

# Plane Localization in 3-D Fetal Neurosonography for Longitudinal Analysis of the Developing Brain

Mohammad Yaqub, Sylvia Rueda, Anil Kopuri, Pedro Melo, A. T. Papageorghiou, Peter B. Sullivan, Kenneth McCormick, and J. Alison Noble

**Abstract**—The parasagittal (PS) plane is a 2-D diagnostic plane used routinely in cranial ultrasonography of the neonatal brain. This paper develops a novel approach to find the PS plane in a 3-D fetal ultrasound scan to allow image-based biomarkers to be tracked from prebirth through the first weeks of postbirth life. We propose an accurate plane-finding solution based on regression forests (RF). The method initially localizes the fetal brain and its midline automatically. The midline on several axial slices is used to detect the midsagittal plane, which is used as a constraint in the proposed RF framework to detect the PS plane. The proposed learning algorithm guides the RF learning method in a novel way by: 1) using informative voxels and voxel informative strength as a weighting within the training stage objective function, and 2) introducing regularization of the RF by proposing a geometrical feature within the training stage. Results on clinical data indicate that the new automated method is more reproducible than manual plane finding obtained by two clinicians.

**Index Terms**—Fetal brain, monogenic signal, plane localization, regression forests, ultrasound.

## I. INTRODUCTION

LONGITUDINAL analysis of the developing brain is an emerging area linking medical image computing and neurology. In the future, this may have impact on clinical management of premature infants and neonates with neurological conditions. Clinicians are exploring tools to track the changes in image-based biomarkers from the womb to the postnatal period [1], [2]. However, promoting and monitoring neurodevelopment remains a significant challenge in the care of premature infants. Recent studies on cohorts of ex-premature

infants have shown improved survival but long term neurodevelopmental outcomes have remained unchanged [3]. Volpe postulated that brain injury in premature infants is an amalgam of both destructive and developmental disturbances [4]. “Destructive” pathologies are currently diagnosed and monitored using cranial ultrasonography by identifying conditions such as intraventricular hemorrhages, infarctions, and hydrocephalus. However, developmental disturbances are more difficult to diagnose using cranial sonography. Current practice of assessment of brain ultrasound images for structural development is at best, subjective.

In a recent study [1], prematurity was shown to have a significant effect on thalamocortical development, specifically, lower gestation being associated with thalamic volume reduction. This volume depletion has been shown to be associated with lower cortical volume, decreased hippocampal volume and reduced fractional anisotropy of the corticospinal tracts [2]. This is therefore our clinical motivation for measuring thalamic size serially, which could potentially help identification of disordered brain development in premature infants. Serial MR imaging in premature infants is clinically impractical and expensive. Cranial ultrasonography is a more practical alternative to finding thalamic size with the benefit of being applicable cot-side. Furthermore, the thalamic shadow can be easily identified in the parasagittal plane on ultrasound images. Motivating this paper, a recent study [1] published preliminary findings suggesting a correlation between neonatal gestational age and the thalamic area measured manually on standard 2-D neonatal cranial ultrasound images acquired at the parasagittal plane. Thalamic area measurements were compared with equivalent fetal markers of a normal population as a reference, suggesting that this might provide a clinically useful way to monitor neuro-development of a neonate. Therefore, to derive normative data, PS planes were manually identified from 3-D ultrasound fetal brain volumes of healthy fetuses [5]. Associations with gestation and fetal thalamic cross-sectional area measurements were however weaker. To strengthen this association, automation of PS plane detection was envisaged. Furthermore, in fetal neurosonography, manual assessment of orthogonal planes is important to ensure a proper diagnosis of brain structures.

Recent advances in ultrasound techniques allow accurate measurement of anatomical structures within the fetal and preterm infant brain raising the prospect of identifying and tracking regions of interest correlating to maturational change at the cot-side. While standardized manual image acquisition has been adopted in neonatal cranial ultrasound practice [6], the ability to examine changes over time and correlate them with a reference

Manuscript received January 15, 2015; revised April 15, 2015; accepted May 11, 2015. Date of publication; date of current version. The work of M. Yaqub, S. Rueda, A. Kopuri, P. Melo, and J. A. Noble was supported by the Wellcome/EPSCRC Centre of Excellence in Medical Engineering—Personalised Healthcare, under Grant WT 088877/Z/09/Z.

M. Yaqub, S. Rueda, P. Melo, and J. A. Noble are with the Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford OX3 7DQ, U.K. (e-mail: mohammad.yaqub@eng.ox.ac.uk; sylvia.rueda@eng.ox.ac.uk; pedro.melo@eng.ox.ac.uk; alison.noble@eng.ox.ac.uk).

A. Kopuri and K. McCormick are with the Department of Paediatrics, University of Oxford, Oxford OX3 9DU, U.K., and also with the Neonatal Unit, John Radcliffe Hospital, Oxford OX3 9DU, U.K. (e-mail: anil.kopuri@wolfson.ox.ac.uk; Kenny.McCormick@ouh.nhs.uk).

A. T. Papageorghiou is with the Nuffield Department of Obstetrics and Gynaecology, University of Oxford, Oxford OX3 9DU, U.K. (e-mail: aris.papageorghiou@obs-gyn.ox.ac.uk).

P. B. Sullivan is with the Department of Paediatrics, University of Oxford, Oxford OX3 9DU, U.K. (e-mail: peter.sullivan@paediatrics.ox.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JBHI.2015.2435651

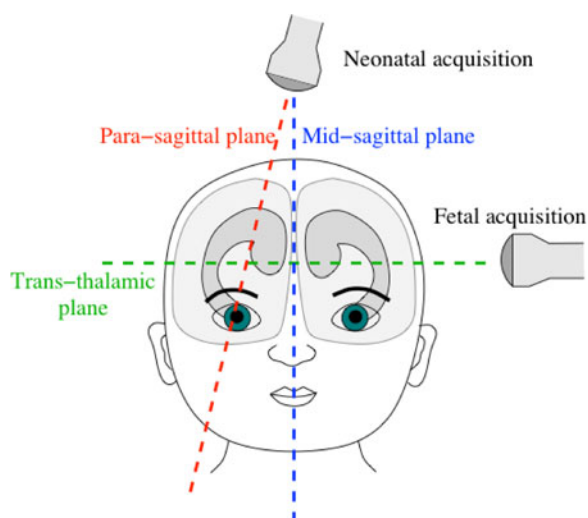


Fig. 1. Standard planes for fetal and neonatal neurosonography. Schematic of the midsagittal plane (blue), parasagittal plane (red), the transthalamic plane (green), and the typical probe orientation for the acquisition of the correct anatomical planes on fetuses and neonates.

healthy fetus at the same gestational age is crucially dependent on accurate and reproducible plane localization. Automation of this process would enable reproducible characterization of structural changes within the developing brain and may allow for assessing the effects of therapeutic intervention on brain growth and maturation.

