Towards Reliable Automatic Characterization of Neonatal Hip Dysplasia from 3D Ultrasound Images

Niamul Quader, Antony Hodgson, Kishore Mulpuri, Anthony Cooper, Rafeef Abugharbieh
BiSICL, University of British Columbia, Vancouver, Canada
niamul@ece.ubc.ca

Abstract. Ultrasound (US) imaging is recommended for early detection of developmental dysplasia of the hip (DDH), which includes a spectrum of hip joint abnormalities in infants. However, the currently standard 2-dimensional (2D) US-based approach to measuring the dysplasia metric (DM), namely the $\alpha$ angle, suffers from high within-hip variability with standard deviations typically ranging between $3^\circ \sim 7^\circ$. Such high variability leads to elevated over- and under-treatment rates in hip classification. To reduce this high variability inherent to the 2D $\alpha$ angle, $\alpha_{2D}$, we propose a 3D US-based DM in the form of a 3D $\alpha$ angle, $\alpha_{3D}$, that more accurately characterizes the morphology of an infant’s hip joint. Our method leverages phase symmetry features that automatically identify the 3D bone/cartilage structures to compute $\alpha_{3D}$. Validating on 30 clinical patient hip examinations, we demonstrate the within-hip variability of $\alpha_{3D}$ to be significantly smaller than $\alpha_{2D}$ (28.9% reduction, $p < 0.01$). Our findings indicate that $\alpha_{3D}$ may be significantly more reproducible than the conventional 2D measure, which will likely reduce misclassification rates.

1 Introduction

Developmental dysplasia of the hip (DDH), which refers to hip joint abnormalities ranging from mild acetabular dysplasia to irreducible hip joint dislocation, affects 0.16% – 2.85% of all newborns [1]. Early arthritis is often associated with DDH [2] so failing to detect and treat DDH in infancy can lead to later expensive corrective surgical procedures. Based on the figures presented in [2], Price et. al [3] estimated that 25,000 total hip replacements per year are attributable to missed early diagnosis in the United States alone. At approximately $50,000 per procedure [4], the direct financial impact of this problem is in the order of $1B/year, not considering the costs of subsequent revision surgeries or other socioeconomic costs.

To diagnose DDH prior to ossification of the femoral head, 2-dimensional (2D) ultrasound (US) imaging is currently recommended over other imaging modalities (e.g. x-ray, magnetic resonance imaging, computed tomography, etc.)
due to its low cost and absence of ionizing radiation [5]. The standard DDH metric obtained from 2D US scans is the angle between the acetabular roof and the vertical cortex of the ilium, referred to as the alpha angle, $\alpha_{2D}$ [5,6]. In general, $\alpha_{2D} > 60^\circ$ indicates a normal hip, whereas $43^\circ < \alpha_{2D} < 60^\circ$ represents borderline to moderate DDH, and $\alpha_{2D} < 43^\circ$ suggests severe DDH [6,7]. However, $\alpha_{2D}$ suffers from high within-hip variability, i.e. the variability between dysplasia metrics (DMs) measured on repeated examinations of the same patient’s hip, with standard deviations of such measurements, $\sigma$, ranging from $3^\circ$ to $7^\circ$ [8]. This may be partly attributable to variations in manually measuring $\alpha_{2D}$ on 2D slices (subjective variability, $\sigma \approx 5^\circ$), but is more likely due to variability in $\alpha_{2D}$ resulting from differences within what is considered clinically acceptable (standard) 2D US scans that are caused mainly by differences in the probe orientation (probe-orientation-dependent variability, $\sigma \approx 7^\circ$) [7]. This high variability in measured $\alpha_{2D}$ leads to significant discrepancies between the initial clinical determination of dysplasia severity and later clinical assessments. Specifically, estimates suggest that 6% – 29% of cases that are later treated were initially regarded as not needing early treatment [9]. Further, there is significant potential for over-treatment since about 90% of US-detected hip dysplasia cases resolve spontaneously [10]. Recently, we have proposed to reduce the subjective variability by automatically extracting $\alpha_{2D}$ from 2D US [11]. Our preliminary results in that work showed a 9% reduction in within-hip variability [11].

