Integration of Local and Global Features for Anatomical Object Detection in Ultrasound

Bahbibi Rahmatullah[1], Aris T. Papageorghiou[2], and J. Alison Noble[1]

[1] Institute of Biomedical Engineering, Dept. of Engineering Science, University of Oxford
[2] Nuffield Dept. of Obstetrics and Gynaecology, University of Oxford, United Kingdom
{bahbibi.rahmatullah,alison.noble}@eng.ox.ac.uk,
aris.papageorghious@obs-gyn.ox.ac.uk

Abstract. The use of classifier-based object detection has found to be a promising approach in medical anatomy detection. In ultrasound images, the detection task is very challenging due to speckle, shadows and low contrast characteristic features. Typical detection algorithms that use purely intensity-based image features with an exhaustive scan of the image (sliding window approach) tend not to perform very well and incur a very high computational cost. The proposed approach in this paper achieves a significant improvement in detection rates while avoiding exhaustive scanning, thereby gaining a large increase in speed. Our approach uses the combination of local features from an intensity image and global features derived from a local phase-based image known as feature symmetry. The proposed approach has been applied to 2384 two-dimensional (2D) fetal ultrasound abdominal images for the detection of the stomach and the umbilical vein. The results presented show that it outperforms prior related work that uses only local or only global features.

Keywords: Ultrasound, Local phase, Monogenic signal, Feature symmetry, Haar features, AdaBoost, Anatomical object detection.

1 Introduction

Ultrasound imaging is considered the simplest, least expensive and most widely used imaging modality in the field of obstetrics. Standard fetal biometric measurements from 2D ultrasound have been extensively used to estimate the gestational age of the fetus, to track fetal growth pattern, to estimate fetal weight and to detect abnormalities. Typically, fetal biometry is determined from standardized ultrasound planes taken from the fetal head, abdomen and thigh. Fetal growth is then assessed from these measurements by using population-based growth charts. The acquisition of standard image planes where these measurements are taken from is crucial to allow for accurate and reproducible biometric measurements, and also to minimize inter- and intra-observer variability.

Detection of medical anatomic structure plays an important role in medical image understanding and application. In our application, the detection of the important anatomical landmarks is one of the pre-defined criteria for qualitative scoring of a

fetal biometric image [1]. For example, the correct abdominal section for taking the abdominal circumference measurement (one of the important fetal biometric measures) should demonstrate the main anatomic landmarks: the stomach (SB) and the umbilical vein (UV). Results from the detection can also provide valuable initialization information for applications such as segmentation and registration. Therefore, the methodology presented in this paper is general and could be adapted for the detection of other anatomical features in medical imaging.

The use of classifier-based object detection with bounding boxes [2] which gained its popularity for generic object detection in natural images has found to be promising in medical anatomy detection [3]. The standard approach is to train the binary classifier by discriminating the anatomic structure of interest from the background. Then, exhaustive scan using the sliding window technique is performed (for all possible translations and a sparse set of scales) to find the anatomical object in the query image. In other words, a classifier is applied to all sub-windows within an image and takes the maximum of the classification score as an indication of the presence or absence of an object. One inherent disadvantage of this approach is the significant increase in computational cost, because of the large number of candidate sub-windows.

In [4], Rahmatullah et al. proposed the use of Haar features for the detection of important anatomical landmarks in fetal abdominal images. However, the local detector was applied on the entire image, resulting in an inefficient computational time of six seconds for an image size of 1024 by 768 pixels. In order to avoid the exhaustive search, our proposed approach includes global features that are formulated to predict the likely location of the object. It is important that the features used are detectable even under changes in image scale, noise and contrast. Moreover, in the case of fetal ultrasound images, the fetus shape and anatomy varies during pregnancy and the image quality diminishes with gestational age (GA) with typically stronger artifacts appearing in the image towards the end of pregnancy. The anatomical object can also appear at any pose in the images.

