A Bayesian Information Flow Approach to Image Segmentation

Akshaya Mishra, Alexander Wong, David A. Clausi, and Paul Fieguth *

Abstract

A novel Bayesian information flow approach is presented for accurate image segmentation, formulated as a maximum a posteriori (MAP) problem as per the popular Mumford-Shah (MS) model. The model is solved using an iterative Bayesian estimation approach conditioned on the flow of information within the image, where the flow is based on inter-pixel interactions and intra-region smoothness constraints. In this way, a localized and accurate Bayesian estimate of the underlying piece-wise constant regions within an image can be found, even under high noise and low contrast situations. Experimental results using 2-D images show that the proposed Bayesian information flow approach is capable of producing more accurate segmentations when compared to state-of-the-art segmentation methods, especially under scenarios with high noise levels and poor contrast.

1 Introduction

Image segmentation has always been a fundamental problem of great interest in the field of computer vision. The success of many high level image processing algorithms such as image analysis [5], surveillance [6], robotic vision and navigation is highly dependent upon the accuracy of image segmentation algorithms. The pixels of an image are often grouped into number of clusters based upon either some region similarity or discontinuity criterion, however there is substantial overlap between these two criteria [14, 15].

Among several existing segmentation techniques [2, 3, 4, 13, 7, 16, 11, 14, 15, 10], Mumford-Shah's [7] (MS) method is a theoretical segmentation model and have been approximated by many researchers [1, 3] for practical segmentation proposes. MS introduced a global model for representing segmented images as piece-wise constant regions. The main difficulty with using the Mumford-Shah approach lies in the complexity of solving the global Mumford-Shah model. To address this issue, Chan and Vese [3] proposed a simplified Mumord-Shah model which approximates the Mumford-Shah model using two piece-wise constant regions, and employs a level set framework to solve for the constant regions. Unfortunately, the Chan-Vese model suffers from the computational expensive re-initialization [3] problems due to the fact that the optimization process is carried out in a level set framework. To overcome the problem associated with Chan-Vese model while finding a global optimum for the Mumford-Shah model, Bresson et al. [1] proposed FAC (fast active contour), a fast global optimization approach for solving the Chan-Vese model without using the level set framework. Hence, FAC does not suffer from the same initialization problems faced by [3]. However, since FAC utilizes a global model, it often fails to segment small homogeneous regions within larger regions. Existing literature to solve MS model are sensitive to local minima, not able to handle embedded topology, not able to identify small regions and complicated to implement.

To address the problems associated with the existing MS model based image segmentation methods, a novel Bayesian information flow (BIF) based technique to solve the MS model is proposed, which is the main contribution of this paper. The BIF is designed specifically to avoid the issues associated with re-initialization and poor separation of small regions faced by existing region-based approaches. The BIF approach solves the Mumford-Shah model using an iterative Bayesian estimation approach to obtain the piecewise constant region. The concept of Bayesian insulation is introduced to adaptively control the flow of information based on the underlying information in the image with continuity on slope as a prior, and is integrated into a Bayesian linear least squares estimation framework to approximate the piecewise constant regions of a given image. The BIF can be interpreted as a scale space representation of an image, where the last scale represents the piece-wise smooth segmented region of the given image.

*Akshaya Mishra, Alexander Wong, David A. Clausi, and Paul Fieguth are with U. of Waterloo, Waterloo, Canada N2L 3G1. {akmishra,a28wong,dclausi,pfieguth}@uwaterloo.ca

978-0-7695-4040-5/10 $26.00 © 2010 IEEE
DOI 10.1109/CRV.2010.46

The paper is organized as follows. Section 2 provides a summary of the Mumford-Shah model. In Section 3, the image segmentation problem is formulated in a Bayesian sense. An iterative Bayesian estimation approach to solving the image segmentation problem is described in Section 4. The BIF theory is formalized in Section 5. A comparative analysis between BIF and two state-of-the-art methods is performed and experimental results are described in Section 6.