To further reduce the within-hip variability, in this paper, we address the crucial probe-orientation-dependent variability problem [7]. More specifically, we propose to characterize DDH based on an intrinsically 3D morphology metric derived directly from 3D US scans, which we argue captures more of the pertinent anatomical structures, while reducing the dependency on probe orientation in imaging those structures. To the best of our knowledge, only one previous work [12] has proposed the use of an intrinsically 3D DM, the acetabular contact angle (ACA). Similar to $\alpha_{2D}$, the ACA represents the angular separation between the acetabular roof ($A$) and the lateral iliac ($I$), except that the ACA is based on the segmented 3D surfaces of $A$ and $I$. Hareendranathan et al.’s [12] method involves a slice-by-slice analysis process that requires manually selecting 4 seed points in each of the 2D US slices in a 3D US volume and manually separating $A$ from $I$. Using such an interactive method would require valuable clinician time and the manual operations introduce within-image measurement variability of approximately $1^\circ$ [12] and inter-scan variability of approximately $4^\circ$ [13].

In this paper, we propose a fully automatic approach for extracting a new 3D DM, the 3D alpha angle, $\alpha_{3D}$, by analogy to a $\alpha_{2D}$ [6]. To the best of our knowledge, our work is the first that proposes a fully automatic approach of extracting a 3D dysplasia metric. In this paper, we: (1) extend our previous phase symmetry feature-based bone/cartilage extraction [11] to 3D, (2) define our new proposed 3D metric, $\alpha_{3D}$, (3) automatically extract $\alpha_{3D}$, (4) demonstrate on real clinical data a significant decrease in within-hip variability of $\alpha_{3D}$ compared to $\alpha_{2D}$. 
2 Methods

The 2D dysplasia metric, $\alpha_{2D}$, is defined as the angle between the fitted straight lines that approximate $A$ and $I$ when viewed on a 2D B-mode US image \[5,6\]. We therefore define an analogous 3D metric, $\alpha_{3D}$, based on the relative orientations of the fitted planar surfaces of $A$ and $I$ (Fig. 1b). Briefly, given a 3D B-mode US image, $U : X \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, where $X = (x, y, z)$ are the voxel coordinates, our approach starts by extracting the bone cartilage structures, $B$ (section 2.1). We then use prior anatomical knowledge of the hip joint to automatically identify the 3D surfaces of $A$ and $I$ within $B$ (section 2.2). Finally, we approximate the average normals across $A$ and $I$, and compute $\alpha_{3D}$ as the angle between these approximated normals (section 2.3).

2.1 Extraction of Bone/Cartilage Structures

To extract the hyperechoic bone/cartilage boundaries, we first compute the local phase symmetry feature, $PS$, from $U$ using responses to a band-pass quadrature filter bank \[14\]. To segment the sheet-like bone and cartilage surfaces from the $PS$ feature volume, we deploy a multi-scale eigen analysis of the Hessian matrix. For eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, voxels on sheet-like structures will exhibit $|\lambda_1| \approx |\lambda_2| \approx 0$, $|\lambda_3| \gg 0$. We enhance the sheet-like features in our $PS$ volume as:

\[
SPS = \begin{cases} 
0, & \text{if } \lambda_3 > 0, \\
(1 - \exp(-R_a^2))(\exp(-R_a^2))(1 - \exp(-S^2)), & \text{otherwise.} \end{cases} 
\]  

(1)

where $R_a = \text{abs}(2 * |\lambda_3| - |\lambda_2| - |\lambda_1|) / |\lambda_3|$ is a blob eliminating term, $R_a = |\lambda_2| / |\lambda_3|$ is a sheet enhancing term and $S = \sqrt{|\lambda_1|^2 + |\lambda_2|^2 + |\lambda_3|^2}$ is a noise cancelling term \[15\].

Bone/cartilage boundaries also tend to attenuate an US beam more than other neighboring structures (e.g. soft-tissue, etc.). We further enhance the bone/cartilage structures in $SPS$ to form $B$ (fig. 1e) using an attenuation-based method \[14\]. Specifically, the enhancement process extracts attenuation and shadowing features from $U$, which are then used to modify $SPS$ such that the bone/cartilage boundaries are enhanced and outliers (e.g. soft-tissue, etc.) are removed \[14\].