Encouraging results in other ultrasound-based applications (for example [5,6]) have shown that a local phase based approach outperforms the conventional intensity-based approach for feature detection in ultrasound images . In this paper, we propose the feature symmetry (FS) measure [7] derived from the local phase-based method as a global (coarse) feature for object detection. A local (fine) detector is then applied only to the locations deemed probable by the global features. To the best of our knowledge, the current study is the first to demonstrate the efficiency and the significance of integrating the feature symmetry in a machine learning framework for the detection of anatomical features in ultrasound images.

2   Local Phase Based Feature Measurement

Local phase information can be obtained by convolving a 1D signal with a pair of band-pass quadrature filters. A common choice of quadrature filters is the log-Gabor filter, which has a Gaussian transfer function when viewed on the logarithmic frequency scale. The log-Gabor filter has a transfer function of the form:

$$G(\omega) = exp\left(-\frac{log^2(\omega/k)}{2log^2(\sigma_\omega)}\right) \tag{1}$$

where k is the centre frequency of the filter and $0 < \sigma_\omega < 1$ is related to the spread of the frequency spectrum in a logarithmic function. Log-Gabor filters allow arbitrarily large bandwidth filters to be constructed while still maintaining a zero DC component in the even-symmetric filters.

The monogenic signal was introduced to calculate local phase in N-dimensional signals [8]. The monogenic signal is generated using the Riesz transform. The spatial representations of these filters are: $h_1(x,y) = -x/(2\pi(x^2+y^2)^{\frac{3}{2}})$ and $h_2(x,y) = -y/(2\pi(x^2+y^2)^{\frac{3}{2}})$. The image $I(x,y)$ is first convolved with an even isotropic band-pass filter $b(x,y)$ that produces the even (symmetric) component of the monogenic signal: $even(x,y) = I_b(x,y) = b(x,y) * I(x,y)$. The bandpassed image $I_b(x,y)$ is then convolved with the Riesz filter to produce the two odd (anti-symmetric) components: $odd_1(x,y) = h_1(x,y) * I_b(x,y)$ and $odd_2(x,y) = h_2(x,y) * I_b(x,y)$. The monogenic signal $I_M(x,y)$ of $I(x,y)$ is often expressed as $I_M(x,y) = [I_b(x,y), h_1(x,y) * I_b(x,y), h_2(x,y) * I_b(x,y)]$.

The stomach and the umbilical vein in fetal ultrasound images, typically appear as dark blobs, with non-uniform intensity and sizes in different scans. Therefore we propose the use of the multi-scale feature symmetry (FS) measure [7] for computing the phase congruency for these two objects, which is defined as:

$$FS(x,y) = \sum_s \frac{\lfloor[|even_s(x,y)| - |odd_s(x,y)|] - T_s\rfloor}{\sqrt{even_s(x,y)^2 + odd_s(x,y)^2} + \varepsilon} \tag{2}$$

where s represents the scale of the band-pass filter, ε is a small positive constant that avoids division by zero, $\lfloor.\rfloor$ operator denotes zeroing of any negative values, and T_s is scale specific noise compression term defined similarly in [5] as $T_s = exp\left(mean\left(log\left(\sqrt{even_s(x,y)^2 + odd_s(x,y)^2}\right)\right)\right)$.