2 Background

Image segmentation can be considered as the optimization of a non-convex energy functional, which typically consists of a measurement term and a prior term. The Mumford-Shah (MS) model [7] represents an image as a set of disjoint piece-wise constant regions. The total energy E^{MS} of the model can be expressed as

$$E^{MS} = \alpha \int\int_{\Omega} |u-f|^2 dA + \beta \int\int_{\Omega \setminus C} |u'|^2 \, dA + \gamma \oint_{C} ds, \tag{1}$$

where $\Omega \subset R^N$ is an open set representing the image domain, f is the observed image, u is the underlying piece-wise constant segmented image, u' is the first derivative of u, C is the boundary of the segmented regions, and parameters α, β and γ are positive weighting constants. The first term of (1) is generally referred to as the data fidelity term and can be considered as the observation for f. The second term of (1) represents the smoothness constraints and can be interpreted as the prior model for u given f. Finally, the third term of (1) asserts a penalty on boundary arclength, forcing boundary smoothness by minimizing the total boundary length. Given these three terms, the MS model provides the theoretical basis for the segmentation of an image into a set of piece-wise constant regions.

Although MS is a very appealing theoretical model, from a practical perspective the minimization of (1) is very challenging. While simplifications have been made by Chan and Vese [3] and Bresson et al. [1] to reduce complexity and to allow for faster convergence rates, re-initialization [3] and poor segmentation of small homogeneous regions [1] problems still exist. The goal of the proposed work is to alleviate these issues through the use of a Bayesian approach to essentially solve the MS minimization problem to account for both local and global characteristics of the image. This is accomplished through the novel concept of Bayesian information flow in a scale space frame work.

3 Problem Formulation

Let X denote a set of sites within a discrete lattice $\mathcal{L}$ and $x \in X$. Let $F = \{F_x | x \in X\}$, $U = \{U_x | x \in X\}$ and $V = \{V_x | x \in X\}$ denote random fields on X, where F_x, U_x, and V_x are random variables taking on values representing the observation state, piece-wise constant state, and high frequency state, respectively, at site x. Let $f = \{f_x | x \in X\}$, $u = \{u_x | x \in X\}$, and $v = \{v_x | x \in X\}$ denote realizations of F, U, and V respectively. Assuming a scale separation of u and v on the basis of degrees of smoothness as implied by respective priors $u_\infty \sim \aleph(0, P_\infty)$, $v_{t-1} \sim \aleph(0, \Sigma_{t-1})$, the set of observations f can be decomposed into a piece-wise constant state u_∞ and a set of high frequency states $V = \{v_1, v_2, \ldots, v_\infty\}$

$$f = u_\infty + \sum_{i=1}^{\infty} v_i. \tag{2}$$

The high frequency states play an important role in multi-scale based image analysis and can be used in scale-space analysis, key feature extraction and denoising, however in the context of image segmentation, the underlying goal is to solve for u_∞ given f. To estimate u_∞ given f, we can formulate the problem in a continuous domain from a Bayesian sense as a maximum a posteriori (MAP) problem,

$$u_\infty = \arg\max_u p\left(u|f\right). \tag{3}$$

Where the posterior distribution $p\left(u|f\right)$ can be defined as

$$p\left(u|f\right) \propto \underbrace{p\left(u\right)}_{prior} \underbrace{p\left(f|u\right)}_{measurement}. \tag{4}$$

In the context of active contours [1, 3, 2], (4) can be re-expressed, by taking the log on both sides of (4), as

$$\underbrace{-\log\left(p\left(u|f\right)\right)}_{Total\ energy} \propto \underbrace{\left(-\log\left(p\left(u\right)\right)\right)}_{Internal\ energy} + \underbrace{\left(-\log\left(p\left(f|u\right)\right)\right)}_{External\ energy}, \tag{5}$$

where $-\log\left(p\left(f|u\right)\right)$ is the external energy or data fidelity term as defined by

$$-\log\left(p\left(f|u\right)\right) = \alpha \int\int_{\Omega} |u-f|^2 dA, \tag{6}$$

