

Model-based Detection of Acoustically Dense Objects in Ultrasound

Jyotirmoy Banerjee and Kajoli B. Krishnan
GE Global Research, Imaging Technologies Lab, JFWTC, Bangalore
 {Jyotirmoy.Banerjee, Kajoli.Krishnan}@ge.com

Abstract

Traditional methods of detection tend to underperform in the presence of the strong and variable background clutter that characterize a medical ultrasound image. In this paper, we present a novel diffusion based technique to localize acoustically dense objects in an ultrasound image. The approach is premised on the observation that the topology of noise in ultrasound images is more sensitive to diffusion than that of any such physical object ($\gg \lambda$). We show that our method when applied to the problem of fetal head detection and automatic measurement of head circumference in 59 obstetric scans compares remarkably well with manually assisted measurements. Based on fetal age estimates and their bounds specified in Standard OB Tables [6], the Gestational Age predictions from automated measurements is found to be within $\pm 2SD$ in 95% and 98% of cases when compared with manual measurements by two experts. The framework is general and can be extended to object localization in diverse applications of ultrasound imaging.

1 Introduction

Sonographers have traditionally located objects of interest by using knobs provided on Ultrasound scanners for optimizing views of acquired images based on perceptual needs. They subsequently used mouse-based caliper to make measurements of anatomical features to assist diagnosis. Operator dependence and reproducibility have continued to be the two big challenges in ultrasound imaging. However, absence of a standardized range of intensity levels such as the Hounsfield Unit as used in X-Ray CT to identify specific tissue types, artifacts induced due to anatomy and imperfect beam forming make automatic analysis of ultrasound images a very challenging problem.

Ultrasonic measurements can be 2D or 3D depending on the anatomy, probe and the method used. Objects of interest in the foreground could be a hyperechoic structure (fetal femur) or boundary (fetal head);

a hypoechoic cyst (in liver) or an isoechoic organ (kidney). Locating the object of interest in an ultrasound image is a necessary first step for computer-assisted algorithms to make automated measurements.

Resolution and uniformity across the field of view of medical ultrasound images is inherently limited by the small differences in acoustic impedance across tissue, frequency and depth dependence of ultrasound in tissue, and speckle noise due to interfering signals from multiple scatterers at the receiver. This leads to significant and variable background clutter in ultrasound images that result in the under-performance of traditional methods of automated detection.

Historically, object detection and segmentation in the domain of medical imaging has tended to focused on MRI and CT. This has been motivated by 1) the need to assist the radiologist to sift through large volumes of data, and 2) well delineated boundaries of anatomical structures in these imaging modalities. In the recent years, there has been a visible interest in taking this experience to medical ultrasound primarily for ultrasonic detection and measurement on the scanner to improve clinical workflow and reproducibility of outcome. Optimization approaches have been applied to fetal image analysis in ultrasound to perform segmentation [2, 7, 8]. Such methods can be susceptible to problems associated with local minima. Carneiro et al. [1] have used a learning based approach, involving large user annotated training data and a discriminant classifier in Probabilistic Boosting Tree.

Ultrasound waves impinging on structures (\gg wavelength- λ) with high characteristic impedance (such as bones) produce relatively high intensity specular echo signal. On the other hand, anatomical features that are small and/or have weaker impedance either produce diffuse echoes or low intensity specular echoes that are subdued by surrounding noise. As a result, application of a diffusion operator significantly alters the topology of regions in an ultrasound image characterized by diffusive and/or weak echoes. Motivated by the observation that the topology of these re-

gion show greater variability in response to multi-level diffusion, we propose the use of variance in topology as an outlier rejection strategy to facilitate object detection. Scale-space approaches have been applied extensively for image enhancement [9]. Our approach is distinct from earlier methods: it is a completely unsupervised multi-level detection framework, which utilizes multi-level diffusion along with regression technique (RANSAC [4]) to detect an object.