Ultrasound-based longitudinal analysis across the perinatal period to allow correlation between clinical scans taken before and after birth remains a challenge. Antenatal scans (for the basic examination at midpregnancy) are acquired in utero with either a 2-D or a 3-D transducer positioned to obtain axial slices of the fetal brain. However, advanced fetal neurosonography examines the fetal brain in three planes (axial, sagittal, and coronal) to allow better understanding of brain development [7], [8]. Direct visualization of sagittal planes on fetuses is possible. However, it does not only require a lot of experience but also some luck to have the fetus in a perfect position to be able to acquire acceptable sagittal images. Therefore, finding a sagittal plane in a 3-D fetal image that is acquired on the axial view can provide a solution to this issue. Neonatal scans are taken *ex-utero*, with a different transducer normally limited to 2-D scans acquired through the anterior fontanelle (AF) to provide angled coronal and sagittal planes (cf., Fig. 1). The question we seek to answer in this paper is how fetal and neonatal scans are correlated. The approach outlined in this paper automatically finds, in a 3-D fetal scan, the equivalent (and nonconventional) 2-D plane, called the parasagittal plane (PS), used in neonatal neurosonography, via a machine-learning approach. Detecting such a plane reliably would allow neurological information to be obtained in theory from around 14 weeks gestational age up to at least eight weeks post birth or potentially six months depending on when the AF closes.

In fetal neurosonography, clinicians usually diagnose not only from the axial view of the brain but also the sagittal and coronal views. This assessment is done manually and more importantly,

it typically does not assess oblique planes of the brain. In an interesting work by Rizzo *et al.* [9], the authors investigated the visualization of standard sagittal and coronal planes in 3-D ultrasound of the fetal brain. They studied the agreement rate of two manual reviewers in finding such planes. Therefore, to reduce subjectivity of finding planes manually and to allow oblique plane localization, it is important to develop automatic methods that can potentially detect planes in 3-D ultrasound.

There have been a small number of publications addressing automatic plane or view localization for clinical workflow improvement. On the other hand, [10] proposes a supervised learning method based on marginal space learning to detect several planes in 3-D echocardiography. A boosting framework [11] and regression forest (RF) [12] have been proposed to locate two, three, and four chamber views in 3-D echocardiography. A machine learning method based on the integrated detection network has been developed to detect several fetal brain planes in 3-D ultrasound volumes and estimate the standard clinical 2-D fetal biometric parameters [13]. In our previous work [14], we developed a semiautomatic method to localize the PS plane on 3-D ultrasound scans of the fetal brain. The method builds on a RF framework to estimate the PS plane parameters relying on only “informative” voxels during training and testing. One issue in [14] was the high intraobserver variability on manual localization of the PS plane. To address this issue, we have defined in this paper an objective protocol to find the PS plane.

Notice that, to-date, machine learning methods have been the most popular approach for plane finding in medical imaging, especially ultrasound. Regression forests [12], [14]–[18] has gained a lot of interest in medical imaging recently in applications such as object detection and view localization. RF consists of several trees trained on a set of training examples and possible features. Each split node in a tree optimizes an objective function to create a weak learner, which splits the training examples into two branches. Leaves contain the list of training examples, which traverse all the way from the root. Training examples at a specific leaf node are used to create a distribution from their continuous variable(s). The distribution at a leaf node on each tree is used during testing to vote for the continuous label for an unseen example. More details about RF can be found in [16].

In this paper, we propose a machine learning technique based on RF to provide accurate parasagittal plane localization. The technique constrains the RF method by: 1) forcing the model to use informative voxels and voxel informative strength as a weight when optimizing the objective function used in the training stage, and 2) introducing regularization of the RF by proposing a geometrical feature within the training stage. This new feature is based on the distance between each voxel and the midsagittal (MS) plane. We propose an automatic method to first localize the MS plane, which is then fed to the RF as a constraint feature. We demonstrate that our approach leads to plane detection accuracies, which are as good as manual plane finding and which show promise for their intended use of standardizing the parasagittal plane acquisition.

This study is an extension of our previous semiautomatic approach [14] with a number of important improvements outlined in the following. First, from a technical perspective, the

new method is fully automatic whereas [14] is semiautomatic. Second, the evaluation is more thorough: We have doubled our dataset size, increased the range of fetal gestation, and evaluation is performed by two manual observers to investigate intra and interobserver variability. The reproducibility of manually selecting the parasagittal plane has never been studied before, as these planes are normally acquired as 2-D images in clinical practice. Therefore, the quantitative evaluation also includes the manual segmentation of the thalamus on a subset of the planes to investigate the sensitivity of the plane finder for thalamic area estimation.

The rest of the paper is organized as follows. In Section II, we present a framework to detect the midsagittal plane, used as a constraint to automatically find the parasagittal plane. Section III introduces the constrained RF for parasagittal plane finding. Section IV outlines the experimental setup used within this framework and Section V reports the quantitative evaluation. Discussion and conclusions are given in Sections VI and VII, respectively.

## II. MIDSAGITTAL PLANE FINDING

One new contribution of the paper is to fully automate the PS plane localization by first localizing the MS plane automatically and using it as a constraint in the PS plane finding learning algorithm. The MS plane is clinically important to assess as it contains several structures which are typically assessed in neurosonographic examination (e.g., corpus callosum and cavum septum pellucidum). In addition, localizing the MS plane is useful itself as it facilitates the study of midline diseases, such as agenesis of corpus callosum or posterior fossa anomalies.

To find the MS plane, we first identify the skull in several 2-D slices of the volume to locate the brain region only (skull stripping). We then localize the midline on these slices since it is the separator between left and right hemispheres. Finally, the MS plane is then constructed as the best plane crossing the correctly detected midlines.

### A. Acquisition Protocol of 3-D Fetal Ultrasound Images

According to the acquisition protocol used within this study, all fetal brain volumes were obtained approximately perpendicular to the left or right side of the fetal head. The acquisition aims at locating and measuring the biparietal diameter, found on the transthalamic plane in the middle of the volume. Fig. 1 shows a schematic of the acquisition protocol used, as well as the location of the midsagittal (MS), parasagittal (PS), and transthalamic planes.