For band-pass filtering with log-Gabor filter, the following parameters produced the best empirical results: $\sigma_\omega = 0.50$ and 3 scales of filter wavelength: [250 150 50] pixels. The filter wavelength scales were selected from the coarse, medium, and fine ranges that would produce local phase images as shown in Fig. 1 (a)-(c). FS values vary from a maximum of 1, indicating a very significant feature, down to 0 indicating no significance. Based on the observation of the validation set, we set the significance value of 0.35. Connected component labeling was then applied to the FS image and the local discriminative detector was applied around the center location of the labeled component. The same FS image was used for both the stomach and the umbilical vein. In our validation set, the size of the stomach component in the labeled image usually varied from 800 to 2500 pixels and the umbilical vein from 300 to 1000 pixels. There might be overlap of potential components for the stomach and the umbilical vein but the task to distinguish them is passed to the local detector trained with the local features. Fig. 1 shows the prediction of the candidate locations of objects based on the global feature symmetry measure.
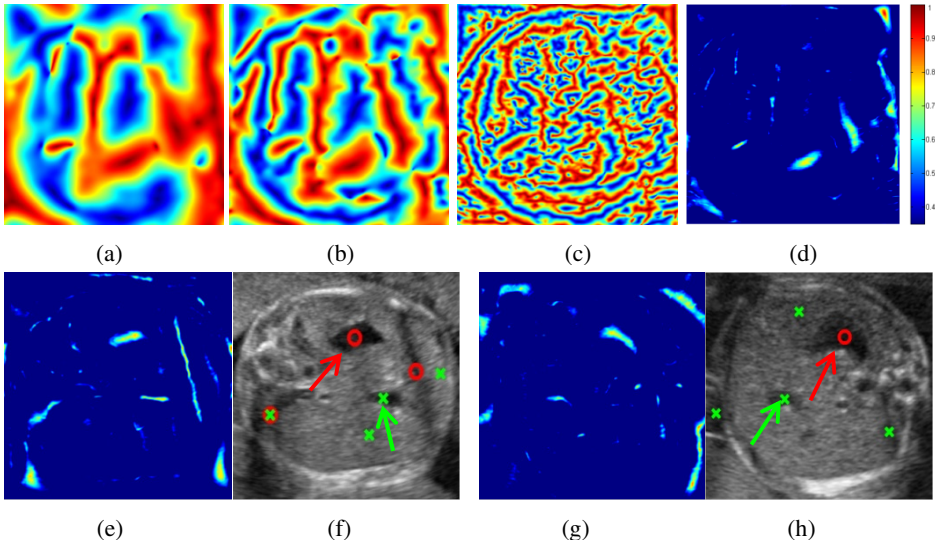
Fig. 1. The top row shows the local phase images produced using (a)-(c) coarse to fine filter scales and (d) the feature symmetry map generated using these scales. The bottom row shows the feature symmetry maps (e) and (g) with significant features (>0.35) that were used to produce the candidate locations for the stomach (*red circles*) and the umbilical vein (*green crosses*) shown superimposed on the original intensity image in (f) and (h). The *red* and *green arrows* denote the correct positions for the stomach and the umbilical vein, respectively.

3   Local Features and Training

The standard approach to local object detection is to classify each image sub-window as foreground (containing the object) or background. In this case, there are two main decisions to be made: what kind of local features to extract from each sub-window, and what kind of classifier to apply to this feature vector.

We propose to use features derived from Haar wavelets for representing the information in sub-window (local) image region. Haar features had been proven to effectively encode the domain knowledge for different types of visual patterns including objects in ultrasound images [3]. Extraction is achieved with high speed of computation due to the use of the integral image [2]. The features are trained as a local classifier using AdaBoost [9], an intuitive and proven effective method in object detection problem. It forms a strong hypothesis through linear combination of weak classifiers that are derived from a pool of extracted features. The training process involves modification of the weight distributions based on the previous classification error in order to focus on the more difficult training samples, thus driving down the classification error. Advantages of AdaBoost algorithm are that it has no parameters to tune other than the number of iterations and the most representative features are automatically selected during the training process, requiring no additional experiments.

Our training data for the classifier is created as follows. For positive training samples, we cropped image regions that contain the anatomical object. The negative samples were extracted randomly from the background and also from images that does not contain SB or UV. The cropped sub-windows were normalized to 100x100. From the training dataset, we extracted the following separate feature sets:

    a. *Local features*: Haar features extracted from intensity image
    b. *Global features*: Unary features acquired from the feature symmetry (FS) image.