2.2 Localization of the 3D Ilium and Acetabulum Surfaces

Our proposed $\alpha_{3D}$ is based only on the surfaces of $A$ and $I$, which are both substructures of the detected $B$ volume. To isolate $A$ and $I$ from other background structures, we use hip anatomy-based priors; the first prior is that $A$ and $I$ are located superior to the spherical femoral head, $F$ (Fig. 1(a,e,f)), while the second prior is that $A$, $I$ and the labrum (cartilage) tend to have a common
junction at the edge of the ilium (Fig. 1(e,g)). We thus start by detecting \( F \), which we estimate by fitting a sphere (with radius \( r \) and center \( c = [c_x, c_y, c_z] \)) in the \( B \) volume using an M-estimator SAmple Consensus (MSAC) algorithm [16]. In fitting the sphere, we set the maximum allowable distance from an inlier point to the sphere as 1mm based on the empirical observation that average radius of an infant’s femoral head is around 10mm [6]. Next, we locate the edge of the ilium, \( i = [i_x, i_y, c_z] \), based on the fact that the bone/cartilage structures around such a junction point, \( i \), has a heterogeneous gradient direction in the image volume. Specifically, we calculate a direction-variability feature (Fig. 1g), \( d \), that captures the variability in orientations (angular values of eigenvectors corresponding to \( \lambda_1 \)) of bone/cartilage surfaces (voxels with \( B > 0 \)). We define \( d \) as,

\[
d = \sum_k \text{std}(\tan^{-1}(u_k/v_k)),
\]

where \( k \in x \) represents the neighboring voxels, and \( u_k \) and \( v_k \) are the \( x \) and \( y \) direction components of eigenvectors corresponding to \( \lambda_1 \). Amongst the voxels located superior to \( c \), we expect the directional variability \( d \) to have the highest response around \( i \) (Fig. 1g); consequently, we assign \( i = [i_{d,x}, i_{d,y}, c_z] \), where \( i_{d,x} \) and \( i_{d,y} \) are the \( x \) and \( y \) coordinates corresponding to the maximum value of \( d \) that is located superior to \( c \).

Having located \( F \) (radius \( r \), center \( c = [c_x, c_y, c_z] \)) and \( i = [i_{d,x}, i_{d,y}, c_z] \), we extract cubic regions-of-interest (ROIs) around \( A \) and \( I \) that are centered at \([i_{d,x} + r/2, i_{d,y} - r/2, c_z]\) and \([i_{d,x} - r/2, i_{d,y}, c_z]\), respectively, with sides of length \( r \) (Fig. 1h). These ROIs simplify the fitting of planes that approximate \( A \) and \( I \) by reducing background structures and by lowering computation complexity.

### 2.3 Computation of the 3D \( \alpha \) Angle

Once we have identified the 3D surfaces of \( A \) and \( I \), we proceed to fit planes that approximate \( A \) and \( I \), which we then use to estimate \( \alpha_{3D} \). We start by evaluating the 3D Radon transforms, \( R_A \) and \( R_I \) of \( A \) and \( I \), respectively. The 3D Radon transform, \( R \), maps an input in \( \mathbb{R}^3 \) into a set of its plane integrals in \( \mathbb{R}^3 \), where the orientations of the planes are represented by elevation angles, \( \theta \), and azimuth angles, \( \phi \) [17]. We estimate the best-fit plane for an input at \( R_m = \max(R) \), i.e., the plane for which the plane integral is maximum, with orientations \( \theta_m, \phi_c \) and unit normal vector \( n = [\cos\theta_m, \sin\theta_m, \sin\phi_m, \cos\phi_m] \) [17]. We use the resulting unit normal vectors \( n_A \) and \( n_I \) to calculate \( \alpha_{3D} \) using the formula

\[
\alpha_{3D} = \cos^{-1}(n_A \cdot n_I)/(|n_A||n_I|).
\]

