# Extraction of Breast Lesions from Ultrasound Imagery: Bhattacharyya Gradient Flow Approach

Mahsa Torkaman[a], Romeil Sandhu[b], and Allen Tannenbaum[a]

[a]Departments of Computer Science and Applied Mathematics & Statistics , Stony Brook University, Stony Brook, NY

[b]Department of Biomedical Informatics, Stony Brook University, Stony Brook, NY

## ABSTRACT

Breast cancer is one of the most commonly diagnosed neoplasms among American women and the second leading cause of death among women all over the world. In order to reduce the mortality rate and cost of treatment, early diagnosis and treatment are essential. Accurate and reliable diagnosis is required in order to ensure the most effective treatment and a second opinion is often advisable. In this paper, we address the problem of breast lesion detection from ultrasound imagery by means of active contours, whose evolution is driven by maximizing the Bhattacharyya distance[1] between the probability density functions (PDFs). The proposed method was applied to ultrasound breast imagery, and the lesion boundary was obtained by maximizing the distance-based energy functional such that the maximum (optimal contour) is attained at the boundary of the potential lesion. We compared the results of the proposed method quantitatively using the Dice coefficient (similarity index)[2] to well-known GrowCut segmentation method[3] and demonstrated that Bhattacharyya approach outperforms GrowCut in most of the cases.

**Keywords:** ultrasound imagery, breast lesion detection, Bhattacharyya distance (BD), GrowCut segmentation, Dice coefficient

## 1. INTRODUCTION

Approximately 1 in 8 American women will develop invasive breast cancer over their lifetime and in spite all efforts and existing technologies, many may not attain long-term remission. Early stage diagnosis is clearly key to improve the cure rate significantly, and the ultrasound modality can be utilized to facilitate this task.

Mammography is among the most effective modalities for detecting and diagnosing breast cancer and is considered to be the "gold standard" for evaluation of breast lesions. However, due to its limitations and drawbacks such as low specificity, can result in unnecessary biopsy operations.[4] Therefore, ultrasound imagery has become one of the important alternative to mammography and has shown increasing interest among both radiologists and researchers for the task of breast lesion detection and classification.[4] Ultrasound can increase overall cancer detection and reduce the number of unnecessary biopsies which is costly and can cause patient anxiety. However as noted in,[5] "it is worthwhile noting that despite all of these advances, it is still the case that no single imaging modality is capable of identifying and characterizing all breast abnormalities and a combined modality approach will continue to be necessary".

CAD systems by providing additional information to radiologists can improve the operator independency of diagnosis and reduce the inter-observer variation rate. In addition, utilizing CAD systems can increase both the detection of early-stage malignancies and recall rate without making any change in positive predictive value for biopsy.[4]

Lesions detection in ultrasound breast imagery is usually driven by a segmentation procedure. The main goal of image segmentation is to decompose an image into its constituent components, each of which associated with a distinct class.[6] Among all diverse image segmentation methods, it is known that some of the most successful ones in practice are ones that minimize the probability of misclassification error.[6] However, one of the drawbacks

Further author information: (Send correspondence to M.T.: E-mail: mtorkaman@cs.stonybrook.edu, Telephone: 1 631 974 6048)

of this type of approach is that usually we do not have much information about the statistics of segmentation classes and consequently misclassification probability may become very complicated to explicitly define. It has been shown that distance measures between probability distributions can provide reasonable alternative criteria, in the sense that they are simpler to formulate and have comparable accuracy to methods which minimize the probability error.[6] Distance between two distributions can be defined in different ways utilizing various measures in probability theory. However, our main purpose is to avoid the potential instability associated with some of these standard measures such as mutual information (MI) and Chi-squared (CS).[7] Bhattacharyya distance is among the standard choices for defining a distance between probability distributions and it has been shown that it provides better results compared to other distant measurements in different applications such as segmentation, object tracking, and signal selection problems.[6,8–12]

