{\eta_s \in NB_2(s)} \psi(l_s, l_{\eta_s}) \exp\left[-\varphi\left(\Delta_{s,\eta_s}(y)\right)\right]\right\} \tag{14}$$

where $\Delta_{s,\eta_s}(y) = \left|y_s - y_{\eta_s}\right| / \min[\mathrm{JND}_P(s), \mathrm{JND}_P(\eta_s)]$ represents the JND-weighted contrast between pixel *s* and its second-order neighbor η_s. $NB_2(s)$ denotes the set of all of the second-order neighbors of *s*. The cost function for the label configuration of the neighboring pixel pair is denoted by $\psi(l_s, l_{\eta_s})$ and defined in (15) for binary segmentation.

$$\psi(l_s, l_{\eta_s}) = \begin{cases} \beta, \text{if } l_s = l_{\eta_s}, \\ -\beta, \text{ otherwise.} \end{cases} \tag{15}$$

where β is a non-negative real number and is the cost used to define the label homogeneity measure of the neighboring pixel pair. Eq. (14) has a form similar to the generalized Potts model [43], except for the additional robust estimation shown in (12) and the JND-weighting operation.

IV. Energy Function for Boundary-Based Image Segmentation

In this section, we develop the energy function for boundary-based segmentation, which is also derived from the desirable properties for good segmentation in terms of a boundary-based evaluation. These properties are listed below,

(i) Region boundary should be smooth and of as small a length as possible. In other words, the boundary should avoid containing too many sharp angles or turns;

(ii) The intensity contrast of a neighboring pixel pair on the two sides of the boundary should be large, while the contrast within a region enclosed by the boundary curve should be small;

(iii) The pixels lying on the boundary curve should be connected.

Criterion (i) represents a property which is characterized by the image contents in both global and local manners. Criteria (ii) and (iii) are properties of a good segmentation in small regions, and can be measured locally.

When designing the energy function $U_B(B(Y))$, due to the huge computational burden for the global boundary-based feature measure, we only employ the local properties described in the boundary-based segmentation evaluation. That is,

$$U_B(B(Y)) = U_{Local}(B(Y)) \tag{16}$$

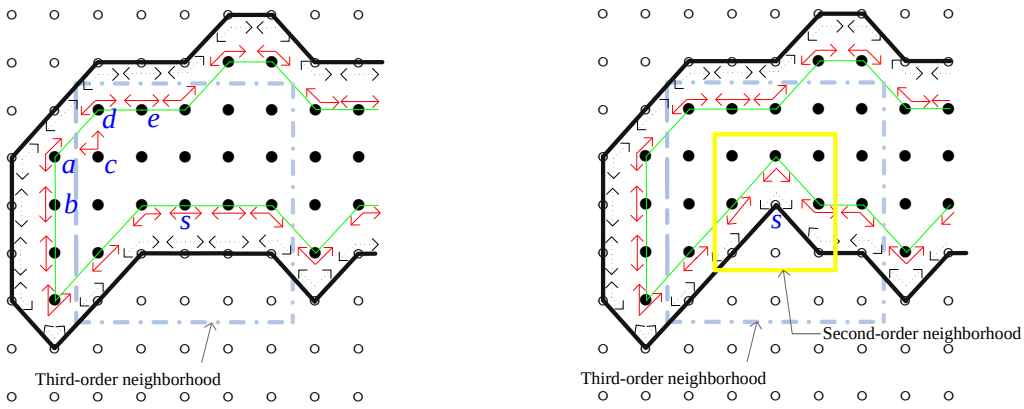
Before deriving the energy function, we first discuss the neighborhood system and define a novel concept called the boundary element in the next sub-section.

A. Boundary Element and Neighborhood System

The energy function of boundary-based segmentation is calculated based on a novel concept, the boundary element, defined for each boundary pixel. A boundary pixel is a pixel for which at least one of its second-order neighbors has a label different from it. A boundary element consists of an angle together with its two directed edges. Suppose *s* is a boundary pixel, the angle of the boundary element of *s* originates from *s*, and two edges of the angle point to the two neighboring boundary pixels of *s* which have the same labels as *s*. The two edges separate the pixels into different regions according to whether their labels are different from or same as that of *s*. Some examples of the angles are shown in Fig. 2. The value of the angle is constrained to be in the range $[0, \pi]$, and is related to the smoothness of the boundary curve. Small angles correspond to sharp turns and therefore to a wiggly boundary, while large angles

correspond to a smooth region boundary. Naturally, we prefer large angles. The edges of the boundary element are related to the connectedness of the boundary pixels and thus the continuation of the boundary contour. The edges also play an important role in determining the cross-boundary pixel pair, which will be discussed in Section IV.B. The boundary elements for two special segmentation configurations, isolated segmentation and interior pixel segmentation, will be discussed in the subsection IV.B.b, where the pixel of interest has either a different label from or the same label as its second-order neighbors.

A boundary element is determined from the interaction between the pixel of interest, say, *s*, and its 8 second-order neighbors. The total energy of a neighborhood system centered at *s* is dependent on the boundary elements of *s* and its 8 neighbors. A change in the label of *s* may affect the boundary elements of its 8 neighbors. Therefore, we define the neighborhood system for boundary-based segmentation as one consisting of the 8 nearest neighbors in the second-order neighborhood system of *s* plus all the second-order neighbors of these 8 pixels, excluding *s*. This is, in fact, a third-order neighborhood system of *s*, $NB_3(s)$, i.e., $NB_3(s) = \{\cup NB_2(\eta_s)\} \cup \{NB_2(s)\} \setminus s$, where " $\setminus s$ " means "excluding *s*". $NB_3(s)$ is shown in Fig. 1, in which the pixels, excluding *s*, included in the rectangular bold dot-dash (light blue) line constitute the third-order neighborhood system of *s* for boundary-based segmentation. In Fig. 1, *s* lies at the center of the neighborhood system. The pixels represented by the solid points belong to one type of region with label zero (0 region). The small circles represent pixels belonging to the region with label one (1 region). The thin solid (green) line and the bold solid (black) line denote the boundary curves of the 0 and 1 regions, respectively. The solid (red) directed edges, paired together with the angle between the edge pair, form the boundary elements of the pixels in the 0 region. The dashed (black) directed edge pairs and the corresponding angles form the boundary elements of the pixels in the 1 region. Now, let us consider the situation when the label of a pixel changes. The impact of the label change of *s* on the boundary elements of its 8 neighbors is shown in Fig. 1 (b), where *s* changes its label from 0 (Fig. 1 (a)) to 1 (Fig. 1 (b)) and results in the change of the boundary elements of its second-order neighbors.