$-\log\left(p\left(u\right)\right)$ is the internal energy and, in the context of active contours [1, 3, 2, 12], is defined as

$$-\log\left(p\left(u\right)\right) =$$

$$\beta_1 \int\int_{\Omega \setminus C} \left|u'\right|^2 dA + \beta_2 \int\int_{\Omega \setminus C} \left|u''\right|^2 dA + \gamma \left| \oint_{C} ds \right|^2, \tag{7}$$

where u' and u'' are the first and second central derivatives of u. Omitting the second order term from (7) makes (5) similar to the MS model (1) such that

$$E^{MS} = -\log\left(p\left(u|f\right)\right). \tag{8}$$

Therefore, minimizing E^{MS} or $-\log\left(p\left(u|f\right)\right)$ is equivalent to maximizing (4). One of the main issues associated with solving the MAP problem posed in (3) is that there is, in general, no closed form solution. An approximate solution to (3) can be obtained by assuming

$$f = u_\infty + V, \;\; u_\infty \sim \aleph(0, P_\infty), V \sim \aleph(0, \Sigma), \tag{9}$$

and then, by applying a Bayesian linear least square estimation to (9), an estimate of u_∞ [9] can be represented as

$$\hat{u}_\infty = \left(P_\infty^{-1} + \Sigma^{-1}\right)^{-1} \Sigma^{-1} f, \tag{10}$$

The main problem in obtaining $\hat{u}_\infty$ from (10) is that we do not know P_∞ and Σ. Therefore, we are adopting an iterative approach to estimate $\hat{u}_\infty$ by enforcing an adaptive smoothness prior P_{t-1} and identity noise covariance matrix Σ at each iteration.

4 Iterative Bayesian Estimation

One approach to solving (10) is to first decompose (2) into u_t and v_t, $t \in [1, \infty]$. Consider a scale separation between u and v based on an adaptive smoothness prior as follows:

$$\hat{u}_{t-1} = u_t + v_t, \;\; u_t \sim \aleph(0, P_t), v_t \sim \aleph(0, \Sigma_t). \tag{11}$$

An iterative estimation strategy based on a Bayesian linear least squares approximation [9, 8] is employed on (11), to estimate u_t as follows:

$$\hat{u}_t = \left(P_{t-1}^{-1} + I\right)^{-1} \hat{u}_{t-1}, \tag{12}$$

where P_{t-1} is the smoothness on u_t at iteration t, and the identity matrix I is the covariance of the high frequency state v_t. One of the issues associated with solving the problem using a Bayesian linear least squares approximation is that u is inherently nonlinear due to its piece-wise constant nature. Therefore we need a prior P_t that can model piece-wise constant regions u, and hence we will not able to find the inverse of P_{t-1}. To address this issue, we introduce the concept of Bayesian information flow for modeling the prior P_{t-1} to account for the nonlinear nature of u.

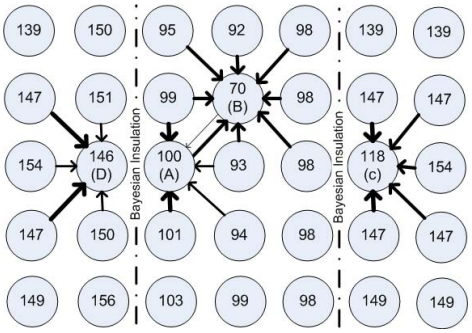

Figure 1. Example of information flow. The circle, embedded value, and arrow width represent the pixel location, pixel intensity and amount of information flow respectively. For clarity, flow is shown only for pixels A, B, C, and D, which represent four important concepts. Little information flows from pixel B to pixel A due to large intensity dissimilarity between them. Also, no information flow exists between the pixels to the left of pixel A and pixel A itself since they belong to different regions. A boundary with no information flow is an insulation boundary. Pixel B is an isolated point and under this circumstance, there is significant information flow from the surrounding pixels into pixel B since the intensity of pixel B can be considered as noise. The way that rough boundaries are handled is illustrated in pixel C, where the information from the pixels to the left of pixel C do not flow into pixel C even though pixel C is more similar to the left side pixel than the pixels on right and top of it to prevent the formation of a rough boundary.