2 Algorithm Overview

The essential idea of this approach is quite simple: The original image is used to derive a family of images $I(x, y, \sigma)$ obtained by convolving the original image $I_o(x, y)$ with a Gaussian kernel $G(x, y, \sigma)$ of Standard deviation σ . $I_{ps}(x, y, \sigma)$ is the point set image containing set of data points representing the image $I(x, y, \sigma)$. As the image is subjected to noise, the data point consists of “inliers”, i.e., data whose distribution can be explained by some set of model parameters, and “outliers” which are data that do not fit the model. Hence given a (usually small) set of inliers across multiple diffused images, the task is to estimate the parameters of the model that optimally describes this data. We use RANSAC [4] as the regression technique. The primary contribution of our approach is that our algorithm utilizes multiple diffused version of the original data as an additional information along with a standard regression technique, to localize the object. The algorithm is essentially composed of three steps: 1) Minimal sample sets (MSSs) are randomly selected from the input dataset and the model parameters are computed using only the elements of the MSS. The cardinality of the MSS is the smallest and sufficient to determine the model parameters (e.g. if the model is a line or ellipse then the cardinality should be at least two or five, respectively). 2) Checks which elements of the entire dataset are consistent with the model instantiated with the parameters estimated in the first step (The set of such elements is called consensus set - CS). 3) Finds instances of objects within a certain class of shapes that are found consistently across the family of images at the same location by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space (that is explicitly constructed by the algorithm for computing the Hough transform), given that the local maxima contains candidates from all the diffused images. The grid size of the accumulator space is fixed based on certain threshold on the variance of parameters (VoP) which is required to put a cap on the variability of the topology. Finally, the algorithm terminates when the probability of finding a bet-

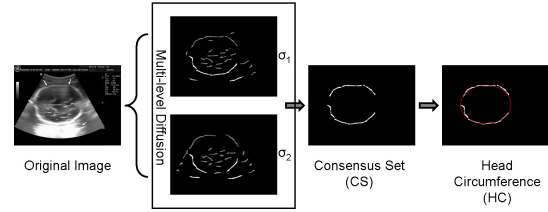


Figure 1. (left-to-right) (a) Typical Fetal Head Image (b) Pre-processed multi-level diffused images (I_{cc}) with Labeled Connected Components. The labels are shown with colour coding where (*white, grey*) correspond to (*upper, lower*) thresholds, respectively) (c) Consensus set (CS) (d) Head Circumference.

ter ranked CS amongst the candidate CS’s drops below a certain threshold.

3 An Application: Fetal Head Detection

We show the fetal head detection problem (see figure 1). Given a new image, our goal is to detect and segment the object (head) of interest and calculate the Head Circumference(HC). To generate the point set image $I_{ps}(x, y, \sigma)$, we introduce a preprocessing module. The flow of the method is as follows: (a) Feature extraction and generating point set images at multiple level of diffusion. (b) Use these images as an input to the (*Diffusion + RANSAC*) algorithm.

3.1 Feature Image & Point Set Generation

To extract the high curvature cranium features, we take the mean curvature ($I_F = \text{div}(\nabla I / \|\nabla I\|)$) of the image (I). The relevant gradient field for this image is the vector field depicting the rate of change of intensity at a point.

Bone has an acoustic impedance of $5.3 \times 10^6 \text{ kg/m}^2\text{s}$, about four times that of soft tissue [3]. If a point is on the boundary of the head (cranium bone) then it will have intensity higher compared to its neighbour, such that the vector field points inward towards that region. Therefore the divergence of the vector field in that region would have a negative value, as the region is a sink. If the region does not belong to cranium(bone), ideally the divergence should be positive and the region is called a source. However, this may not always hold true as the image is perturbed by noise.

To generate the point set image I_{ps} we follow a straight forward approach. Feature image I_F alone may capture only the structural information. As the cranium is of high intensity, we use the intensity information as well. A combination of feature image I_F and the diffused image I is used to generate a binary image. This binary image can be looked upon as a point set image

I_{ps} containing connected components where each connected component is a collection of points. To speed up the algorithm, we no more deal with the points, instead consider connected component as the basic entity for the regression algorithm. For the sake of convenience we rename the point set image I_{ps} as the connected component image I_{cc} .

Although the cranium is structurally uniform, its intensity on the image varies according to the alignment of the edges with respect to the ultrasound beam. Considerable amount of variability in the data is also introduced by the operator in setting imaging parameters such as focus and depth and scanner adjustments such as gain and dynamic range. Mindful of this scenario, we label the binary image I_{cc} with the intensity information. To capture the intensity information the diffused image I is thresholded at multi-levels, where ϕ_I^k is the value at the k th threshold level. The intensity information is embedded in two steps :

Step 1: In order to capture maximum ground, the connected component image $I_{cc}(x, y)$ is generated with the lowest k value ($k = lower$)

$$I_{cc}(x, y) = \begin{cases} 1 & \text{if } I_F(x, y) < \phi_{I_F} \text{ and } I(x, y) > \phi_I^{lower} \\ 0 & \text{else} \end{cases}$$

where ϕ_{I_F} and ϕ_I are the respective thresholds on I_F and I . The $I_F(x, y)$ is set less than ϕ_{I_F} as the region with negative value represent the cranium (see section 3.1). The cranium is the positive intensity region, thus $I(x, y)$ is set greater than ϕ_I^{lower} .