(a) (b)

Fig. 1 Boundary elements, second-order and third-order neighborhood systems. (a) Typical boundary elements and third-order neighborhood system; (b) impact of label change of *s* on the boundary elements of its second-order neighbors.

From Fig. 1, we can see that the boundary curve of a region is determined by the boundary pixels together with their edge pairs. However, not all the boundary pixels contribute directly to the boundary curve, and it is possible that the impact of some boundary pixels is "hidden" by the neighboring boundary pixels when forming the curve. For example, pixel *c* in Fig. 1 (a) is a boundary pixel, but the boundary curve, the thin solid (green) line, does not pass through *c* and its two edges, $\vec{ca}$ and $\vec{cd}$. This is because the boundary elements *a* and *d* together with their edges, $\vec{ab}$, $\vec{ad}$ and $\vec{da}$, $\vec{de}$, hide the contribution of *c* to the curve. However, the change in the label of *c* will affect the boundary elements of *a* and *d*, as mentioned above. Thus, we still need to consider the boundary element and the energy corresponding to pixel *c* when updating the pixel labels, the parameters of the MRF and therefore the shape of the boundary curve. In other words, boundary pixel *c* impacts the boundary curve in an indirect or implicit manner. The energy function corresponding to boundary elements will be discussed in next sub-section.

From Fig. 1, we notice that the angles of the boundary elements reflect the variation in the boundary shape, and the directed edges represent the interaction and the relative locations of the neighboring boundary pixels.

B. Energy Function for Boundary-Based Segmentation

As indicated earlier, we prefer smooth boundaries, i.e., gentle "turns". This corresponds to large angles, and the energy function is, thus, designed as a monotonically decreasing function of the angle value. At the same time, a "reasonable" turn, which results from the significant contrast of pixel intensities across the boundary, should also be maintained. Furthermore, the intensity contrast of pixels on the same side of the boundary and belonging to the same region should be small. In this way, we may make the smoothness measure of the boundary curve dynamic and adaptive to non-stationary image content. Thus, the energy function of a boundary element is composed of three terms: the first one is related to the angle between the two edges, the second one is related to the intensity contrast across the boundary, and the third one is related to the contrast on the same side of the boundary which is called intra-pie slice contrast. We define a "pie slice" next.

a. Cross-boundary contrast and intra-pie slice contrast

Consider a boundary pixel, *s*. A pie-slice of *s*, by definition, consists of *s*, as the origin, and some other pixels in the second-order neighborhood system of *s*. These pixels must have the same labels as *s*, and there are no pixels with different labels from *s* in the pie slice. The cross-boundary contrast is calculated as the intensity difference between pixels in a pie slice of *s* and the pixels in the second-order neighborhood system of *s* but with labels different from *s*. All possible configurations of the single pie slice in a second-order neighborhood system are shown in Fig. 2. The configuration in each figure is valid when rotated by $\pi/2$. The configurations of the multi-pie slices, where a second-order neighborhood system contains more than one pie slice, can be determined in a similar manner and an example of it is shown in Fig. 3.

In Fig. 2, *s* is the pixel of interest. Crosses "$\times$" represent the pixels with the same label as *s* (including *s*), and form a pie slice of *s*. The remaining pixels have labels different from *s*, and are represented by "+", "-" and "*". The cross-boundary contrast of a pie slice is determined by averaging the contrast of the so-called cross-boundary pixel pairs, which are determined by the two edges of the boundary element. More specifically, in the figure the cross-boundary pixel pairs corresponding to the pixels represented by

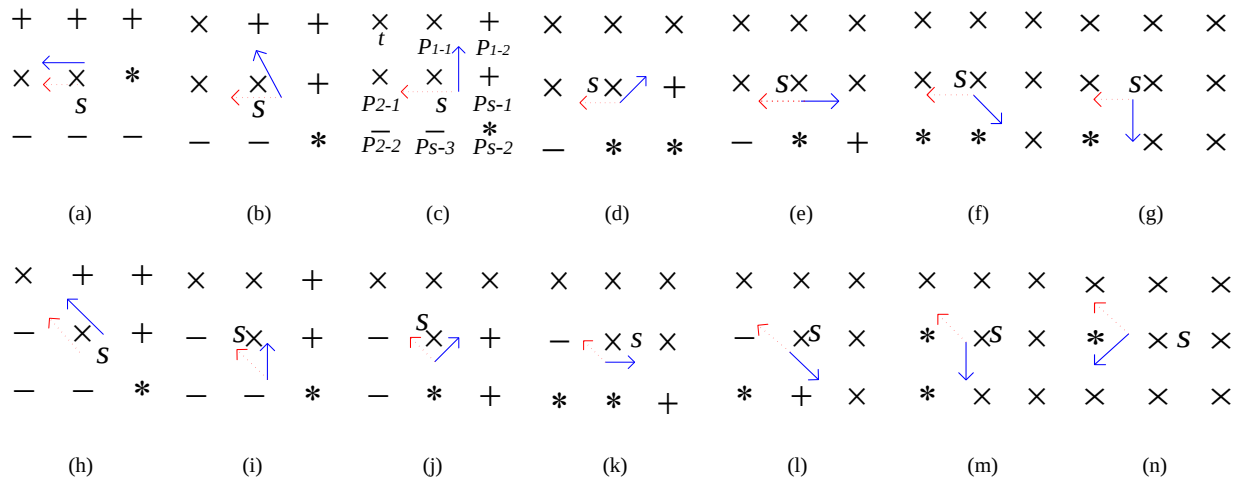


Fig. 2. Possible configurations of the single pie slice in a second-order neighborhood system. The configuration in each figure is valid when rotated by $\pi/2$.