### B. Preprocessing

The process of skull localization starts by identifying several axial slices in the fetal head. These slices are those forming the middle region of the volume (above and below the middle slice, i.e., the transthalamic plane). We then process each slice independently to find the skull as described in Section II-C. We then check if the detected 2-D skull contours on the selected slices are consistent (i.e., long enough and with similar angles).

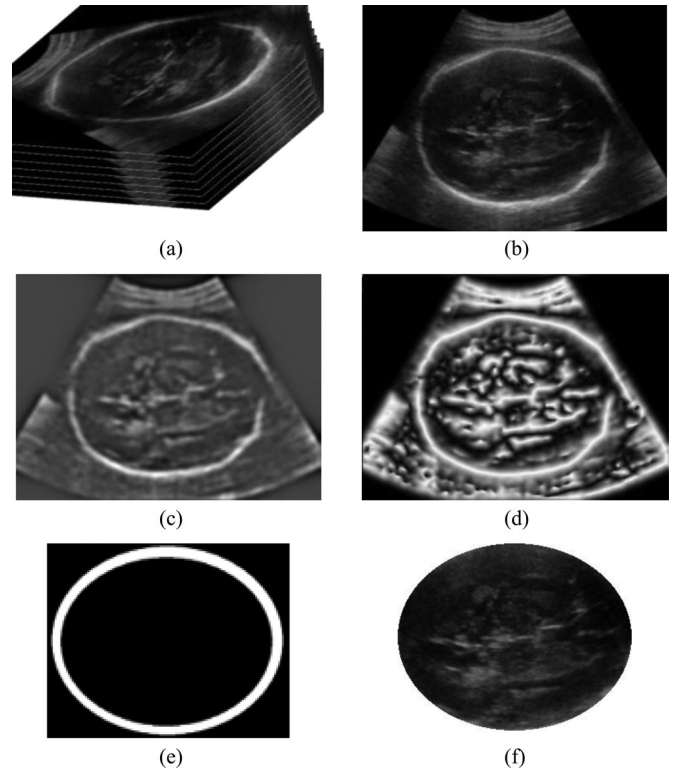


Fig. 2. Preprocessing, skull detection and brain region extraction (skull stripping). (a) List of 2-D axial slices forming a fetal 3-D volume. (b) Generic axial slice of a fetal brain at 23 weeks. (c) Output of the amplitude on the axial slice in Fig. 2(b). (d) Output of the local phase for Fig. 2(b). (e) Skull template example. (f) Extracted brain region using the template in Fig. 2(e) on Fig. 2(b).

Fig. 2(a) and (b) shows a stack of slices and a representative slice, respectively.

Prior to extracting the skull (and hence the midline), we derive the 3-D monogenic signal [19] representation for each volume. Using the monogenic signal is a powerful tool to deal with the appearance inhomogeneities inherent to ultrasound data, which make relying on intensities a hard problem [20]. The monogenic signal of an image  $I(x, y, z)$  is defined as

$$I_M(x, y, z) = \left( I_b(x, y, z), h_1(x, y, z) \otimes I_b(x, y, z), h_2(x, y, z) \otimes I_b(x, y, z), h_3(x, y, z) \otimes I_b(x, y, z) \right), \quad (1)$$

where  $I_b(x, y, z) = b(x, y, z) \otimes I(x, y, z)$  is the band-passed-filtered image by the Gaussian derivative filter  $b(x, y, z)$  [21] and  $\otimes$  denotes the convolution operation. From the monogenic signal, local energy (or amplitude)  $A(x, y, z)$  and local phase  $\varphi(x, y, z)$  can be derived as

$$A = \sqrt{I_b^2 + (h_1 \otimes I_b)^2 + (h_2 \otimes I_b)^2 + (h_3 \otimes I_b)^2} \quad (2)$$



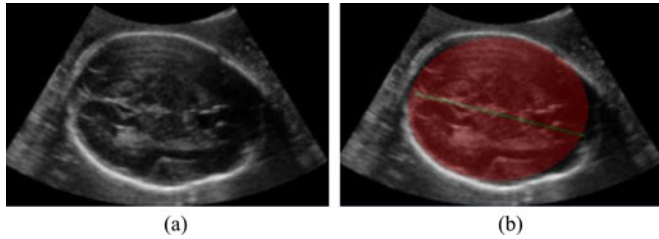


Fig. 3. Brain (red) and midline (green) localization. (a) 2-D fetal brain axial slice of a fetus at 27 weeks. (b) Detection of the brain region and the midline for Fig. 3(a).

and

$$\varphi = \arctan \left( I_b / \sqrt{(h_1 \otimes I_b)^2 + (h_2 \otimes I_b)^2 + (h_3 \otimes I_b)^2} \right). \quad (3)$$

The local energy provides a filtered version of the original image, resulting in a visible reduction of the speckle pattern and a clearer edge definition (i.e., skull). The local phase represents structural image information and one of its key properties is that it is invariant to contrast. Hence, local phase determines the type of a feature, whereas local energy determines that there is a feature at one particular pixel. A feature is defined for points of local phase congruency [22]. Fig. 2(c) shows the output of the amplitude on one 2-D slice and Fig. 2(d) shows the local phase corresponding to the same image. The local phase is used to enhance the visibility of the midline inside the brain as a long approximately straight line.

### C. Skull and Brain Localization

To find the skull on a 2-D axial brain image, several skull templates were constructed at different sizes to cover a range of fetal gestational ages. Fig. 2(e) illustrates one of the templates used within the framework. We then compute the normalized cross-correlation (NCC) between the output of the amplitude [see Fig. 2(c)] of the 2-D image and each skull template. The template which has the maximum NCC response is used to localize the brain region by masking the 2-D original ultrasound image, amplitude image and local phase image by the best-matched template of the brain. An extracted brain region is shown in Fig. 2(f) after skull stripping.

Note that the masking does not have to be exact, since we want to retain only the midregion of each slice containing the midline. In further processing, voxels outside the brain regions are neglected and the ones inside the skull are only used for MS and PS plane localization.

### D. Midline Localization

Having found the brain region, the next task is to locate the midline. The midline or falx is a structure that appears in the middle of the brain separating the left and right hemispheres. It is typically seen as a bright line in the middle of each brain ultrasound axial slice as in Fig. 3. To find the MS plane, one needs to locate the midline on several slices and then find the plane matching the midline on these slices.