In this work, we exploited the geometric active contours technique whose evolution was derived by employing the Bhattacharyya based energy functional. The level set technique (which can handle curves topological changes automatically) was employed for the implementation of the curve evolution.[13] Here we assume that the number of classes for which probability distribution is computed is two (the selected lesion and the background), and the evolution equation is derived through maximizing the level set implementation of the Bhattacharyya distance functional. In order to make the evolution more stable, we employed the regularized form of the evolution equation, which is defined by constraining the length of the active contours (section 2.1). We used Dice coefficient (similarity index)[2] in order to compare the proposed method to well-known GrowCut segmentation method.[3] GrowCut was chosen for the purposes of comparison since it is the default segmentation method now used in 3D Slicer. We have demonstrated both quantitatively and visually that the proposed method in most of the cases is closer to the reference ultrasound images compared to the result of GrowCut segmentation.

## 2. BHATTACHARYYA CURVE EVOLUTION

### 2.1 Cost Functional

In this section we sketch the mathematical details of the level-set formulation of Bhattacharyya cost functional. In order to facilitate the computations related to derivation of the evolution flow, we assume that domain of the given image $I(x)(x \in \Omega)$ is composed of two classes (object of interest and background). These classes can be represented by the characteristic functions $v$:

$$\begin{aligned} v_{in}(x) &= H(-\Phi(x)) \\ v_{out}(x) &= H(\Phi(x)) \end{aligned} \tag{1}$$

where $H$ is the Heaviside function and $\Phi : \Omega \to R$ is the level set function. Specifically,

$$H(\epsilon) = \{1, \ \text{if } \epsilon \geq 0 \text{ and } 0, \text{ if } \epsilon < 0\} \tag{2}$$

It is possible to extend this methodology to images with multiple homogenous regions (multiclass).[6]

The level set technique was employed as a numerical solution to the evolution equation.[14] One of the advantages of level sets is that they can handle topological changes of the evolving contour, such as merging and breaking, naturally. The main idea of underlying this approach is to define the evolving contour as the zero level set of the graph of a surface.

We first need to transform the input image into a vector-valued image $J$ by applying a proper transformation function on the image. The simplest choice for a transformation function is the identity function in which the feature image $J$ is same as the input image itself. Partial derivatives (gradient transform), multiresolution partial derivatives of the image (wavelet transform),[15] vector field of local displacements of the gray levels,[16] the local moments,[17] and local fractal dimension[18] are among other possible choices for the transformation feature. Using the gradient as a transformation function is briefly discussed in appendix A.

In our case, the optimal solution $\Phi^*(z)$ $(z \in R^n)$, which minimizes the information between the lesion and its background is given as follows:

$$\Phi^*(x) = \arg \inf_{\Phi(z)} \{\widehat{B}(\Phi(x))\} \tag{3}$$

where $\widehat{B}$ is the *Bhattacharyya coefficient* defined by

$$\widehat{B}(\Phi(x)) = \int_{z \in R^N} \sqrt{P_{in}(z|\Phi(x))\ P_{out}(z|\Phi(x))} dz. \tag{4}$$

The Bhattacharyya distance is a symmetric distance between probability distributions and is defined as $-\log$ of the Bhattacharyya coefficient. Minimizing the Bhattacharyya coefficient is same as maximizing Bhattacharyya distance since log is an increasing function. Note that $P_{in}$ and $P_{out}$ are probability density functions of predefined image features. These probability densities may be computed using a kernel-based estimation method:[19]

$$P_{in}(Z|\Phi(x)) = \frac{\int_\Omega K_{in}(z - J(x)) v_{in}(x) dx}{\int_\Omega v_{in}(x) dx} \tag{5}$$

$$P_{out}(Z|\Phi(x)) = \frac{\int_\Omega K_{out} v_{out}(x) dx}{\int_\Omega v_{out}(x) dx}, \tag{6}$$

where, $K_{in}$ and $K_{out}$ are two normalized scalar-valued functions known as kernels. The Gaussian kernel is one of the most frequently used kernels in the estimation of probability densities.[19,21] The cost functional (4) without any regularization is sensitive to noise caused by the measurement or errors in the data. In order to decrease this sensitivity, the solution was regularized by constraining the length of the active contour. The resulting expression is given by:

$$\Phi^*(x) = \arg\inf_{\Phi}\{\widehat{B}(\Phi(x)) + \alpha \int_\Omega ||\nabla H(\Phi(x))|| dx\} \tag{7}$$