"-" and "×" are determined by the dashed (red) edge, those corresponding to the pixels represented by "+" and "×" are determined by the solid (blue) edge, and those corresponding to the pixels represented by "*" and "×" are determined by both of the two types of edges mentioned above. The cross-boundary pixel pair configuration in each figure is valid when rotated by $\pi/2$.

Suppose there are $N_{PS}(s)$ pie slices included in a second-order neighborhood system with the origin at s. The cross-boundary contrast of the i^{th} pie slice consists of two terms, one including s and another without s, as shown in (17)

$$Contr(s)_{CB-i} = \frac{1}{N_r + N_j}\left[\sum_{r=1}^{N_r}\varphi\left(\frac{\left|y_{P_{r-1}} - y_{P_{r-2}}\right|}{\min[\mathrm{JND}_P(P_{r-1}), \mathrm{JND}_P(P_{r-2})]}\right) + \sum_{j=1}^{N_j}\varphi\left(\frac{\left|y_s - y_{P_{s-j}}\right|}{\min[\mathrm{JND}_P(s), \mathrm{JND}_P(P_{s-j})]}\right)\right] \quad (17)$$

where the first summation is related to the cross-boundary contrast calculated from the cross-boundary pixel pairs corresponding to the pixels in the i^{th} pie slice of s (excluding s), and the second summation is related to that corresponding to s. Here, y_k denotes the intensity value of pixel k. P_{r-1} and P_{r-2} denote the boundary pixels inside and outside the i^{th} pie slice of s (excluding s). They form the r^{th} cross-boundary pixel pair, in which P_{r-2} is the closest neighboring boundary pixel of P_{r-1} and has a label different from P_{r-1},

and N_r denotes the number of such pairs. Here, $i = 1,...,N_{PS}(s)$ with $N_{PS}(s) \le 4$, where $N_{PS}(s)$ denotes the number of pie slices in the second-order neighborhood system of *s* and 4 is the maximum number of pie slices in a second-order neighborhood system. P_{s-j} are the boundary pixels outside the i^{th} pie slice and have the labels different from *s*. *s* and P_{s-j} form the j^{th} cross-boundary pixel pair, and N_j denotes the number of such pairs. By annotating the pixel in Fig. 2 (c), we show an example for calculating the cross-boundary contrast using (17), where the four "×" pixels form a pie slice and the cross-boundary pixel pairs are determined by the solid and dashed edges. Here, pixel *t* is not involved in the calculation since it is not a boundary pixel when we calculate the cross-boundary contrast of *s*. As mentioned before, in this case there is only one pie slice in the second-order neighborhood system.

The intra-pie slice contrast of the i^{th} pie slice is defined as

$$Contr(s)_{PS-i} = \frac{1}{N_{PS-i}(s)} \sum_{j=1}^{N_{PS-i}(s)} \varphi\left(\frac{\left| y_{i-j} - \mu_{PS-i} \right|}{\mathrm{JND}_P(j)} \right) \tag{18}$$

where $N_{PS-i}(s)$ is the number of pixels included in the i^{th} pie slice in the second-order neighborhood system of *s*. μ_{PS-i} is the average intensity value of the pixels in the i^{th} pie slice, and y_{i-j} is the observed pixel intensity of the j^{th} pixel in the i^{th} pie slice.

b. Energy function for the second-order neighborhood system

In this subsection, we develop the energy function for the second-order neighborhood system of the pixel of interest.

As mentioned before, the energy function is designed to encourage a large turn angle and large cross-boundary contrast. At the same time, we would like to have the contrast of the pixel intensity within each pie slice to be as small as possible. We first find all the pie slices in each second-order neighborhood system of the pixel of interest *s*, and then the contrasts across boundary and inside the pie slice are

determined by (17) and (18) for each pie slice. The energy corresponding to a second-order neighborhood system can then be calculated as the summation of the energy of each pie slice,

$$\phi(s) = \sum_{i=1}^{N_{PS}(s)} \exp\left[-a_3 \cdot \alpha_i(s) - a_4 \cdot Contr(s)_{CB-i} + a_5 \cdot Contr(s)_{PS-i}\right] \tag{19}$$

where $\alpha_i(s)$, $Contr(s)_{CB-i}$ and $Contr(s)_{PS-i}$ are the angle value, cross-boundary contrast and intra-pie slice contrast of the i^{th} pie slice, respectively. The weights a_3, a_4 and a_5 are non-negative real numbers.

Fig. 3 shows an example for calculating the energy of a multi-pie slice segmentation, where there exist two pie slices Δasb and $\Delta csed$. The corresponding energy is

$$\begin{aligned}\phi(s) = &\exp\left\{-a_3 \cdot \alpha_{\Delta asb} - a_4 \cdot Contr(s)_{CB-\Delta asb} + a_5 \cdot Contr_{PS-\Delta asb}(s)\right\} \\ &+ \exp\left\{-a_3 \cdot \alpha_{\Delta csed} - a_4 \cdot Contr(s)_{CB-\Delta csed} + a_5 \cdot Contr_{PS-\Delta csed}(s)\right\}\end{aligned} \tag{20}$$

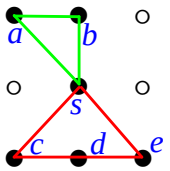


Fig. 3 A typical multi-pie slice configuration, where there exist two pie slices Δasb and $\Delta csed$.

We can see that the multi-pie slice configurations, which correspond to complicated and less preferred segmentation, have more terms than the single pie slice case. The larger the $N_{PS}(s)$, the more "messy" the segmentation is. We show in the appendix that the energy function (19) assigns higher energy and therefore lower probability to the multi-pie slice configurations, which is consistent with our expectation.