300 rounds of boosting were performed on each feature set. This is based on the performance on the validation set. The resulting first five features chosen for each object are shown in Fig. 2.
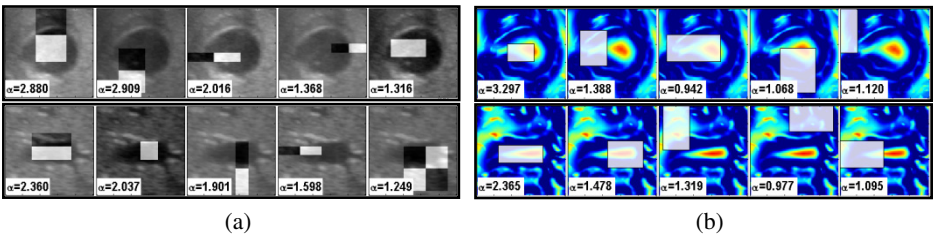

        (a)   (b)

Fig. 2. The first five selected features by AdaBoost for the stomach (*top row*) and the umbilical vein (*bottom row*) detection are shown superimposed on example images from the training set. The features are calculated by summing the value in white regions and subtracted with the totals in grey regions (if any). (a) Local features from intensity image and (b) global features from feature symmetry (FS) image are shown with each feature classifier weight (α).

4   Experimental Setup and Results

4.1   Data Acquisition

Fetal abdominal images for this work were randomly selected from a clinical study database where data has been obtained from over 4000 healthy pregnant women at low risk of impaired fetal growth who are scanned up to six times from $<14_{+0}$ weeks to term. All women were screened at study entry with particular focus on excluding known risk factors for IUGR (e.g. smoking, chronic illness) and over-growth (e.g diabetes). Multiple pregnancy or major fetal abnormality were excluded from the study. All ultrasound examinations were performed using a Philips HD9 ultrasound machine with a 2-5MHz 2D probe by ultrasonographers trained to follow standardized data acquisition procedures. All images were saved in a DICOM format with a size of 1024 x 768 pixels. Details of datasets is shown in Table 1 where images were labeled and divided (no overlaps) after consultation with trained sonographers.

Table 1. Details of the number of positive (+) and negative (-) images in the training, validation and testing datasets

	Train+	Train-	Valid+	Valid-	Test+	Test-
SB	633	2073	50	50	2283	101
UV	224	851	50	50	2284	100

4.2   Results and Performance Analysis

The experiment was implemented in MATLAB running on a Pentium Xeon® 3.4 GHz machine with 3GB of memory. We experimented with three different detection methods: "Local", "Global" and "Hybrid". In the "Local" and the "Global" methods, we exhaustively scanned the image at multiple scales using the sliding window method. Features from each sub-window were extracted and classified according to its trained model (local Haar features or global FS). For "Hybrid", the detector trained with the local features was applied at the probable locations and scales predicted by the global features (FS). We empirically set that a ground-truth object is considered detected if 75% of its area is covered by the output boxes.

We compared the performance of the three methods using ROC curves, as shown in Fig. 3 and summarize the performance metric; area under the curve (AUC), accuracy, sensitivity, specificity and execution time, shown in Table 2. From the results, we found that the method using global and local features is computationally efficient and eliminates many false positives caused by using local features alone, and these are illustrated qualitatively in Fig. 4. The umbilical vein detection is a harder problem hence the lower AUC value and accuracy compared to the stomach detection. This is due to the presence of other similar looking blood vessels in the abdomen causing false positive in the detection. However the detection accuracy is increased by 9.75% with the integration of global feature with the local features in the "Hybrid" method.

Table 2. The performance of three different methods in the detection of the stomach and the umbilical vein in fetal abdominal images