5 Bayesian Information Flow Theory

The challenge with solving the aforementioned Bayesian estimation problem to account for the inherent nonlinear nature of u lies in modeling the prior P_{t-1} such that u can be approximated by a set of piece-wise linear functions. In general, the exact prior knowledge of the objects in an image is not known a priori. Therefore, a more intuitive and common approach is to utilize continuity in slope and curvature of image intensity and object boundary as prior constraints, as done by Chan and Vese [3] and Bresson et al. [1]. The key issue associated with relying solely on these prior constraints is that they do not account well for the nonlinear nature of u and as such result in poor inter-region separability. This ineffective inter-region separability often generates over and under segmented regions.

To address the important issue of poor inter-region separability, we shall introduce the theory of Bayesian information flow. The underlying concept of Bayesian information flow is to adaptively control the flow of information within an image such that inter-region separability is preserved during the iterative estimation process. Intuitively, based on the criteria of inter-region separability, the flow of information between two pixels should increase as the intensity similarity between the pixels increases, as intensity similarity is a good indicator of intra-region pixels. Furthermore, intuitively, there should be no flow of information between pixels from different regions. Motivated by these intuitions, the concept of information flow can be illustrated as shown in Fig. 1. Little information flows from pixel B to pixel A due to large intensity dissimilarity between them. Also, no information flow exists between the pixels to the left of pixel A and pixel A itself since they belong to different regions. The boundary that restricts no information flow will be referred to as an insulation boundary. Isolated points are handled as demonstrated by pixel B, where all neighboring points are different from itself. Under this circumstance, there is significant information flow from the surrounding pixels of pixel B into pixel B since the intensity of pixel B can be considered as noise. Handling of boundary smoothness is illustrated by pixel C, where the information from the pixels to the left of pixel C do not flow into pixel C even though pixel C is more similar to left side pixel than the pixels to the right and above of it.

Based on the aforementioned Bayesian information flow theory in conjunction with the prior constraints, the prior P_{t-1} can be expressed as [8]

$$P_t^{-1} = \kappa(L^T L), \tag{13}$$

where L is a penalty matrix (penalty on non-smooth region) and is represented as:

$$\|Lu_t\| = \int\int \left\{ \left| \frac{\partial u_t}{\partial x} \right|^2 + \left| \frac{\partial u_t}{\partial y} \right|^2 + \left| \frac{\partial^2 u_t}{\partial x^2} \right|^2 + \left| \frac{\partial^2 u_t}{\partial y^2} \right|^2 \right\} dA. \tag{14}$$

In discrete format L can be decomposed as:

$$L = [L_x; L_y; L_{xx}; L_{yy}]. \tag{15}$$

Where L_x can be expressed as:

$$L_x = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}, \tag{16}$$

L_y, L_{xx} and L_{yy} and can be expressed in the similar fashion. The κ is the information flow matrix that controls the flow of information between pixels and accounts for nonlinearity in u by putting insulation between the boundary of the objects. To determine κ, we first compute the potential information flow κ^* from site (a, b) to site (i, j) can be expressed as

$$\kappa^*_{i,j}(a, b) = \frac{1}{k_{\kappa^*}} \exp\left(-\left|f_{i,j} - u_{a,b}\right|\right), \tag{17}$$

where k_{κ^*} is a normalization constant to make $\kappa^*_{i,j}$ in the range [0,1]. Based on the distribution of the potential information flow $\kappa^*_{i,j}$, the insulation threshold T is defined to determine the information flow $\kappa_{i,j}$ to prevent the flow of information from unwanted measurements to the considered measurement as shown in Fig. 1,

$$\kappa_{i,j}(a, b) = \begin{cases} \kappa^*_{i,j}(a, b) & \text{if } \kappa^*_{i,j}(a, b) < T \\ 0 & \text{otherwise,} \end{cases} \tag{18}$$

For implementation purposes, the insulation threshold T is defined by the median of the image variance, which is calculated by first computing the variance of local region by sliding a window of 12×12, and then taking the median among all the computed local variance. To accommodate for boundary constraints, the BIF approach defines two additional constraints:

1. An isolated point should be merged with its surrounding neighbors, i.e.,

$$\kappa_{i,j} = 1 \text{ if } \forall(\kappa_{i,j}) = 0, \tag{19}$$

2. In cases where some elements of $\kappa_{i,j}$ are zero, a smoothness criterion is employed to evaluate

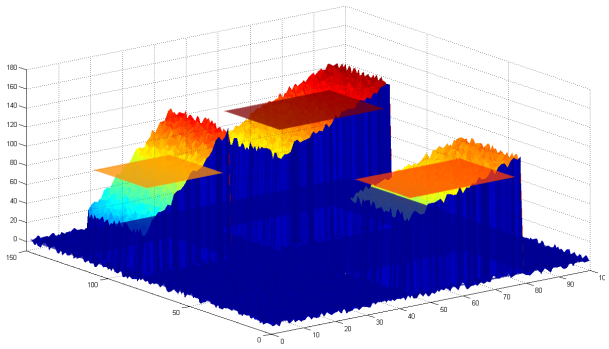

Figure 2. Illustration of Bayesian information flow based image segmentation for a synthetic image contaminated with additive Gaussian noise. The continuous surface plot of the image along with the segmented piece-wise constant region (three self-similar segmented regions) are plotted for demonstration of BIF. It can be observed that BIF is able to identify three constant regions for the three distinct objects and one constant region for the background, while perfectly preserving inter-region separability.

which neighbor elements are influencing the current measurement. Let $\aleph$ and $\aleph_c$ denote the set of neighbor elements having non-zero and zero values of $\kappa_{i,j}$ respectively. The influence of $\aleph$ and $\aleph_c$ on the current measurement, s_1 and s_2 respectively, can then be computed as the integral of the potential information flow κ^* normalized by arc-length,

$$s_1 = \frac{\int\limits_{\aleph} \kappa^*_{i,j} d\aleph}{\int\limits_{\aleph} ds}, \quad s_2 = \frac{\int\limits_{\aleph_c} \kappa^*_{i,j} d\aleph_c}{\int\limits_{\aleph_c} ds}. \tag{20}$$

The value $\kappa_{i,j}$ can then be determined as,

$$\kappa_{i,j} = 1 - \kappa_{i,j} \text{ if } s_1 < s_2 \tag{21}$$

An illustration of Bayesian information flow as applied to a synthetic image contaminated with additive Gaussian noise is shown in Fig. 4, which is displayed as a surface plot to better illustrate the convergence to constant regions. The BIF approach is able to identify three constant regions for the three distinct objects and one constant region for the background, while perfectly preserving inter-region separability. The algorithm for the BIF is provided in Algorithm. 1. Note that computing $\left(P_{t-1}^{-1} + I\right)^{-1}$ for a large image is not feasibility due to computational and numerical reasons. Therefore, BIF uses a modified conjugate gradient technique to solve (12).

Algorithm 1 $[u_t]$ = Function BIF(f)

1: $t = 0, \hat{u}_t = f$, $t = t + 1, \hat{u}_t = 0$
2: Compute the L matrix using (15)
3: **while** $\hat{u}_t \neq \hat{u}_{t-1}$ **do**
4: Compute the information flow matrix κ.
5: Compute P_{t-1}^{-1} from (13)
6: Compute $\hat{u}_t = \left(P_{t-1}^{-1} + I\right)^{-1} \hat{u}_{t-1}$ (12)
7: t = t+1
8: **end while**

6 Experiments

The goal of this section is to investigate the effectiveness of the proposed BIF segmentation approach at producing well-localized and accurate segmentations. To achieve this goal, a set of experiments is performed involving the segmentation of synthetic images and real clinical images that are shown in the first column of Fig. 4. The set of synthetic images consists of two images depicting rice and stone respectively (rows 1 and 2 of Fig. 4), and are designed to test BIF's ability to deal small homogenous regions in the presence of varying illuminance. The set of real clinical images consists of two different computed tomography (CT) head images (rows 3 and 5 of Fig. 4), and a positron emission tomography (PET) head image (row 4 of Fig. 4). This set of clinical images are designed to test the BIF method's ability to handle real world data with complex embedded regions (in the case of the CT images) and a large number of artifacts (in the case of the PET image).