In many practical situations, there exist two special segmentation configurations, isolated segmentation and interior pixel segmentation, as mentioned before. Isolated segmentation corresponds to a special boundary element, where the angle of the isolated segmentation is -2π and the cross-boundary contrast is determined by averaging the JND-weighted intensity differences between *s* and its 8 second-order neighbors (with the robust function considered). The intensity contrast within the pie slice for the

isolated segmentation is zero since the pie slice includes only one pixel, *s* itself. Admittedly, the interior pixel is not a boundary pixel and does not have the boundary element as defined. However, for updating the pixel label we still assign a special boundary element to this type of configuration, where the angle of the interior pixel is set to be 2π and the corresponding cross-boundary contrast is zero. The intensity contrast within the pie slice can be calculated using (18), where *s* and its 8 neighboring pixels included in the second-order neighborhood system construct a pie slice.

c. Energy function for boundary-based segmentation

As mentioned before, a change in the label of *s* might impact the boundary elements and therefore the energy function values of the 8 second-order neighborhood systems centered at the 8 second-order neighbors of *s*, plus the second-order neighborhood system of *s*. Let $NB_{2+s}(s)$ denote the set of pixels in the second-order neighborhood system of *s* plus *s* itself, so $NB_{2+s}(s) = NB_2(s) \cup \{s\}$, where $NB_2(s)$ represents the 8 second-order neighbors of *s*, as defined in Section III. B. Thus, the energy function values of the second-order neighborhood systems of the pixels in $NB_{2+s}(s)$ are necessary to calculate the energy corresponding to the label configuration of *s*. Therefore, energy function (16) can be expressed as

$$U_B(B(Y)) = \sum_{s=1}^{M} \left[\frac{1}{N_H(s)} \sum_{h=1}^{N_H(s)} \phi(h) \right] \tag{21}$$

where $\sum_{h=1}^{N_H(s)} \phi(h) / N_H(s)$ is the average energy of the second-order neighborhood systems included in the third-order neighborhood system of *s*. This energy function takes into account the impact of the change in the label of *s* on the local region. The pixels in $NB_{2+s}(s)$ are indexed by *h*, and $N_H(s)$ denotes the number of pixels in $NB_{2+s}(s)$. We can see that $N_H(s) = 9$ if *s* is an internal pixel of an image.

To speed up the computation, only the energy function of the second-order neighborhood system of *s* is calculated in this paper, which is sufficient to produce satisfactory results and is therefore implemented in

the experiments. Thus, we have a simplified energy function for boundary-based segmentation,

$$U_B\left(B(Y)\right)=\sum_{s=1}^{M}\phi(s) \tag{22}$$

V. Overall Objective Function and Optimization

Having defined the energy functions for region-based and boundary-based segmentations, we formulate the overall optimization problem. According to the previous discussion, the image segmentation problem can be transformed to an optimization problem with respect to the pixel label and region boundary configurations,

$$\begin{aligned}\left(\hat{L},\hat{B}\right)&=\underset{L\in\Omega_L,B\in\Omega_B}{\arg\min}\,U\left(L(Y),B(Y)\right)=\underset{L\in\Omega_L,B\in\Omega_B}{\arg\min}\left\{U_L\left(L(Y)\right)+U_B\left(B(Y)\right)\right\}=\underset{L\in\Omega_L,B\in\Omega_B}{\arg\min}\left\{U_{Global}\left(L(Y)\right)+U_{Local}\left(L(Y)\right)+U_{Local}\left(B(Y)\right)\right\}\\&=\underset{L\in\Omega_L,B\in\Omega_B}{\arg\min}\left\{\begin{array}{l}-\sum_{s=1}^{M}\left(a_1\cdot\varphi\left[\Xi_{s|\Gamma^s}(\mu_1,\mu_0)\right]-a_2\cdot\frac{1}{N_1(s)+N_2(s)}\left[\Phi_{s|\Gamma^s}(0)+\Phi_{s|\Gamma^s}(1)\right]\right)\\-\gamma_1\sum_{s=1}^{M}\left[\sum_{\eta_s\in NB_2(s)}\psi(l_s,l_{\eta_s})\exp\left[-\varphi\left(\Delta_{s,\eta_s}(y)\right)\right]\right]\\+\gamma_2\sum_{s=1}^{M}\left[\sum_{i=1}^{N_{PS}(s)}\exp\left(-a_3\cdot\alpha_i(s)-a_4\cdot Contr(s)_{CB-i}+a_5\cdot Contr(s)_{PS-i}\right)\right]\end{array}\right\}\end{aligned} \tag{23}$$

where γ_1 and γ_2 are two hyper-parameters controlling the contributions of the label and the boundary energy functions to the total energy. Finding the optimal configuration of *L* and *B* includes the minimization of (23) with respect to *L* and *B* followed by the maximum likelihood estimation (MLE) of β if γ_1 and γ_2 are given. The procedure proceeds in an iterative manner. The JND weights only need to be calculated once before segmentation. We call this scheme HVS-driven segmentation scheme 1 (HDSS-1).

The hyper-parameters γ_1 and γ_2 can be determined either by trial-and-error or by using the method presented in [22]. In our work, we use the trial-and-error method to choose these parameters and the experimental results show that the performance of the algorithm is not sensitive to these parameters. To further reduce the effort of choosing the hyper-parameters, the optimization of the energy function is carried out in two steps. In the first step, only the energy terms related to region-based segmentation are minimized. In the second step, the boundary energy function is minimized which is based on the

segmentation result from the first step. Steps 1 and 2 are iterated until we achieve a satisfactory result. (In our work, the number of iterations is determined by trial-and-error, but the experiments show that the performance of the algorithm is not sensitive to the number of iterations if the number is larger than a certain value.) We see that in this scheme only the hyper-parameter γ_1 needs to be chosen. We call this scheme the HVS-driven segmentation scheme 2 (HDSS-2). The two-step procedure is given by

$$\text{Step 1:} \qquad \hat{L} = \arg\min_{L\in\Omega_L}\left\{\begin{array}{l} -\sum_{s=1}^{M}\left(a_1\cdot\varphi\left[\Xi_{s|\Gamma^s}(\mu_1,\mu_0)\right] - a_2\cdot\frac{1}{N_1(s)+N_2(s)}\left[\Phi_{s|\Gamma^s}(0)+\Phi_{s|\Gamma^s}(1)\right]\right) \\ -\gamma_1\sum_{s=1}^{M}\left[\sum_{\eta_s\in NB_2(s)}\psi(l_s,l_{\eta_s})\exp\left[-\varphi\left(\Delta_{s,\eta_s}(y)\right)\right]\right] \end{array}\right\}$$

$$\text{Step 2:} \qquad \hat{B} = \arg\min_{B\in\Omega_B}\left\{\sum_{s=1}^{M}\left[\sum_{i=1}^{N_{PS}(s)}\exp\left(-a_3\cdot\alpha_i(s) - a_4\cdot Contr(s)_{CB-i} + a_5\cdot Contr(s)_{PS-i}\right)\right]\right\} \qquad (24)$$

In our experiments, besides the implementation of the above mentioned segmentation schemes, we will also present the results when only the boundary element is considered during the optimization, that is, when only the energy function in Step 2 of (24) is optimized. We call this segmentation method the boundary element-based segmentation (BEBS).