### 2.2 Derivation of Evolution Equation via Gradient Flow

In this section we explain the derivation of the evolution equation through maximizing the Bhattacharyya distance between PDFs of the defined classes. In order to minimize Equation 4, we need to compute the first variation of $\widehat{B}(\Phi(x))$ with respect to $\Phi(x)$.

$$\frac{\delta \widehat{B}(\Phi(x))}{\delta(\Phi(x))} = \frac{1}{2} \int_{z \in R^N} (\frac{\delta P_{in}(z|\Phi(x))}{\delta \Phi(x)} \sqrt{\frac{P_{out}(z|\Phi(x))}{P_{in}(z|\Phi(x)))}} + \frac{\delta P_{out}(z|\Phi(x))}{\delta \Phi(x)} \sqrt{\frac{P_{in}(z|\Phi(x))}{P_{out}(z|\Phi(x))}} \tag{8}$$

where the derivative of $P_{in}$ and $P_{out}$ are:

$$\frac{\partial P_{in}(z|\Phi(x))}{\partial \Phi(x)} = \frac{\partial(\Phi(x))}{A_{in}} (P_{in}(z|\Phi(x)) - K_{in}(z - J(x))) \tag{9}$$

$$\frac{\partial P_{out}(z|\Phi(x))}{\partial \Phi(x)} = \frac{\partial(\Phi(x))}{A_{out}} (K_{out}(z - J(x))) - P_{out}(z|\Phi(x))) \tag{10}$$

$A_{in}$ and $A_{out}$ are areas inside and outside the contour respectively and are given by $\int_\Omega v_{in}(x) dx$ and $\int_\Omega v_{out}(x) dx$. Substituting $P_{in}$ and $P_{out}$ differentiations in the original equation (Equation 8) we have:

$$\frac{\partial \widehat{B}(\Phi(x))}{\partial \Phi(x)} = \delta(\Phi(x)) S(x) \tag{11}$$

$$\begin{aligned} S(x) =& \frac{1}{2} \widehat{B}(\Phi(x))(A_{in}^{-1} - A_{out}^{-1}) + \frac{1}{2} \int_{z \in R^N} K_{out}(z - J(x)) \frac{1}{A_{out}} \sqrt{\frac{P_{in}(z|\Phi(x))}{P_{out}(z|\Phi(x))}} dz \\ &- \frac{1}{2} \int_{z \in R^N} K_{in}(z - J(x)) \frac{1}{A_{in}} \sqrt{\frac{P_{out}(z|\Phi(x))}{P_{in}(z|\Phi(x))}} \end{aligned} \tag{12}$$

Following the computations above, the resulting gradient flow is:

$$\Phi_t(x) = \frac{\partial \widehat{B}(\phi(x))}{\partial \Phi(x)} = -\delta(\Phi(x)) S(x) \tag{13}$$

where, $t$ is an artificial time parameter and $\delta$ is the delta function.

In order to reduce the effect of noise, one can work with the regularized version of the optimal active contour (Equation 7). The gradient flow of the regularized version can be computed following the same procedure as above:

$$\Phi_t(x) = \delta(\Phi(x))(\alpha\kappa - S(x)) \tag{14}$$

Where, $\kappa$ is the curvature of the active contour. We implemented the regularized version of the evolution flow (Equation 14) in order to alleviate the sensitivity of the flow towards noise.

## 3. EXPERIMENTS AND RESULTS

In this section, we demonstrate some experimental results of the proposed segmentation method just described. The initial curve $C$ (which is defined by the user) will evolve towards the object of interest (the lesion in this case) according to Equation 14 and stops after a certain number of iterations at the boundary of the chosen lesion. We compared the results of the proposed method with the well-known GrowCut region growing algorithm. The GrowCut algorithm is an interactive segmentation algorithms that employs cellular automaton as the underlying method.[3] GrowCut was chosen for the purposes of comparison since it is the default segmentation method in 3D Slicer.