We used the Hough transform to locate the midline on each axial slice of an enhanced amplitude image, called  $A_E$ , and defined as  $A_E = A \times \varphi$ , where  $A$  and  $\varphi$  are defined as in (3) and (4), respectively. We enhanced the amplitude image with local phase information because we found that due to the low contrast of ultrasound, speckle, shadows and attenuation of signal, the midline does not always have a strong signal but is locally ridge-like. This corresponds to local phase values of  $\pi/2$  and  $3\pi/2$ , for ridges going up or down, respectively. Multiplying the amplitude image by the local phase image enhances the structure of interest and the visualization of the midline is clearer.

The output of this step is the longest line found on every 2-D slice within the brain volume as illustrated in Fig. 3. However, the longest line on some slices does not necessarily represent the midline since some slices contain several structures that appear bright within the region of interest. This might lead to finding incorrect lines on some slices. To handle this, we select lines as follows. We first keep the longest 12 lines from all slices (this number was experimentally chosen). We then find the median angle amongst these lines. The lines whose angles have a small difference with respect to the median angle are retained and the others rejected. The output of this step is a line on a few brain slices corresponding to the brain midlines. An example of the detected midline is shown in Fig. 3(b) as a green line.

### E. Midsagittal Plane Detection

Having found one line on each of the axial slices in a subset of the volume, we compute all points on every line as  $p\_line(x, y, z)$ . We then fit a plane to pass through all the line points on the retained axial slices. We use least-squares-fitting to compute the plane normal to these points. The output of this step is an MS plane normal and a location on the plane. This is used later within the RF technique to compute distance features from voxels to the MS plane in each volume.

## III. CONSTRAINED RF FOR PARASAGITTAL PLANE FINDING

### A. Problem Description

Given a 3-D fetal ultrasound scan, the midsagittal plane is detected as described in Section II. The goal now is to use the fetal brain volume and the midsagittal plane to locate the parasagittal plane. We propose a machine learning method to solve this problem. The method is based on the RF technique. In the training stage, a set of 3-D training images with their corresponding manual planes are used to build the forest. In our formulation, we can generate a training example for each voxel. A training example for voxel  $v_i$  has the following form:

$$v_i \rightarrow \{f_1, f_2, f_3, \dots, f_n\}, d = (d_x, d_y, d_z).$$

Here,  $f_j$  is a feature computed for the voxel  $v_i$ .  $n$  is the total number of possible features.  $d = (d_x, d_y, d_z)$  is a distance vector which represents the continuous class label for the training example.  $d$  is the perpendicular distance between  $v_i$  and the manual PS plane. The PS plane is represented as three parameters for the normal vector and three parameters to represent a point on the plane. We compute  $d$  as in [12].

The list of training examples from several ultrasound volumes alongside their continuous class labels are used to train an RF model. The forest is then used to test unseen ultrasound volumes.

### B. Sampling Informative Voxels

During the RF training on image volumes, millions of features need to be computed on typically millions of examples (voxels in our case). RF is responsible for finding the best set of features to build each tree. Although many of the training examples are noninformative, RF treats all training examples equally during training. During a node split in the forest, an objective function is optimized over all training examples which reached that node. If several of these examples are noninformative, the objective function could fail to find a good split. As a result, this could create weaker trees and hence a weaker forest which means lower regression accuracy. Additionally during testing of an unseen volume, relying on all testing examples (voxels) to vote for plane parameters could potentially yield erroneous results because many voxels are noninformative.

We only sample voxels from inside the brain since we extract the brain region (see Section II). This allows us to exclude all voxels from outside the brain that do not necessarily contain meaningful information. In fact, due to the existence of maternal tissues and the appearance heterogeneity outside the fetal brain region, voxels from outside the brain are misleading to the learning algorithm. Therefore, we restrict the learning algorithm to utilize brain voxels only to locate the PS plane.

We also propose a technique which finds the informative voxels within the fetal brain and uses their strength to weigh their influence during tree training and testing. We present a solution to overcome a common problem observed in machine learning in ultrasound of otherwise generating a sub-optimal RF because of the existence of many noninformative voxels.

An informative voxel is a voxel that is discriminative within its neighborhood. Here, we use feature asymmetry (FA) to highlight informative voxels motivated by its successful application to 3-D echo feature enhancement and detection, such as [23], [24]. FA is computed as

$$FA = \frac{1}{N} \sum_s \frac{||\text{odd}_s| - |\text{even}_s| - T_s|}{\sqrt{\text{odd}_s^2 + \text{even}_s^2 + \varepsilon}} \quad (4)$$

where  $\text{even} = I_b$ ,  $\text{odd} = (h_1 \otimes I_b, h_2 \otimes I_b, h_3 \otimes I_b)$ ,  $[\cdot]$  sets to zero the negative values,  $s$  represents the scale,  $N$  is the total number of scales considered,  $\varepsilon$  is a constant used to avoid division by zero (typically  $\varepsilon = 0.01$ ), and  $T_s$  is an orientation independent threshold that controls the spurious responses to noise at scale  $s$  [22]. FA computes the step edge strength at a specific voxel and provides a value in the range  $[0, 1]$ , thus providing a direct measure of the strength at a specific voxel. The region-of-interest (ROI) highlighted in red in Fig. 4 shows an area with speckle texture while the ROI highlighted in green shows an area containing fetal brain structures (part of the falx). Most importantly, even within areas of apparent weak image appearance (red ROI in Fig. 4); FA identifies the voxels which

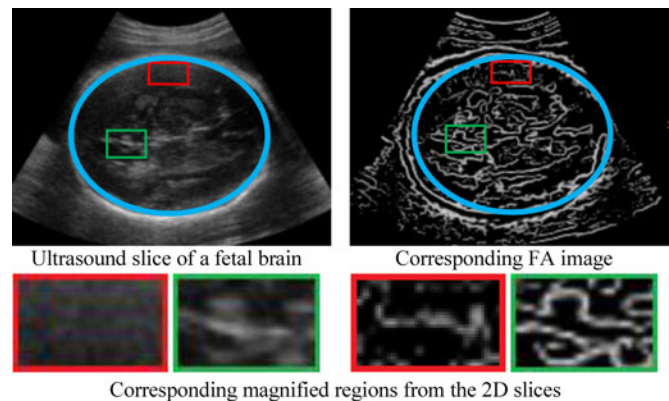


Fig. 4. Localization of informative voxels. Blue ellipse represents the brain region which we sample from. Red is an ROI with less informative voxels while green is an ROI with highly informative voxels. Ultrasound slice of a fetal brain. Corresponding FA image. Corresponding magnified regions from the 2-D slices.

may be discriminative. Fig. 4 motivates the choice of using FA measure to represent voxels informative strength.