Step 2: The intention is to label the connected components in the image I_{cc} with labels associated with the intensity levels. In this method we use a two-level threshold. d_i^l is the i th connected component (Ω_i) with label l in the image I_{cc} such that:

$$l = \begin{cases} lower & \text{if } I(x, y) < \phi_I^{upper} \quad \forall (x, y) \in \Omega_i \\ upper & \text{if } I(x, y) \geq \phi_I^{upper} \quad \exists (x, y) \in \Omega_i \end{cases}$$

The multi-level labeling is possible because logically a high intensity region is always contained in a low intensity region, forming a tree structure. Further, the labeling based on intensity can be looked upon as embedding the contrast information on the connected component image I_{cc} . As the cranium is usually of high intensity, we induce the labeling information as an additional constraint on the regression algorithm for better initialization (discussed in the section 3.2).

3.2 The (Diffusion + RANSAC) Algorithm

The input image which is composed of Q connected components is indicated by $D = \{d_1^l, \dots, d_Q^l\}$ where label $l \in (lower, upper)$ and we will indicate a MSS with the letter s . Let $\varphi(\{d_1^l, \dots, d_h^l\})$ be the parameter vector estimated using the set of data $\{d_1^l, \dots, d_h^l\}$,

where $h \geq k$ and k is the cardinality of the MSS. Invariably a connected component has at least five points which is necessary and sufficient condition to draw a unique ellipse. Hence k is set to one. As the cranium has high intensity, we ensure that at least one connected component in the MSS should have the label $l = upper$.

The model \mathbf{M} is defined as: $\mathbf{M}(\varphi) \stackrel{def}{=} \{d \in R^d : f_M(d; \varphi) = 0\}$, where φ is a parameter vector and f_M is a function containing all the points that fit the model \mathbf{M} instantiated with the parameter vector φ . We define the error associated with the datum d with respect to the model $\mathbf{M}(\varphi)$ as the distance from d to $\mathbf{M}(\varphi)$: $e(d, \mathbf{M}(\varphi)) = \sqrt{LSE}/pn$, where LSE is least square distance function [5], p the scale factor is the perimeter of the ellipse and n the normalizing factor is the number of points considered for the ellipse fit. Using this error metric, fetal head characteristics and information from the clinical accepted scan-plane protocol, we define the CS as:

$$S(\varphi) \stackrel{def}{=} \{d \in D : e(d, \mathbf{M}(\varphi)) \leq \delta, p_{min} \leq p(\varphi) \leq p_{max}, \eta(\varphi) \leq \eta_{max}, \angle(\varphi) \leq \angle_{max}\}$$

where δ is a threshold on the cost of ellipse fit, which is inferred from the nature of the problem. (p_{max}, p_{min}), η_{max} and \angle_{max} are the bounds on the perimeter p , eccentricity η and orientation \angle , respectively. The perimeter and eccentricity bounds are extracted using the standard OB tables [6]. The limit on the orientation is based on scan-plane guidelines. The variance of parameter (VoP) is computed using only the elements that in parameter space φ are consistent across all the diffused images I_{σ_k} :

$$var(\hat{\varphi}/S_i) \stackrel{def}{=} \{E\{(\bar{\varphi} - \hat{\varphi}_j)^2\} : (\bar{\varphi} - \hat{\varphi}_j) < \phi_{(VoP)}, \varphi_j \in S_j \quad (\exists S_j) \in (\forall I_{\sigma_k})\} \quad (1)$$

where $\phi_{(VoP)}$ is the size of the accumulator grid in the parameter space. As mentioned in section 2, we down select all the CS based on VoP that are consistent in shape, size and location across the diffused images (see equation 1). The algorithm terminates after finding the CS with the best cardinality amongst the remaining CS's. The cardinality of the CS is defined by the number of discrete points from the fitted ellipse that lie on the CS. The points on the ellipse circumference are discretized based on constant angular span. The angular discretization of the ellipse circumference, has a normalizing effect on the cardinality, across various scales of ellipses. Once the fetal head is localized, the measurement process utilizes an ellipse fitting technique [5] to determine the Head Circumference.

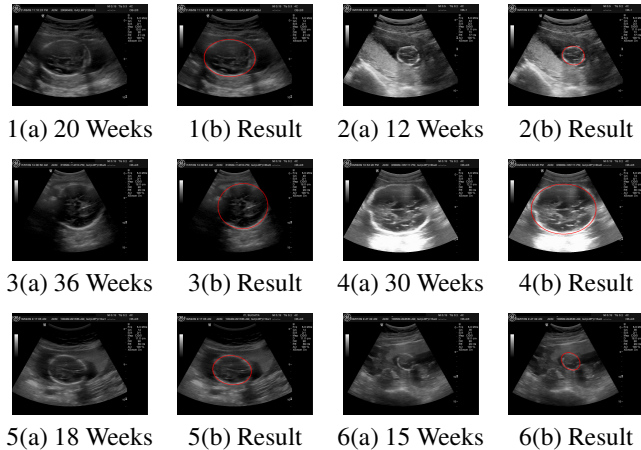


Figure 2. Results across all the three trimesters.