Figure 1 demonstrates a qualitative comparison between the ground truth image and the resulting segmentations generated by the proposed method and GrowCut method on a sample input data. As can be seen, the Bhattacharyya approach (green contour) outperforms the GrowCut algorithm (red contour) in capturing the lesion boundary. In order to compare the results quantitatively, we computed the Dice coefficient with respect to the ground truth for both approaches. The results are 0.73 ad 0.92 for the GrowCut and Bhattacharyya approaches, respectively 2. We ran the proposed approach on 20 sample ultrasound images obtained from database of ultrasound images of breast cancer provided by the Department of Radiology of Thammasat University and Queen Sirikit Center of Breast Cancer of Thailand. We compared Dice coefficient of these samples with respect to the ground truth data for both Bhattacharyya and GrowCut approaches. As can be seen in Figure 3, except in one case (case 14), Bhattacharyya's performance is close or in most cases better than the GrowCut method. The evolution flow based on Bhattacharyya approach may fail to capture the lesion's boundary in cases in which intensity values do not provide enough information to the evolving contours. As an example, Figure 4 shows the case 14 in which Grow-cut approach outperform the Bhattachayya. The reason is that the darker homogeneous region inside the lesion tends to maximize the pdfs and the evolution flow mistakenly detects this area as the lesion.

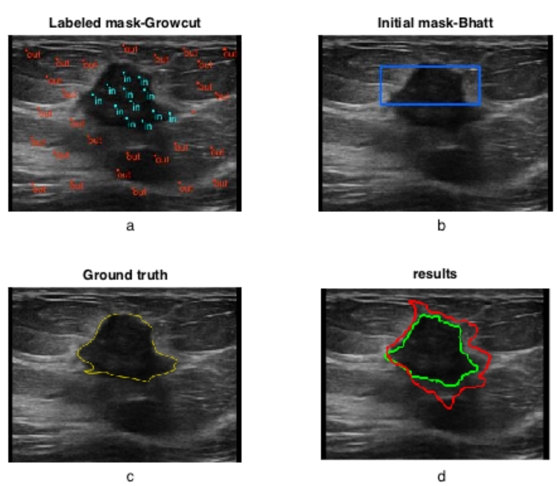


Figure 1: a) initial mask: Growcut b) initial mask: Bhattacharyya c) ground truth d) green contour: Bhattacharyya; red contour: Growcut