We use FA to create stronger trees during training by: 1) masking out all voxels which have FA measure below a specific threshold and 2) using the FA value to weight the importance of each training example when finding a splitting score in a tree node.

The traditional function that is optimized is

$$\arg \min \text{err} = \sum_{v \in S} (z - \bar{z})^2 - \sum_{i \in \{L, R\}} \sum_{v \in S^i} (z - \bar{z})^2 \quad (5)$$

where  $z$  is the variable that defines the optimal split of a training set. Instead, we normalize the traditional variable by the FA measure, i.e.,  $z' = \frac{z}{FA}$ . Note that we divide rather than multiply since we are minimizing an objective function. We find the best split of a training set  $S$  ( $v$  is a training example) to left ( $L$ ) and right ( $R$ ), denoted ( $S_L$  and  $S_R$ ).  $\bar{z}$  is the mean value of  $z$  for all training points reaching a specific tree node.

### C. Feature Sets

Careful selection of features is an important part in any machine learning algorithm. In the medical image analysis literature, several types of features have been used [11], [17], [18], [25]–[29]. In this study, we use traditional appearance features which have shown promising results in other medical imaging domains [17], [29]. However, because of the complexity of our problem, we propose a constrained distance feature set to aid appearance features when localizing the PS plane.

1) *Appearance Features*: Appearance features are computed within a neighborhood of each voxel. We deploy features like voxel intensity, mean intensity within a cuboid and difference of mean intensity of two cuboids. These types of features have been successfully used in related classification and RF applications and showed promising results [17], [29].

2) *Constrained Distance Features*: The use of appearance features within RF creates a learner which works well if image appearance is clear and not ambiguous. However, in ultrasound imaging this is not always the case as previously discussed. To

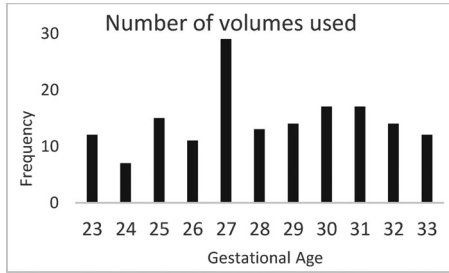


Fig. 5. Number of volumes used for each gestational age considered within this study.

address this issue we introduce a geometric constraint expressed as a distance feature to help constrain the RF training. This type of feature embeds the perpendicular distance between the voxel of interest and the midsagittal plane. Given the midsagittal plane normal  $n = (n_x, n_y, n_z)$ , a point on the plane  $p = (p_x, p_y, p_z)$  and a point in space  $v = (v_x, v_y, v_z)$ , the perpendicular distance  $D$  between  $v$  and the plane parameterized by  $n$  and  $p$  can be expressed as

$$D = |n \cdot w| / |n| \quad \text{where} \quad w = v - p. \quad (6)$$

This feature clusters training examples depending on their distance from the midsagittal plane. However, this distance feature alone is not sufficient to capture the variability of plane appearance in different fetuses especially when dealing with different gestational ages. Therefore, it is important to incorporate both appearance and distance features inside the RF and to let the optimization of the objective function choose a feature at a specific depth depending on its ability to split the training points properly.

#### D. Localizing the Parasagittal Plane

In the testing stage, the informative voxels in a test image traverse the trees until each voxel ends up in a leaf. Only voxels that reach confident leaves are allowed to vote for the output plane parameters. The confidence level is measured using the variance of the training examples that reached a specific leaf during training. The mean parameters of all informative voxels reaching confident leaves are used to output the final plane parameters by fitting a mean plane which represents the estimated parasagittal plane position.

### IV. EXPERIMENTAL SETUP

#### A. Dataset and Manual Annotation

161 3-D fetal brain ultrasounds between 23 and 33 weeks of gestation from normal fetuses were acquired with a Philips HD9 ultrasound machine [5]. These fetal brain volumes were not acquired specifically for our study but for standard fetal biometry, and hence were not optimized to our application. The number of volumes used in each gestational week is presented in Fig. 5. Volumes vary in size but they are approximately  $(250 \times 250 \times 200)$  voxels. There were no specific settings used to acquire the 3-D images because these images were acquired for

a clinical study [5] and before the start of our analysis study. We did not use volumes for fetuses older than 33 weeks because this study was set up to correlate to neonatal scans of preterm babies, and the age range of interest was until 33 weeks of gestation, as these are defined to be more critical than later on.

The quantitative evaluation reported in this paper is more extensive than in our earlier work [14]. Two clinicians participated in the evaluation by manually selecting the PS and MS planes. Clinician 1 manually localized the PS plane twice to derive intraobserver variability. Clinician 2 manually localized the PS plane once to investigate interobserver variability. Our model was trained by the second manual localization of the PS planes performed by the clinician 1. A strict protocol to find the PS plane was defined prior to manual selection of planes. Manual localization of the MS plane was carried out once to allow us to investigate the accuracy of detecting this plane automatically.

#### B. Evaluation Measures

To evaluate the accuracy of plane-finding, two evaluation metrics were used. The first is the angular distance between the manual and the detected planes defined as

$$\theta = \cos^{-1} \frac{n_m \cdot n_d}{|n_m| |n_d|} \quad (7)$$

where  $n_m$  is the normal to the manually defined plane and  $n_d$  is the normal to the automatically found plane. The angular distance provides insight into how close two planes are in the 3-D space but does not take into consideration their line of intersection - two planes may have a small angle between them but their intersection can be far outside the brain which means the error in diagnostic plane finding is large. To circumvent this problem, we include, as a second evaluation metric, the Euclidian distance between the centers of the manual and the automatic PS planes. Since the fetal brain appears approximately in the middle of the volume in most cases (given this requirement is part of the scanning protocol), the distance metric expresses how far the center of the two planes are from each other. The smaller the distance and angle, the higher the accuracy. In addition, we measured the angle between the manual and automatic MS planes as well as the distance between the center of the MS plane to the center of the volume.

To evaluate the PS plane localization, we report six comparisons: 1) manual-to-manual plane agreement, which includes intra and interexpert plane comparison; 2) manual-to-automatic plane agreement with the RF trained on appearance features only; 3) manual-to-automatic plane agreement using the RF trained on appearance features only but constrained by the use of informative voxels (CRF-FA); 4) manual-to-automatic plane agreement using RF trained on all examples but constrained by automatic MS plane distance features (CRF-Dist); 5) manual-to-automatic plane agreement using RF constrained by the use of informative voxels and automatic MS plane distance features (CRF-FA-Dist-A); and finally, 6) manual-to-semiautomatic plane agreement using RF constrained by the use of informative voxels and manual MS plane distance features (CRF-FA-Dist-M).