4 Experimental Results

We collected a total of 59 fetal head images from a Clinical Site that uses *GE LOGIQ P3* ultrasound scanner. In our numerical experiments, we generally choose the parameters as follows: $k = 2$ (two level diffusion), $\sigma_1^2 = 15$, $\sigma_2^2 = 5$ (see figure 1b). One set of manual measurements is made by the expert who performed the scan. The other set is made by a radiologist offline. The gestational age (GA) corresponding to the Head Circumference (HC) computed by the algorithm is compared with the gestational age corresponding to the manual measurements made by two experts based on [6] and given by: With Expert I : 48/59 within $\pm 1SD$, 56/59 within $\pm 2SD$, 59/59 within $\pm 3SD$; With Expert II : 36/59 within $\pm 1SD$, 58/59 within $\pm 2SD$, 59/59 within $\pm 3SD$. The range and variability of the measurements by the three systems (Expert I, II and Automated Algorithm) is shown in figure 4. GA from [6] with automated measurements is plotted along with GA's with $\pm 2SD$ bounds corresponding to manual measurements by the two experts in figure 3. Figure 4 shows the relative spread in the measurement. On a 2.6 GHz Intel® Processor with 3.5 GB of RAM, the cpu time is ≈ 1.6 seconds.

5 Conclusion

We have proposed a novel diffusion based technique to localize an acoustically dense object in an ultrasound image. This approach is used to automatically detect, segment and measure fetal head from obstetric scans.

References

- [1] G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu. Detection and measurement of fetal anatomies from ul-

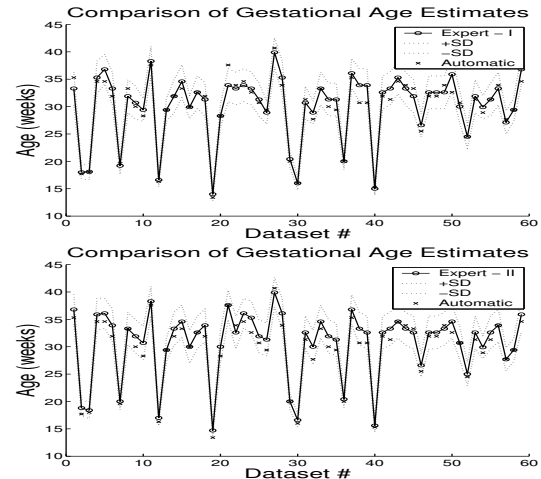


Figure 3. Experimental results with two experts.

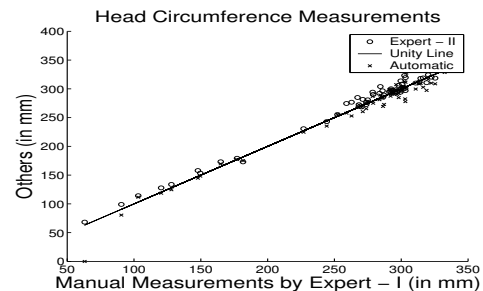


Figure 4. Variability range of HC measurements.

- trasound images using a constrained probabilistic boosting tree. *MedImg*, 27(9):1342–1355, September 2008.
- [2] V. Chalana, T. Winter II, D. Cyr, D. Haynor, and Y. Kim. Automatic fetal head measurements from sonographic images. In *Acad Radiology*, pages 628–635, 1996.
- [3] R. Farr and P. Allisy-Roberts. *Physics of Medical Imaging*. Harcourt Publishers Limited, 1998.
- [4] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 24(6):381–395, June 1981.
- [5] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(5):476–480, 1999.
- [6] F. Hadlock, R. Deter, R. Harrist, and S. Park. Estimating fetal age: Computer assisted analysis of multiple fetal growth parameters. *Radiology*, 152:497–501, 1984.
- [7] C. Hanna and A. Youssef. Automated measurements in obstetric ultrasound images. In *ICIP*, pages III: 504–507, 1997.
- [8] S. Jardim and M. Figueiredo. Segmentation of fetal ultrasound images. *Ultrasound in Medicine and Biology*, 31(2):243–250, 2005.
- [9] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. In *CVWS*, pages 16–22, 1987.