To the best of our knowledge, HDSS-1, HDSS-2 and BEBS are distinctive from existing segmentation algorithms in terms of the design of the objective functions from region- and boundary-based perspectives.

VI. EXPERIMENTS

In this section, comparative results of the segmentation of two types of medical datasets, mammogram and MRI brain images, and one natural image are shown. The comparisons are carried out between the presented algorithms and several representative methods. To save the space, we will not describe the algorithms for comparison purpose in this paper.

The mammogram used in the experiment is from the Digital Database for Screening Mammography (DDSM) [44]. DDSM has 2620 cases available in 43 volumes. A case consists of between 6 and 10 mammograms, and resolution of the images is 50 microns/pixel. The MRI brain image is T1-weighted and

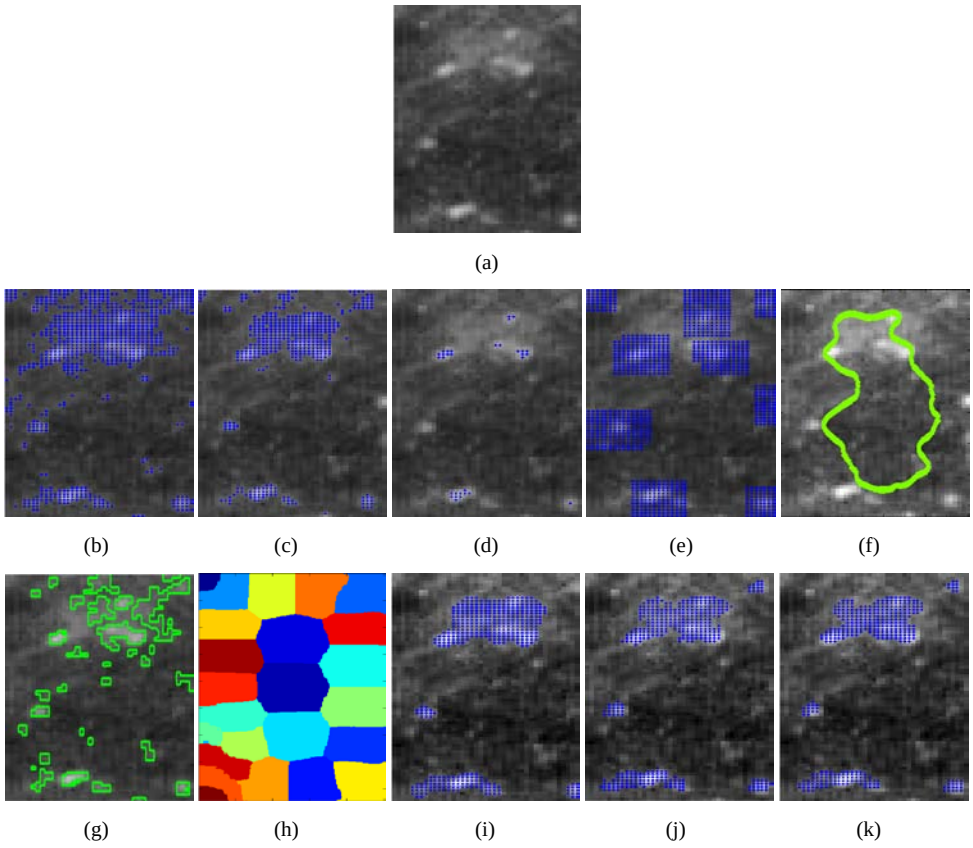
is from SUNY Upstate Medical University. We use Cameraman image as an example of the natural image.

In mammogram segmentation, our goal is to find the pixels which represent lesions. The segmented positive pixels by the presented algorithms are marked in blue. In segmenting the MRI brain image, we wish to segment the white matter (WM) from the gray matter (GM) and cerebrospinal fluid (CSF). The segmented non-WM tissues are shown using purple and black colors. For the Cameraman image, we want to segment out the man as the foreground from the background, i.e., the building and the meadow. We also treat the camera as a background object and do not attempt to segment it out. The segmentation result is represented by a binary image, where the dark part is the foreground.

We test HDSS-1 and BEBS for mammogram data, HDSS-1, HDSS-1 and BEBS for MRI brain image data, and HDSS-1for Cameraman image. To be fair, we use the same initialization for all the algorithms implemented in the experiments. We set a_1, a_2, a_3 and a_5 to be 1, and a_4 to be 15. In HDSS-1, γ_2 is chosen to be 1. $\gamma_1's$ in HDSS-1 and HDSS-2 are set to be 0.1. From the experimental results, we notice that the segmentation performance is not sensitive to the choice of these parameters. We use the iterative conditional modes (ICM) algorithm [46] as optimization method for all the three algorithms.

Fig. 4 shows the segmentation results of the mammogram with lesions, where the blue points denote the segmented positives. The regions enclosed by the light green curves in Figs. 4 (f) and (g) correspond to the segmented lesion regions by LSEWRI and RACM, respectively.