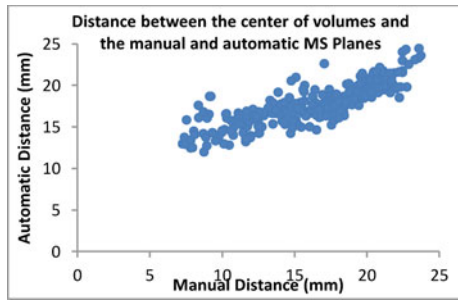


Fig. 6. Distribution of distances for the manual MS planes (x-axis) and the automatic MS planes. The distance is measured as the distance between the center of the volume to the center of the plane.

### C. Implementation Details

We performed three-fold cross-validation to ensure that every volume is included in testing. In all experiments we optimized the method parameters on the classic RF and fixed them for the proposed RFs to allow a fair comparison. We used 15 trees with a maximum depth of 20. We optimized 300 candidate features during node creation. We set the FA threshold to 0.2 to allow, as much as possible, informative voxels during the training stage. In addition, when comparing multiple FA thresholds, visual inspection showed that the chosen threshold provides a suitable spread of voxels in different images. This implies informative voxels from inside the brain were selected during the training and testing stages. We used the 3-D Gaussian derivative [21] of scale  $\sigma$ , as a filter to compute FA [see (4)] and we experimentally set the scales to  $\sigma = 4$ . Finally, we used the C# language to build the methods in a parallel fashion.

## V. RESULTS

### A. MS Plane Localization

We compared the manual MS planes to the automatically generated ones. The mean angle between the two was found to be  $5.6 \pm 2.6^\circ$ . Although the angle between two planes provides an insight on how close they are to each other, it does not provide a physical distance between them. Therefore, we measured the distance between the center of the volume and the center of the manual and automatic MS planes. Fig. 6 shows the correlation between the manual and automatic distances. The mean of the absolute difference between the manual and automatic distances is  $2.1 \pm 1.9$  mm. The correlation coefficient between the distances is 0.83 (p-value < 0.005).

### B. Manual to Manual PS Plane Localization

The first two rows in Table I show the mean angle and distance between the manual planes localized by the same clinician (M-M intra) and by two clinicians (M-M inter), respectively. M-M intra reflects the intraobserver variability in detecting the PS plane while M-M inter evaluates the interobserver variability.

### C. Automatic PS Plane Localization

Table I presents the mean and standard deviation of angles and distances between different PS plane localization methods.

TABLE I  
MEAN  $\pm$  STANDARD DEVIATION OF ANGLES AND DISTANCES BETWEEN PLANES

	Angle ( $^\circ$ )	Distance (mm)
M-M Intra	$6.8^\circ \pm 2.6$	$8.6 \pm 4.5$
M-M Inter	$10.7^\circ \pm 4.1$	$11.0 \pm 4.4$
M-RF	$16.9^\circ \pm 7.8$	$12.5 \pm 4.1$
M-CRF-FA	$11.2^\circ \pm 4.0$	$11.1 \pm 3.9$
M-CRF-Dist	$11.1^\circ \pm 4.2$	$11.1 \pm 3.6$
M-CRF-FA-Dist-M	$9.3^\circ \pm 3.3$	$9.5 \pm 3.8$
M-CRF-FA-Dist-A	$10.3^\circ \pm 3.7$	$10.2 \pm 4.0$

M-M intra is manual to manual planes localized by the same clinician. M-M inter is manual to manual planes localized by two clinicians. M-RF is manual to automatic planes (localized by the traditional RF method). M-CRF-FA is manual to automatic planes (localized by the RF method that is trained using informative examples). M-CRF-Dist is manual to automatic planes (localized by the RF method that is constrained by the automatic MS plane). M-CRF-FA-Dist-M is manual to automatic planes (localized by the RF method that is trained on informative examples and constrained by the manual MS plane). M-CRF-FA-Dist-A is manual to automatic planes (localized by the RF method that is trained on informative examples and constrained by the automatic MS plane).

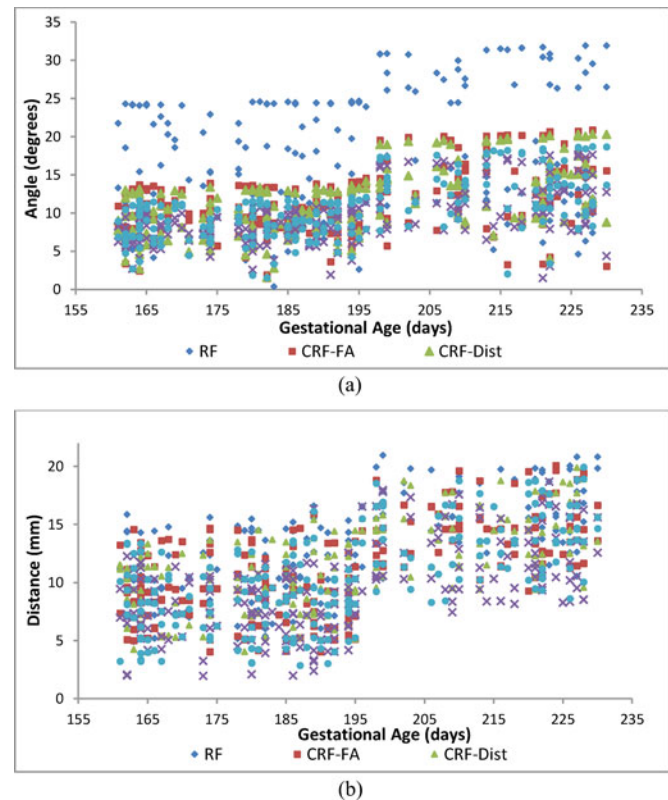


Fig. 7. Distribution of angles and distances on all fetal volumes for the different RFs with respect to the fetal gestational age.