For comparison purposes, two state-of-the-art image segmentation methods were also evaluated. The tested segmentation methods include the fast region based global minimization approach proposed by Bresson et al. [1], denoted as FAC, and the level set active contour without re-initialization approach proposed by Chunming et al. [4], denoted as LS. FAC was chosen as it shares the same underlying Mumford-Shah model as BIF and thus provides a good comparison of different approaches to solving Mumford-Shah. LS was chosen as it addresses the shortcomings related to re-initialization faced by level-set methods. All tested segmentation methods were implemented using the parameters proposed in the respective papers. For all experiments the value of β_1 was set to be 1.

6.1 Experiment 1: Segmentation Results

In the first set of experiments, BIF, LS [4], and FAC [1] were used to segment the sets of synthetic images and real clinical images (Fig. 4). Visually, the segmentation results produced using BIF (2^{nd} column

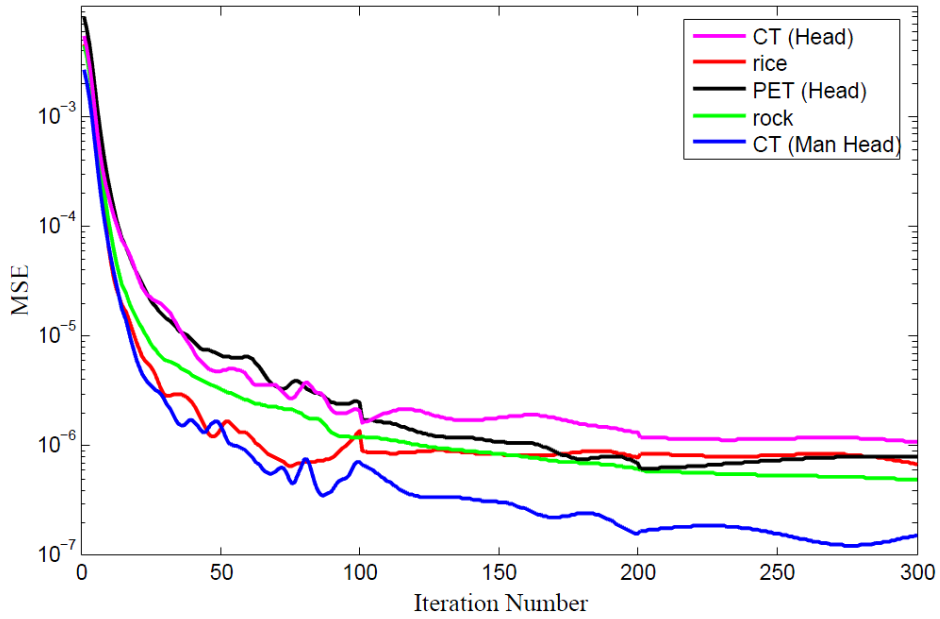

Figure 3. An illustration of convergence rate of BIF based image segmentation for five images. The y- axis represents the iteration-to-iteration mean square error (MSE) in intensity. MSE decreases in an exponential manner and plateaus after around 125 iterations for all cases.

Table 1. Execution Time (ET) (in seconds) of BIF, FAC [1], and LS [4].

	BIF	FAC	LS
Rice	75	50	891
Stone	240	170	1721
CT2 (Head)	89	64	948
PET (Head)	80	49	876
CT1 (Head)	77	45	905

of Fig. 4) are better localized and more accurate than the other tested methods. In the case of the rice image (1^{st} row of Fig. 4), the BIF method was able to segment each of the individual grains of rice with high accuracy, while the FAC method misses several grains of rice and the LS method produces very poor segmentation results. In the case of the stone image, both BIF and FAC are able to segment each of the individual stones, while LS produces the worst results of the tested images as LS is not able segment the individual stones as stones are packed very tightly together, and thus is treated as a single object from a topological perspective by LS. In the case of the PET head image, the individual regions are segmented with good locality and inter-region separability using BIF, while FAC and LS perform poorly. In particular, LS had significant problems dealing with the dominant streaking artifacts. This illustrates the effectiveness of BIF even under artifact-laden scenarios. Finally, in the case of the CT head images, BIF correctly segments small homogeneous regions when compared to the other two methods. For example, much of the embedded objects in the image is missed by FAC, while completely ignored by LS. These experiments demonstrate the effectiveness of BIF at providing accurate and well-localized segmentations under different imaging conditions.