Figure 2: Dice Coefficient comparison, a) Grow Cut, b) Bhattacharyya

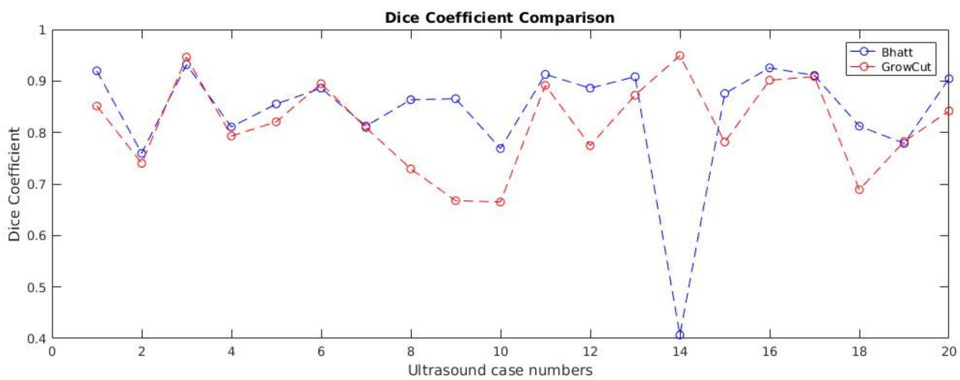


Figure 3: Dice Coefficient comparison for 20 ultrasound cases

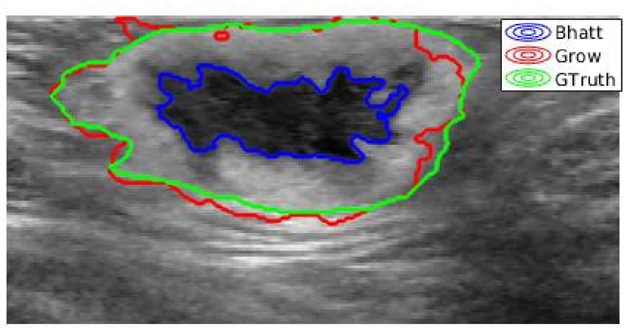


Figure 4: The case in which Grow cut outperform Bhattacharyya

## 4. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

The segmentation method we employed was derived by maximizing the Bhattacharyya distance between the pdfs defining the interior and exterior of the evolving active contour.[6] The initial curve $C$ , evolves towards the object of interest (the lesion in this case) according to Equation (14) and stops after a certain number of iterations at the boundary of the lesion. Future work will center around more automatic initialization of the initial contour as well as localizing the evolution flow in order to avoid failures in cases such as case 14. Finally, more user friendly code will need to be written together with a graphical user interface that may be employed by practicing clinicians.

## APPENDIX A. PARTIAL DERIVATIVES AS A FEATURE IMAGE

As it was discussed in section 2, vector-valued feature image $J$ can be computed by means of applying an appropriate transformation function $T$ on the input image before starting the segmentation process. One of the simplest choice for transformation function is identity function ($T(I(x)) = I(x)$) and in this case the problem is reduced to histogram based segmentation problem.[22]

Assume the intensity pattern of the object of interest is same as its background. In order to detect the object in this case, one need to use the information related to relative displacement of the patterns with respect to each other. Choosing a transformation function which uses gray-level intensity values alone such as identity function is not a reasonable choice in this case. Information regarding the relative displacement of the patterns with respect to each other can be achieved by setting the transformation function $T = \nabla$. Figure 5 shows the result of applying the evolution flow (Equation 14) on an image which the object and its background have the same intensity distributions. It can be seen that the algorithm fails to detect the object (circular central region) in case of identity transformation function $T(I(x)) = I(x)$ and succeeds when the transformation function is set to gradient operator.

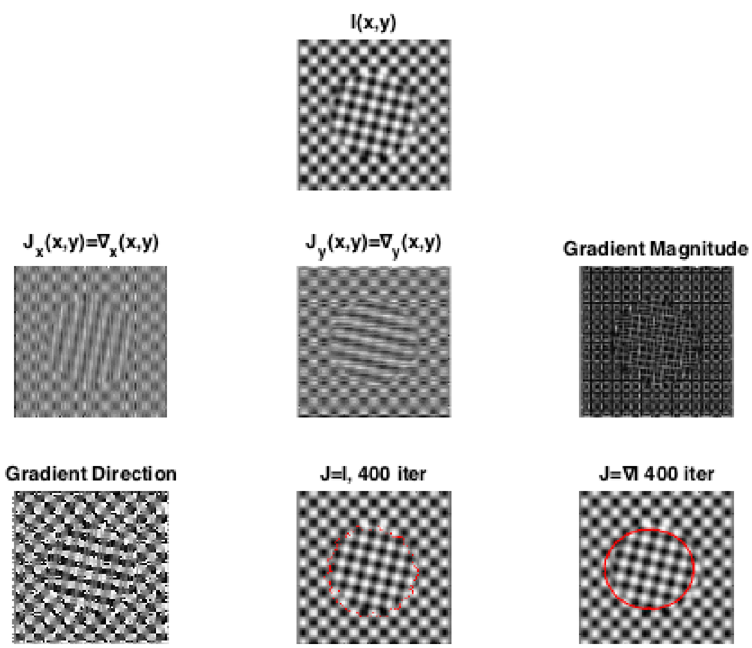


Figure 5: Gradient operation as a feature function

## Acknowledgements

This project was supported by AFOSR grants (FA9550-15-1-0045 and FA9550-17-1-0435), ARO grant (W911NF-17-1-049), grants from the National Center for Research Resources (P41-RR-013218) and the National Institute of Biomedical Imaging and Bioengineering (P41-EB-015902), NCI grant (1U24CA18092401A1), NIA grant (R01 AG053991), and a grant from the Breast Cancer Research Foundation.