The best result is achieved using the semiautomatic technique CRF-FA-Dist-M; this is the constrained RF which is trained on informative examples and distance features from the manual MS planes. The automatic version of this (CRF-FA-Dist-A) achieves comparable results while being automatic. To investigate the effect of gestational age on the localization of the PS plane, Fig. 7(a) shows the distribution of angles between manual and automatic methods at all gestational ages we have. We also

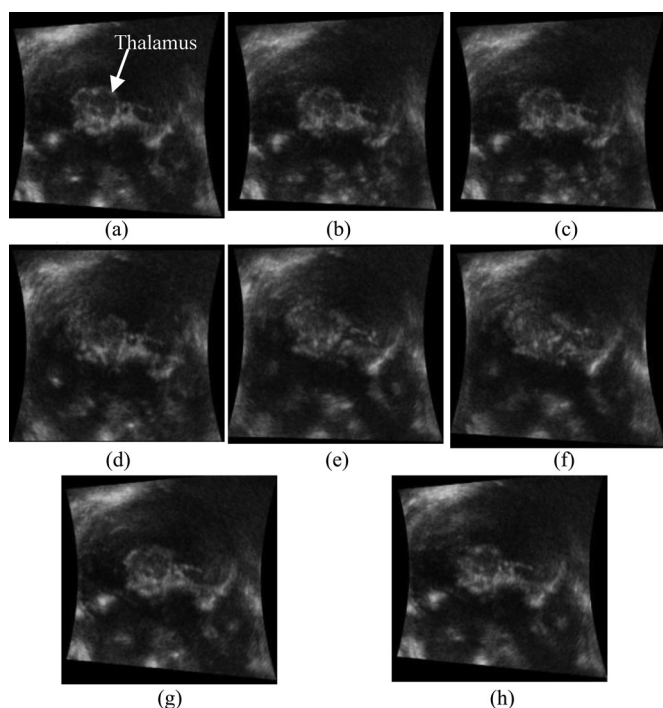


Fig. 8. Visual comparison between the three manual PS planes (a)–(c) and the automatic ones (d)–(h) on one fetal brain volume. Manual 1 and manual 2 were localized by clinician 1 while manual 3 was generated by clinician 2. (a) Manual 1. (b) Manual 2. (c) Manual 3. (d) RF. (e) CRF-FA. (f) CRF-Dist. (g) CRF-FA-Dist-M. (h) CRF-FA-Dist-A.

present in Fig. 7(b) the distribution of distances for the proposed methods. Finally, Fig. 8 shows the PS planes localized manually and by the proposed methods for visual comparison.

## VI. DISCUSSION

### A. Importance of the Preprocessing

Since the quality of the ultrasound images suffers from several issues like shadowing, weak edges, speckle, we have employed the monogenic signal to create suitable features for analysis. This is not a major new contribution to this paper as this form of preprocessing has been done before. However, it is an important step which contributes to the success of our method.

### B. MS Plane Localization

We presented an automatic method to detect the MS plane in 3-D ultrasound images of the fetal brain. We first locate the skull and extract the brain region. We then localize the midline at several slices. Midlines at several brain slices are then used to fit the best plane which represents the estimated MS plane. Evaluation results confirm that we can detect the MS plane with high accuracy.

### C. Automatic PS Plane Localization

As shown in Table I and Fig. 7, the classic RF struggles to perform well especially at later gestations. The use of informa-

tive voxels and voxels within the brain (CRF-FA) improves the PS plane localization results compared to the conventional RF. Adding the newly proposed MS plane distance feature (CRF-Dist) also improves localization over the classic RF. However, both CRF-FA and CRF-Dist are still less accurate than the interobserver manual planes agreement. On the other hand, combining both ideas (CRF-FA-Dist-A) not only improves the localization accuracy but also provides better localization than the interobserver variability obtained by two clinicians.

The best proposed method is CRF-FA-Dist-M; however, it is semiautomatic. In addition, it is marginally better than its automatic equivalent (CRF-FA-Dist-A). Therefore, taking into account its accuracy and no need for user interaction, CRF-FA-Dist-A is the most promising method of all.

Note that in Fig. 7, the accuracy of the localization of the PS plane decreases with increased gestational age. This is mainly because fetal brains of later gestation are bigger and therefore the thalamus and its surrounding structures are larger in size and hence present more room for localization errors. In fact, if we compare our previous work in [14] with our current study, we find that the localization accuracy in [14] using CRF-FA-Dist-M is better than here since that paper used a smaller gestational age range (23–27 weeks) and smaller number of fetal brain volumes (87 volumes).

## VII. CONCLUSION AND FUTURE WORK

This paper introduces for the first time an automatic method to localize the PS plane in 3-D ultrasound volumes of the fetal brain. This is a very difficult task for a sonographer to do well. Results show that the automatic method provides more precise and reproducible PS planes than the interexpert variability between two clinicians.

One way to possibly improve the MS plane localization would be to detect the midline on axial and coronal slices and then fit the best plane between these. This could provide a more accurate orientation of the MS plane since it is detected using different views. The improvement on the MS plane detection could also improve the PS plane detection as suggested by Table I where best results were obtained with manual MS plane detection in the main algorithm.

Although we showed that finding the PS plane is important to facilitate the measurement of the thalamus, the PS plane contains several other structures which would be interesting to assess in the future, such as the lateral ventricle wall and choroid plexus. In addition, we applied our method to volumes from one ultrasound machine, but the method is generalizable to images from other ultrasound machines and potentially other imaging modalities (e.g., MRI). Finally, to validate the utility as a tool for biomarker assessment, a thalamic area comparison on manual and automatic planes needs to be investigated. This is the subject of future clinical work.

## ACKNOWLEDGMENT

The authors would like to thank the International Fetal and Newborn Growth Consortium, INTERGROWTH-21st, Nuffield



Department of Obstetrics and Gynaecology, John Radcliffe Hospital, University of Oxford, Oxford, U.K., who provided the medical images.