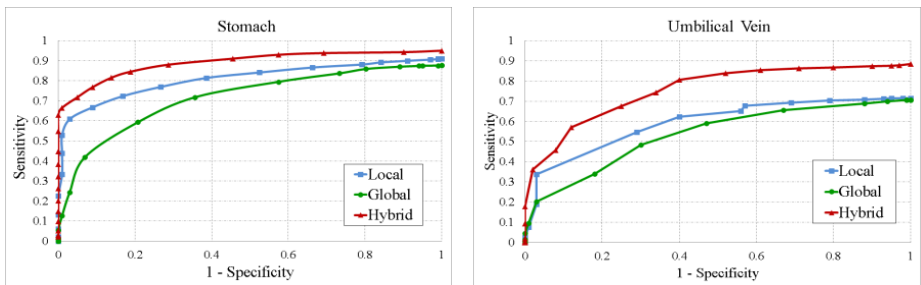|	AUC		Accuracy (%)		Sensitivity (%)		Specificity (%)		Mean Execution
	SB	UV	SB	UV	SB	UV	SB	UV	Time (secs)
Local	0.80	0.57	78.94	62.80	60.84	54.59	96.04	71.00	10.27
Global	0.71	0.53	69.28	57.99	59.35	33.98	79.21	82.00	10.65
Hybrid	0.88	0.75	82.75	72.55	66.49	57.09	99.01	88.00	0.94

Fig. 3. ROC plot for the detection of the stomach (*left*) and the umbilical vein (*right*). We see that combining the global and local features improves detection performance.

| (a) FS map | (b) Stomach detection | (c) FS map | (d) UV detection |
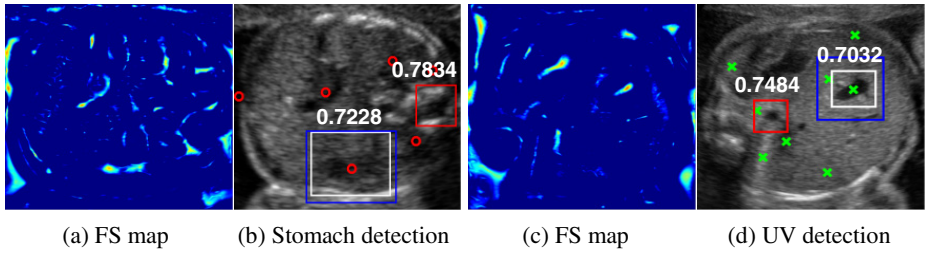
I) False positive results by "*Local*" method (*red boxes*) corrected with true positives by "*Hybrid*" method (*blue boxes*) along with the detection scores.

| (a) FS map | (b) Stomach detection | (c) FS map | (d) UV detection |
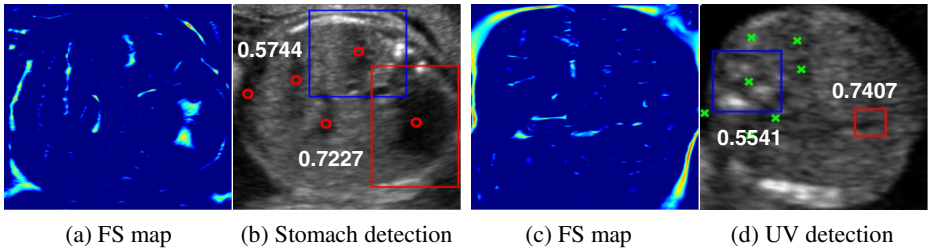
II) False positive results (high detection scores) by "*Local*" method (*red boxes*) corrected with true negatives (low detection scores) by "*Hybrid*" method (*blue boxes*).

| (a) FS map | (b) Stomach detection | (c) FS map | (d) UV detection |
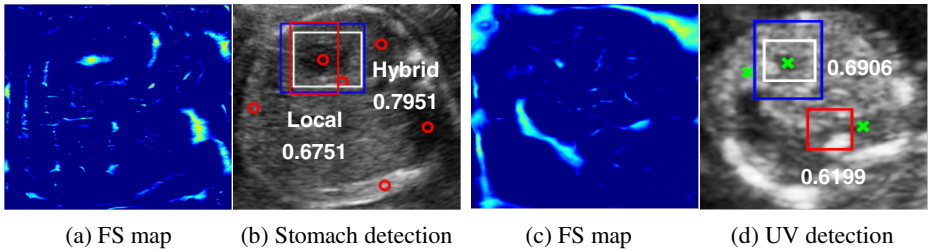
III) False negative results (low detection scores) by "*Local*" method (*red boxes*) corrected with true positives (high detection scores) by "*Hybrid*" method (*blue boxes*).