## REFERENCES


[1] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bulletin Calcutta Mathematical Society*, 99–109 (1943)

[2] A. Zijdenbos, B.M. Dawant, R. Margolin and A. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE Transactions on Medical Imaging* **13(4)**, 716–724 (1994).

[3] V. Vezhnevets and V. Konouchine., "GrowCut: Interactive multi-label ND image segmentation by cellular automata, " *In Proc. of Graphicon* **1(4)**, 150–156 (2005).

[4] J. Jesneck, J. Lo, J. Baker, "Breast mass lesions: computer-aided diagnosis models with mammographic and sonographic descriptors," *Radiology* **244(2)** , 390–398 (2007).

[5] Andreea G.I., Pegza R., Lascu L., Bondari, S. Stoica Z. and Bondari A., "The role of imaging techniques in diagnosis of breast cancer," *J. Curr. Health Sci* **37(2)**, 241–248 (2011).

[6] O. Michailovich, Y. Rathi and A. Tannenbaum, "Image segmentation using active contours driven by the Bhattacharyya gradient flow," *IEEE Transactions on Image Processing* **16(11)**, 2787–2801 (2007).

[7] T. Leung and J. Malik, "Representing and recognizing the visual appear- ance of materials using three-dimensional textons," *Int. J. Comput. Vis.* **43(1)**, 29–44 (2001).

[8] F. Goudail, P. Refregier and G. Delyon, "Bhattacharyya distance as a contrast parameter for statistical processing of noisy optical images," *J. Opt. Soc. Amer. A* **21(7)**, 1231–1240 (2004).

[9] D. Freedman and T. Zhang, "Active contours for tracking distributions," *IEEE Trans. Image Proc.* **13(4)**, 518–526 (2004).

[10] I. R. Khan and F. Farbiz, "A new similarity measure and back-projection scheme for robust object tracking," *In Communications and Information Technologies (ISCIT)*, 412–417 (2010).

[11] Y. Lou, A. Irimia, P. Vela, M. Chambers, J. Van Horn, P. Vespa and A. Tannenbaum, "Multimodal deformable registration of traumatic brain injury MR volumes via the Bhattacharyya distance," *IEEE Transactions on Biomedical Engineering* **60(9)**, 2511–2520 (2013).

[12] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.* **15(1)**, 52–60 (1967).

[13] Osher S. and Sethian J. A., "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulation," *Journal of computational physics* **79**, 12–49 (1998).

[14] Osher S. and Sethian J. A., "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," *Journal of computational physics* **79**, 12–49 (1998).

[15] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11(7)**, 674–693 (1989).

[16] B. K. P. Horn and B. G. Schunck., "Determining optical flow," *Artif. Intell.* **17**, 185–203 (1981).

[17] M. Tuceryan, "Moment based texture segmentation," *Pattern Recognit. Lett.* **15**, 659–668 (1994).

[18] B. B. Chaudhuri and N. Sarkar., "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell* **17(1)**, 72–77 (1995).

[19] B. W. Silverman., "Density Estimation for Statistics and Data Analysis," *CRC Press* **26**, (1986).

[20] Cheng, L. T., Burchard, Paul, Merriman, Barry and Osher, Stanley., "Motion of curves constrained on surfaces using a level-set approach," *Journal of Computational Physics* **175(2)** , 604–644 (2002).

[21] M. C. Jones and J. S. Marron and S. J. Sheather., "A brief survey of bandwidth selection for density estimation," *J. Amer. Statist. Assoc.* **91(433)**, 401–407 (1996).

[22] C. A. Glasbey, "An analysis of histogram-based thresholding algorithms," *Graph. Models and Image Process.* **55**, 532–537 (1993).