## REFERENCES

- [1] A. Kopuri, M. Yaqub, S. Rueda, P. Sullivan, K. McCormick, and A. Noble, "Cranial ultrasound derived "thalamic area" as a marker for brain growth in premature infants and comparison with similar markers from a normal fetal population," in *Proc. Pediatric Acad. Soc. Asian Soc. Pediatric Res.*, 2014.
- [2] G. Ball, J. P. Boardman, D. Rueckert, P. Aljabar, T. Arichi, N. Merchant *et al.*, "The effect of preterm birth on thalamic and cortical development," *Cerebral Cortex*, vol. 22, pp. 1016–1024, May 2012.
- [3] T. Moore, E. M. Hennessy, J. Myles, S. J. Johnson, E. S. Draper, K. L. Costeloe, N. Marlow, "Neurological developmental outcome in extremely preterm children born in England in 1995 and 2006: The EPICure studies," *BMJ*, vol. 345, 2012.
- [4] J. J. Volpe, "Brain injury in premature infants: A complex amalgam of destructive and developmental disturbances," *Lancet Neurology*, vol. 8, pp. 110–124, 2009.
- [5] A. T. Papageorghiou, E. O. Ohuma, D. G. Altman, T. Todros, L. C. Ismail, A. Lambert, Jaffer, A. Yasmin, E. Bertino, M. G. Gravett, M. Purwar, J. A. Noble, R. Pang, C. G. Victora, F. C. Barros, M. Carvalho, L. J. Salomon, Z. A. Bhutta, S. H. Kennedy, J. Villar, "International standards for fetal growth based on serial ultrasound measurements: The fetal growth longitudinal study of the INTERGROWTH-21st project," *Lancet*, vol. 384, pp. 869–879, 2014.
- [6] J. M. Rennie, *Neonatal Cerebral Ultrasound*. Cambridge, U.K.: Cambridge Univ. Press, 1997.
- [7] G. Malinger, A. Monteagudo, G. Pilu, I. E. Timor-Tritsch, and A. Toi, "Sonographic examination of the fetal central nervous system: Guidelines for performing the "basic examination" and the "fetal neurosonogram"," *Ultrasound Obstetrics Gynecology*, vol. 29, pp. 109–116, 2007.
- [8] A. I. L. Namburete, R. V. Stebbing, B. Kemp, M. Yaqub, A. T. Papageorghiou, and J. Alison Noble, "Learning-based prediction of gestational age from ultrasound images of the fetal brain," *Med. Image Anal.*, vol. 21, pp. 72–86, 2015.
- [9] G. Rizzo, A. Capponi, M. E. Pietrolucci, A. Capece, E. Aiello, S. Mammarella, D. Arduini, "An algorithm based on OmniView technology to reconstruct sagittal and coronal planes of the fetal brain from volume datasets acquired by three-dimensional ultrasound," *Ultrasound Obstetrics Gynecology*, vol. 38, pp. 158–164, 2011.
- [10] L. Xiaoguang, B. Georgescu, Z. Yefeng, J. Otsuki, and D. Comaniciu, "AutoMPR: Automatic detection of standard planes in 3D echocardiography," in *Proc. Int. Symp. Biomed. Imag.*, 2008, pp. 1279–1282.
- [11] J. Domingos, E. Lima, P. Leeson, and A. Noble, "Local phase-based fast ray features for automatic left ventricle apical view detection in 3D echocardiograph," in *Proc. 3rd Int. MICCAI Workshop, Med. Comput. Vis.*, pp. 119–129, 2013.
- [12] K. Chykeyuk, M. Yaqub, and A. Noble, "Class-specific regression random forest for accurate extraction of standard planes from 3D echocardiography," in *Proc. 3rd Int. MICCAI Workshop, Med. Comput. Vis.*, pp. 53–62, 2013.
- [13] M. Sofka, J. Zhang, S. Good, S. Zhou, and D. Comaniciu, "Automatic detection and measurement of structures in fetal head ultrasound volumes using sequential estimation and integrated detection network (IDN)," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1054–1070, May 2014.
- [14] M. Yaqub, A. Kopuri, S. Rueda, P. Sullivan, K. McCormick, and J. A. Noble, "A constrained regression forests solution to 3d fetal ultrasound plane localization for longitudinal analysis of brain growth and maturation," in *Machine Learning in Medical Imaging*. New York, NY, USA: Springer, 2014, pp. 109–116.
- [15] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*. Boston, MA, USA: Chapman & Hall, 1984.
- [16] A. Criminisi and J. Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*. New York, NY, USA: Springer, 2013.
- [17] A. Criminisi, J. Shotton, D. Robinson, and E. Konukoglu, "Regression forests for efficient anatomy detection and localization in CT studies," in *Proc. MICCAI Workshop Med. Comput. Vis.*, pp. 106–117, 2010.
- [18] O. Pauly, B. Glocker, A. Criminisi, D. Mateus, A. M. Moller, S. Nekolla *et al.*, "Fast multiple organs detection and localization in whole-body MR dixon sequences," in *Proc. Med. Image Comput. Comput. Assisted Intervention Conf.*, pp. 239–247, 2011.
- [19] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001.
- [20] M. Mulet-Parada and J. A. Noble, "2D+T acoustic boundary detection in echocardiography," *Med. Image Anal.*, vol. 4, pp. 21–30, 2000.
- [21] D. Boukerroui, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *J. Math. Imag. Vis.*, vol. 21, pp. 53–80, 2004.
- [22] P. Kovess, "Image features from phase congruency," *J. Comput. Vis. Res.*, vol. 1, pp. 1–26, 1999.
- [23] V. Grau, H. Becher, and A. Noble, "Phase-based registration of multi-view real-time three-dimensional," in *Proc. Med. Image Comput. Comput. Assisted Intervention Conf.*, 2006, pp. 612–619.
- [24] K. Rajpoot, A. Noble, V. Grau, and N. Rajpoot, "Feature detection from echocardiographic images using local phase information," in *Proc. Med. Image Understanding Anal.*, pp. 13–17, 2008.
- [25] K. Chykeyuk, M. Yaqub, and J. A. Noble, "Novel context rich LoCo and GloCo features with local and global shape constraints for segmentation of 3D echocardiograms with random forests," in *Medical Computer Vision: Recognition Techniques and Applications in Medical Imaging*. Berlin, Germany: Springer, 2013, pp. 59–69.
- [26] A. Criminisi, K. Juluru, and S. Pathak, "A discriminative-generative model for detecting intravenous contrast in CT images," in *Proc. Med. Image Comput. Comput. Assisted Intervention Conf.*, pp. 49–57, 2011.
- [27] A. Criminisi, J. Shotton, and S. Bucciarelli, "Decision forests with long-range spatial context for organ localization in CT volumes," in *Proc. MICCAI Workshop Probabilistic Models Med. Image Anal.*, 2009.
- [28] A. Montillo, J. Shotton, J. Winn, J. E. Iglesias, D. Metaxas, and A. Criminisi, "Entangled decision forests and their application for semantic segmentation of CT images," in *Proc. Inform. Process. Med. Imag.*, pp. 184–196, 2011.
- [29] M. Yaqub, M. K. Javaid, C. Cooper, and J. A. Noble, "Investigation of the role of feature selection and weighted voting in random forests for 3-d volumetric segmentation," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 258–271, Feb. 2014.

Authors' photographs and biographies not available at the time of publication.