6.2 Experiment 2: Convergence

In the second set of experiments, we empirically study the convergence rate of BIF to understand its convergence rate. The mean square error (MSE) at each iteration compared to the previous iterations over 300 iterations for the five test images and for three set of additive Gaussian noise with variance of 0, 0.001, and 0.005 (with a dynamic range of [0,1]) is calculated. The iteration-to-iteration MSE (intensity difference between previous and current iteration) with respect to the iteration number performed for the test images with additive Gaussian noise with variance of 0.005 are shown in Fig. 3 (for other test cases the results were similar). As expected, the change in MSE decreases in an exponential manner as the iteration

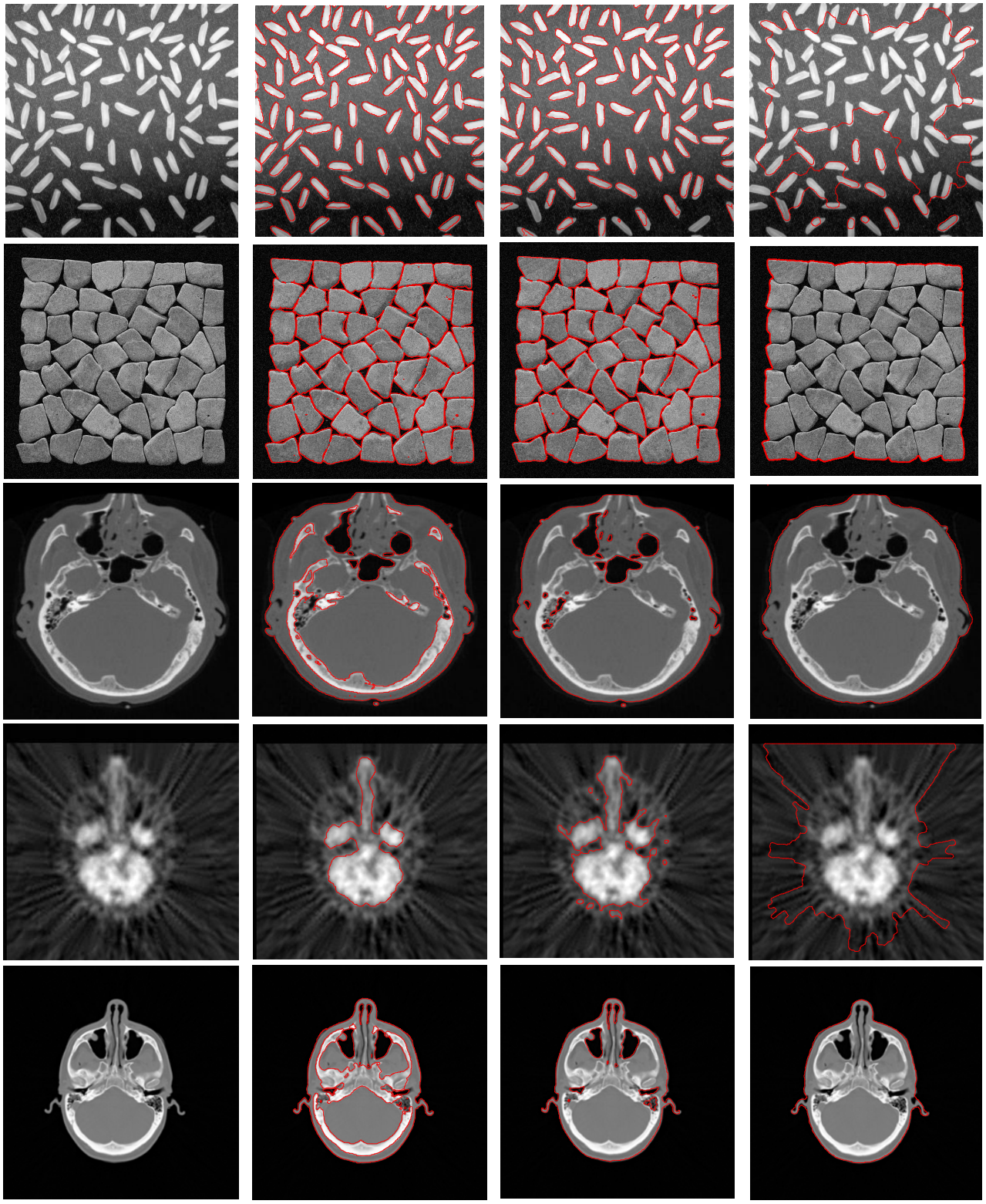

Figure 4. Performance of BIF compared to other methods across two synthetic and three real-world clinical images. First column shows the original image. $2^{nd}, 3^{rd}$, and 4^{th} columns illustrates the segmentation result of BIF, FAC [1] and LS [4].