| (a) FS map | (b) Stomach detection | (c) FS map | (d) UV detection |
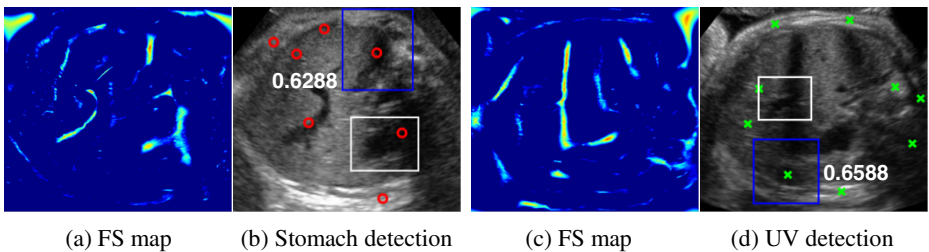
IV) Misdetections which occurred mostly in late GA images. The correct location of the objects could not be identified through the FS map due to major shadowing effect over the objects.

Fig. 4. Examples of the results achieved using "Hybrid" method. White box represents the ground truth. (a) and (c) show the FS maps that input the candidate locations for stomach (*red circles*) and UV (*green crosses*), respectively.

5 Conclusion

This paper presents a novel feature set for the detection of anatomical objects in fetal ultrasound image. We integrated the unary features extracted from local phase image global feature within a machine learning framework that trains a local classifier using local Haar features. This provides a computationally cheap step before invoking a local object detector to be applied in plausible locations and scales. The proposed method exhibits good generalization capability when tested on 2384 images with an accuracy of 82.75% and 72.55% for the detection of the stomach and the umbilical vein, respectively. In average, it runs 9 times faster than the typical local object detector with the sliding window approach. Our future works will focus on the extension of the proposed method for the detection of objects in other fetal biometry scan areas and in 3D ultrasound volumes.

References

1. Salomon, L.J., et al.: Feasibility and reproducibility of an image-scoring method for quality control of fetal biometry in the second trimester. Ultrasound in O&G 27, 34–40 (2006)
2. Viola, P., Jones, M.J.: Robust Real-Time Face Detection. International Journal of Computer Vision 57, 137–154 (2004)
3. Karavides, T., Leung, K.Y.E., Paclik, P., Hendriks, E.A., Bosch, J.G.: Database guided detection of anatomical landmark points in 3D images of the heart. In: 2010 IEEE International Symposium on Biomedical Imaging, Rotterdam, pp. 1089–1092 (2010)
4. Rahmatullah, B., Sarris, I., Papageorghiou, A., Noble, J.A.: Quality control of fetal ultrasound images: Detection of abdomen anatomical landmarks using AdaBoost. In: 2011 IEEE International Symposium on Biomedical Imaging, Chicago, IL, pp. 6–9 (2011)
5. Rajpoot, K., Vicente, V.V., Noble, J.A.: Local-phase based 3D boundary detection using monogenic signal and its application to real-time 3-D echocardiography images. In: 2009 IEEE International Symposium on Biomedical Imaging, Boston, MA, pp. 783–786 (2009)
6. Grau, V., Noble, J.A.: Adaptive Multiscale Ultrasound Compounding Using Phase Information. In: Duncan, J.S., Gerig, G. (eds.) MICCAI 2005. LNCS, vol. 3749, pp. 589–596. Springer, Heidelberg (2005)
7. Kovesi, P.: Symmetry and Asymmetry From Local Phase. In: Tenth Australian Join Conference on Artificial Intelligence, Brisbane, pp. 185–190 (1997)
8. Felsberg, M., Sommer, G.: The monogenic signal. IEEE Transactions on Signal Processing 49, 3136–3144 (2001)
9. Freund, Y., Schapire, R.E.: A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer & System Sciences 55, 119–139 (1997)