From the figures, we can see that BEBS and HDSS-1, shown in Figs. 4 (j) and (k), yield better results than the other methods. Otsu thresholding, (b), yields too many false alarms. Obviously, for an image in which the intensity contrast is not very high, like the mammogram, intra-class variance measure is insufficient for yielding good segmentation. The advantage of GADCA (c) is that it converges quickly, but it yields many false positives. IMSA (d) may converge to local extrema and misses many lesions. HOSLW method (e) can find the lesions efficiently, but it still generates false alarms and fails to determine the shape of lesion which, however, plays a very important role in discriminating the benign



(a)

(b) (c) (d) (e) (f)

(g) (h) (i) (j) (k)

Fig.4. Original mammogram and the segmentation results. (a) Original mammogram with lesions; (b) segmentation by Otsu thresholding; (c) segmentation by GADCA; (d) segmentation by IMSA; (e) segmentation by HOSLW; (f) segmentation by LSEWRI; (g) segmentation by RACM; (h) segmentation by MNCut; (i) segmentation by conventional MRF; (j) segmentation by BEBS; (k) segmentation by HDSS-1.

tumors from the malignant ones. Moreover, the segmentation performance of HOSLW depends on how accurately we can estimate the number of lesion pixels, which is usually not available in real-world applications. The LSEWRI method, shown in (f), also performs poorly and yields many mis-segmentations. RACM (g) finds all the lesions but with many false alarms. Besides, it also fails to determine the lesion shapes. The number of iterations of the two level set-based algorithms is set to be 500, which is sufficient for them to converge. For the MNCut method, we tried several numbers of segments, but did not observe any satisfactory results. A typical segmentation is shown in (h). Conventional MRF (i) does not find all the lesions and it also fails to determine lesion shape. This is because the conventional MRF only emphasizes intra-region homogeneity and label smoothness, such that it is too conservative and works poorly when the image contents are complex. Besides, all the methods used for comparison purpose fail to mimic the adaptation of HVS to the complexity and non-stationarity of the image contents.

As shown in Fig. 4 (j), BEBS performs satisfactorily, but the lesion contour is not smooth. This is because the boundary energy function does not emphasize the label homogeneity of the neighboring pixels as well as the global contrast of the image. On the contrary, HDSS-1 (k) integrates the boundary information, global contrast and pixel label homogeneity, and therefore yields a better result.

Fig. 5 shows the results for MRI brain image segmentation, where the non-brain background is first removed from the image and segmentation is carried out only on the brain pixels. Similar to the mammogram case, our presented algorithms yield better segmentations than the representative ones, in terms of lower mis-segmentation rate and higher accuracy in object boundary determination. In the figures, white regions denote the WM. The regions enclosed by the light green curves in Figs. 5 (d) and (e) correspond to the segmented WM by the two level set-based methods with the number of iterations being 2000, which is sufficient for them to converge.

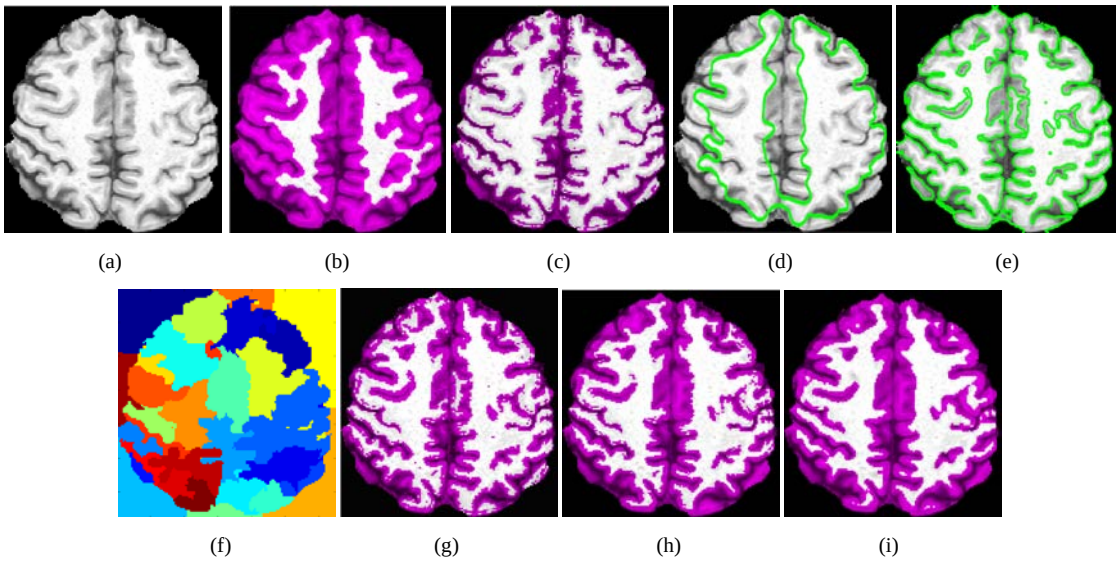
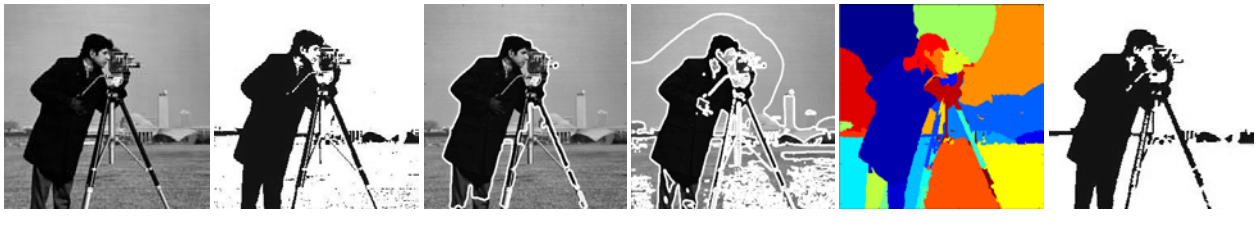


Fig.5. Original MRI brain image and the segmentation results. (a) Original MRI image; (b) segmentation by conventional MRF; (c) segmentation by Otsu thresholding; (d) segmentation by LSEWRI; (e) segmentation by RACM; (f) segmentation by MNCut; (g) segmentation by BEBS; (h) segmentation by HDSS-1; (i) segmentation by HDSS-2.