number increases. In all cases, the MSE plateaus after 125 iterations, indicating that fast convergence rates can be achieved using BIF.

Furthermore, a quantitative analysis of the execution time of BIF compared to FAC and LS is provided in Table. 1. An unpaired two-tailed t-test was performed on BIF and LS when compared to FAC. The two-tailed p-values of BIF and LS compared to FAC are 0.4 and 0.001, respectively, implying that the execution time of BIF is statistically insignificant and the execution time of LS is statistically extremely significant when compared to that of FAC.

7 Conclusions and Future Work

This paper proposed a novel image segmentation algorithm based on Bayesian information flow (BIF) approach for accurately segmenting an image into number of piece-wise constant regions. BIF solves the Mumford-Shah model by employing an iterative Bayesian estimation technique with an information flow prior to robustly identify the piecewise constant regions of an image. The performance of BIF compared to other two state-of-the-art segmentation techniques is demonstrated on several natural and synthetic images. BIF provided better segmentation performance among its peers. Future research will involve extending BIF to multiple dimensions.

8 Acknowledgements

This research has been sponsored by the Natural Sciences and Engineering Research Council (NSERC) of Canada through individual Discovery Grants as well as GEOIDE (GEOmatics for Informed DEcisions) which is a Network of Centres of Excellence under NSERC.

References

[1] X. Bresson, S. Esedoglu, P. Vandergheynst, J. Thiran, and S. Osher. Fast global minimization of the active contour/snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151–167, 2007.

[2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–97, 1997.

[3] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[4] L. Chunming, X. Chenyang, G. Changfeng, and M. D. Fox. Level set evolution without re-initialization: A new variational formulation. In *CVPR*, volume 1, pages 430–436, 2005.

[5] A. Mishra, A. Wong, W. Zhang, D. Clausi, and P. Fieguth. Improved interactive medical image segmentation using enhanced intelligent scissors (eis). In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, August 2008.

[6] D. Mukherjee, N. Ray, and S. Acton. Level set analysis for leukocyte detection and tracking. *IEEE Transactions on Image Processing*, 13(4):562572, 2004.

[7] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *Computer Vision and Pattern Recognition*, pages 22–26, June 1985.

[8] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Companies, 3rd edition, 1991.

[9] W. Penny, J. Kilner, and F. Blankenburg. Robust Bayesian general linear models. *NeuroImage*, 36(3):661–671, 2007.

[10] L. Shafarenko, M. Petrou, and J. Kittler. Automatic watershed segmentation of randomly textured color images. *IEEE Transactions on Image Processing*, 6(11):1530–1544, November 1997.

[11] K. Siddiqi, S. Zucker, Y. Bérubé Lauzière, and A. Tannenbaum. Area and length minimizing flows for shape segmentation. In *Computer Vision and Pattern Recognition*, pages 621–627, Washington, DC, USA, 1997. IEEE Computer Society.

[12] G. Storvik. A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing. *TPAMI*, 16(10):976–986, 1994.

[13] H. Tek, F. Akova, and A. Ayvacy. Region competition via local watershed operators. In *Computer Vision and Pattern Recognition*, pages II: 361–368, 2005.

[14] Q. Yu and D. Clausi. SAR sea-ice image analysis based on iterative region growing using semantics. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12):3919–3931, 2007.

[15] Q. Yu and D. Clausi. IRGS: image segmentation using edge penalties and region growing. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 30(12):2126–2139, 2008.

[16] S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/mdl for multiband image segmentation. *PAMI*, 18(9):884–900, 1996.