From the figures, we can see that the conventional MRF-based method (b) is conservative, as mentioned before. It puts more weight on pixel label homogeneity so misses many fine structures of the WM. Otsu (c) yields an unsuitable threshold such that some GM and CSF regions are segmented into the

WM part. LSEWRI (d) produces a large number of mis-segmentations. RACM (e) includes some GM into WM and also yields segments with very small size in the right half part of the brain image. MNCut (f) fails to characterize the fine structures of the WM and GM. The result from BEBS (g) is satisfactory, but it also yields some isolated pixels as well as some false positives in the bottom and boundary of the image. In contrast, HDSS-1 (h) and HDSS-2 (i) produce better segmentations in terms of reduced number of isolated pixels and the strengthened homogeneity of the neighboring pixel labels.

Fig. 6 shows the results of segmenting the Cameraman image. For visualization purposes, the results of the Otsu and HDSS-1 algorithms are shown as a binary image. From the figures, we can see that Otsu (b) yields a good segmentation of the human body, but generates many mis-segmentations in the background building and meadow parts. Most of them are segments with small sizes. LSEWRI (c) generates no mis-segmentation in the meadow but misses part of the human legs and also segments out a large area of background between the man and camera. RACM (d) yields good segmentation of the man. It even finds some fine structures of the image. But this method generates a large number of mis-segmentations in the meadow. We adjusted the number of segments of the MNCut method (e), and observed that it generates many homogeneous patches. Obviously, some post-processing algorithms, like region merging, need to be applied to find the human from these patches. (f) shows the result of the presented HDSS-1 method. We can see that the man is segmented out with fewer mis-segmentations, especially in the meadow part, when compared to other methods. Also, a smaller portion of the camera is segmented out by this method.



(a) (b) (c) (d) (e) (f)

Fig.6. Original cameraman image and the segmentation results. (a) Original image; (b) segmentation by Otsu thresholding; (c) segmentation by LSEWRI; (d) segmentation by RACM; (e) segmentation by MNCut; (f) segmentation by HDSS-1.

VII. CONCLUSION

Image segmentation is a very important but challenging problem for computer vision and image analysis. This paper developed an image segmentation approach driven by HVS properties, where the objective function for segmentation was designed by considering the preference of HVS to good segmentation from both region-based and boundary-based perspectives. The metrics were encoded into the MRF and the JND model was used to calculate the contrast of the image contents. Comparative performance evaluation was carried out via the experiments between the three variations of the presented algorithm and several representative segmentation and clustering algorithms available in the literature. The results show that the presented algorithms resulted in highly encouraging performance in terms of segmentation efficiency, robustness and convergence speed.

There are many future research directions that are worth pursuing. An investigation on the segmentation with multiple region types or multiple pixel labels, instead of the binary labels, 0 and 1, as discussed in this paper, will be an interesting extension to the algorithm. This would, of course, require the presented objective function to be adjusted accordingly. In our current work, we mainly discussed the problem of hard or crisp segmentation, that is, a pixel belongs to either region 0 or region 1. Future research on the fuzzy segmentation, based on the objective function presented in this paper, will be another interesting research topic. Perhaps, the scheme of designing the objective function for fuzzy MRF, as discussed in [47], will be very helpful. Finally, in this paper we employed ICM for the optimization task, and the research on approaches which can further improve the computational speed and segmentation accuracy will be both theoretically and practically useful. The graph-cut based methods [4][43] may be promising options.

APPENDIX

In this appendix, we show that the multi-pie slice configuration, with $N_{PS}(s)$ pie slices, may have a large angle value which can be obtained by adding the angles of the $N_{PS}(s)$ pie slices together, but will

unnecessarily result in a smooth boundary and thus a lower energy due to the exponential function in (19).

We assume that the cross-boundary contrast and the interior contrast of each pie slice are the same as each other. Then (19) can be written as

$$\phi(s)=\sum_{i=1}^{N_{PS}(s)}\exp\left[-a_3\cdot\alpha_i(s)-a_4\cdot Contr(s)_{CB-i}+a_5\cdot Contr(s)_{PS-i}\right]=\exp\left[Contrast(s)\right]\sum_{i=1}^{N_{PS}(s)}\exp\left(-a_3\cdot\alpha_i(s)\right) \tag{25}$$

where $Contrast(s)=-a_4\cdot Contr(s)_{CB-i}+a_5\cdot Contr(s)_{PS-i}$. Suppose we have a favorite segmentation with a single pie slice in the second-order neighborhood system of *s*. Its contrast is equal to $Contrast(s)$, and its angle value, $\alpha_{favorite}(s)$, equals the summation of the angle values in (25), i.e., $\alpha_{favorite}(s)=\sum_{i=1}^{N_{PS}(s)}\alpha_i(s)$. Then, it is not difficult to prove that

$$\exp\left(-a_3\cdot\alpha_{favorite}(s)\right)=\exp\left(-a_3\cdot\sum_{i=1}^{N_{PS}(s)}\alpha_i(s)\right)\le\sum_{i=1}^{N_{PS}(s)}\exp\left(-a_3\cdot\alpha_i(s)\right) \tag{26}$$

given that $\alpha_i(s)\ge 0$. Therefore, we have

$$\phi(s)_{favorite}\le\phi(s) \tag{27}$$

Thus, the favorite segmentation has lower energy than the multi-pie slice configuration and therefore has a higher probability to survive.

REFERENCES

[1] Y. Zhang (Editor), *Advances in Image And Video Segmentation*. IRM Press, USA, May 2, 2006.

[2] P. K. Sahoo, S. Soltani, A. K. C.Wong, and Y. C. Chen, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing archive*, vol. 41, no. 2, Feb., 1988.

[3] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag, 2001.

[4] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, Feb. 2004.

[5] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 1549–1560, Apr. 1995.

[6] D. Mumford and J. Shah, “Optimal approximation by piecewise smooth functionals and associated variational problems,” *Commun. Pure Appl. Math.*, vol. 42, pp. 577–685, 1989.

[7] D. Pham, C Xu, and J. Prince, “A survey of current methods in medical image segmentation,” *Annual Review of Biomedical Engineering*, vol. 2, pp. 315-337, 2000.

[8] N.R.Pal and S.K.Pal, “A review on image segmentation techniques,”*Pattern Recognition*,vol.26, no.9, pp. 1277-1294, 1993.

[9] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[10] Y. Ma, H. Derksen, W. Hong, and J. Wright, “Segmentation of multivariate mixed data via lossy coding and compression,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 29, no. 9, pp. 1546-1562, 2007.

[11] J. Canny, “A computational approach to edge detection”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[12] L. Shapiro and G. Stockman, *Computer Vision*. Prentice-Hall, Inc. 2001.

[13] D. Geman, “Stochastic model for boundary detection,” *Image and Vision Computing*, vol. pp. 61-65, 1987.

[14] D. Geman, S. Geman, Chr. Graffigne, and P. Dong, “Boundary detection by constrained optimization,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 609-628, 1990.

[15] D. Geman and S. Geman, “Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721-741, 1984.

[16] M.Kass, A.Witkin,and D. Terzopoulos,“Snakes: active contour models,”*Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[17] D. Cremers, M. Rousson, and R. Deriche, “A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape,” *Int. J. Comput. Vis.*, vol. 72, no. 2, pp. 195–215, 2007.

[18] J. Freixenet, X. Munoz, D. Raba, J. Marti and X. Cuffi, “Yet another survey on image segmentation: region and boundary information integration,” in *Proc. of the 7th European Conference on Computer Vision-Part III*, pp. 408 – 422, 2002.

[19] X. M. Pujol, “Image segmentation integrating color, texture and boundary information,” *Ph.D. Thesis*, University of Girona, Spain, 2003.

[20] J. Fan, D.K.Y. Yau, A.K. Elmagarmid, and W.G. Aref, “Automatic image segmentation by integrating color-edge extraction and seeded region growing,” *IEEE Trans. Image Processing*, vol. 10, no. 10, pp. 1454 – 1466. 2001.

[21] J. Wu and A. C. S. Chung, "A segmentation model using compound Markov random fields based on a boundary model," *IEEE Trans. Image Processing*, vol. 16, no. 1, Jan. 2007.

[22] M. S. Alili and D. Ziou, "An approach for dynamic combination of region and boundary information in segmentation," in *Proc. 19th International Conference on Pattern Recognition*, 2008.

[23] Z. Tu and S.-C. Zhu, "Image segmentation by data-driven Markov chain Monte Carlo," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 657–673, May 2002.

[24] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[25] S. X. Yu and J. Shi, "Multiclass spectral clustering," in *Proc. IEEE Int. Conf. Computer Vision*, 2003, vol. 1, pp. 313–319.

[26] B. Kim and D. Park, "Novel non-contrast-based edge descriptor for image segmentation," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 9, Sep. 2006.

[27]] J. Chen, T. N. Pappas, A. Mojsilovic´, and B. E. Rogowitz, "Adaptive perceptual color-texture image segmentation," *IEEE Trans. Image Processing*, vol. 14, no. 10, Oct. 2005.

[28] J. Zhang and D. Wang, "Image segmentation by multi-grid Markov random field optimization and perceptual considerations," *Journal of Electronic Imaging*, vol. **7**, 52, 1998.

[29] P.A.Devijver, and J.Kittler, *Pattern Recognition: A Statistical Approach*. Prentice Hall, Englewood Cliffs, London, 1982.

[30] A. Murat Bagct, Y. Yardimci, and A. E. Cetin, "Detection of micro-calcification clusters in mammogram images using local maxima and adaptive wavelet transform analysis," in *Proc. Of International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*, vol.4, pp. IV-3856- IV-3859, 2002.

[31] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Sys., Man., Cyber*, vol. 9, pp.62–66, 1979.

[32] P. Liao, T.Chen, and P. Chung, "A fast algorithm for multilevel thresholding," *J. Inf. Sci. Eng.*, vol. no. 5, pp.713–727, 2001.

[33] C. Li, C. Xu, C. Gui, and M.D Fox, "Level set evolution without re-initialization: a new variational formulation," in *Proc. IEEE Computer Society Conference on Computer Vision and Patter Recognition (CVPR)*, vol. 1, pp. 430- 436, 2005.

[34] C. Li, C. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Processing*, vol. 17, no.10, pp.1940-1949, 2008.

[35] T. Cour, F. Benezit, and J. Shi, "Spectral segmentation with multi-scale graph decomposition," in *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[36] Gerhard Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (2nd edition)*. Springer, Feb., 27, 2006.

[37] W. Lin, "Computational models for just-noticeable difference," in *H. R. Wu and K. R. Rao (Editors), Digital Video Image Quality and Perceptual Coding*, ,Taylor & Francis, 2005.

[38] Y. Jia, W.Lin, and A. A. Kassim, "Estimating just-noticeable distortion for video," *IEEE Trans. Circuits and System for Video Technology*, vol. 16, no. 7, pp. 820-829, Jul. 2006.

[39] X. Zhang, W. Lin, and P. Xue, "Just-noticeable difference estimation with pixels in images," *Journal of Visual Communication and Image Representation (ELSEVIER)*, vol. 19, pp. 30-41, 2008.

[40] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Trans. Circuits and Systems for Video Technology*, vol., 15, no. 7, pp. 900-909, Jul. 2005.

[41] Z. Liu, L. J. Karam, and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," *IEEE Trans. Image Processing*, vol. 15, no. 7, pp. 1763-1778, Jul. 2006.

[42] X. Yang, W. Lin, Z. Lu, E. P. Ong, and S. Yao, "Just-noticeable-distortion profile with nonlinear additivity model for perceptual masking in color images," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, pp. 609-612, 2003.

[43] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.

[44] [online]. Available: http://marathon.csee.usf.edu/Mammography/Database.html

[45] W. M. Diyana, J. Larcher, and R. Besar, "A comparison of clustered micro-calcifications automated detection methods in digital mammogram," in *Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, Page(s): II - 385-8 6-10, Apr., 2003.

[46] J. Besag, "On the statistical analysis of dirty pictures" (with discussions), *Journal of the Royal Statistical Society, Series B*, vol. 48, pp. 259-302, 1986.

[47] F. Salzenstein and W. Pieczynski, "Parameter estimation in hidden fuzzy Markov random fields and image segmentation," *Graphical Models and Image Processing*, vol. 59, no. 4, pp. 205–220, Jul., 1997.