### 3 Results and Discussion

**Data Acquisition and Experimental Setup:** In this study, two orthopedic surgeons at British Columbia Children’s hospital participated in collecting 3D US images from 30 hip examinations (15 US sessions \( \times 2 \) hips/US session = 30 hip examinations) belonging to 12 patients (single US session for 9 patients and two US sessions for 3 patients), obtained as part of routine clinical care under the appropriate institutional review board approval. To investigate repeatability,
i.e., within-hip variability, each hip examination consisted of five repeated 3D US image acquisitions. Our proposed $\alpha_{3D}$ angles were automatically calculated for each of the five 3D US volumes. Furthermore, each of the infants underwent an independent regular clinical care 2D US scanning session at the radiology department, where the infants were scanned by the radiologist on duty. In every 2D US session, the radiologist would acquire repeated 2D US images and make manual measurements of $\alpha_{2D}$ (3 to 6 measurements per hip) on all images that the radiologist judged to adequately show the key anatomical structures needed to measure $\alpha_{2D}$; subsequently, these were recorded in and retrieved from the patient chart. As in a typical 2D US scan session (Fig. 2(a,b)), the 3D scans were acquired in the coronal plane, where the ilium (located superior to the femoral head, Fig. 1a) appears towards the left of the femoral head in an US image (Fig. 1(e,f)).

**Validation Scheme:** We compared the automatically extracted measure, $\alpha_{3D}$, with the currently standard manual 2D measurement, $\alpha_{2D}$. Specifically, we analyzed the following for each of the 30 hip examinations: (1) the correlation between the averages of $\alpha_{2D}$ and $\alpha_{3D}$ on individual hips (i.e., between $\text{mean}(\alpha_{2D})$ and $\text{mean}(\alpha_{3D})$), (2) the discrepancy between the averages of $\alpha_{2D}$ and $\alpha_{3D}$ (i.e., $\text{mean}(\alpha_{2D}) - \text{mean}(\alpha_{3D})$), and (3) the within-hip variability, $\sigma$, in the measurements (i.e., $\text{std}(\alpha_{2D})$ and $\text{std}(\alpha_{3D})$) on individual hips.
Correlation between $\alpha_{2D}$ and $\alpha_{3D}$: In terms of agreement, $\text{mean}(\alpha_{3D})$ showed good correlation with $\text{mean}(\alpha_{2D})$ across all the 30 hip examinations (Fig. 3a, correlation coefficient $R = 0.87$ (95% confidence interval: 0.74 to 0.94)). This suggests that the proposed $\alpha_{3D}$ measures an aspect of morphology similar to what the current $\alpha_{2D}$ measures.

Discrepancy between $\alpha_{2D}$ and $\alpha_{3D}$: The difference between the two metrics, $\text{mean}(\alpha_{2D}) - \text{mean}(\alpha_{3D})$ in the 30 hip examinations, was significant ($p < 0.01$, mean: 5.17°, SD: 3.33° with a bias towards $\alpha_{3D}$ being smaller than $\alpha_{2D}$).

Fig. 2: Qualitative results. (a), (b), (d) and (e) show example variability of $\alpha_{2D}$ and $\alpha_{3D}$ from two 2D and two 3D US images from a hip examination of patient 3DUS004 ($\alpha_{2D} = 47°$ and 56°, $\alpha_{3D} = 44.8°$ and 45.2°). The higher variability in the input 2D US images (and $\alpha_{2D}$ values) can be seen in the manually aligned 2D US images in (e) compared to the variability in the manually aligned 3D US images (and $\alpha_{3D}$ values) in (f).

Variability of Metrics, $\sigma$: The automatic $\alpha_{3D}$ shows a statistically significant improvement in variability compared to the manual $\alpha_{2D}$ ($p = 0.0053$, $\text{mean}(\sigma_{3D}) = 2.19°$, $\text{mean}(\sigma_{2D}) = 3.08°$ (qualitative result in Fig. 2 and box-plot in Fig. 3b)). This 28.9% reduction in variability suggests that probe position variation has
a larger effect on variability in the DM than manual processing of the 2D US (9% improvement with automatic image processing within a 2D US in our previous study [11]). The residual variability of $\alpha_{3D}$ ($\sigma_{3D} \approx 2^\circ$) seems to be small enough to be diagnostically valuable, given that the typical range from normal to dysplastic hip is around $17^\circ$ [6,7]. Furthermore, the variability of $\alpha_{3D}$ appears substantially lower than the reported variability of the recently described 3D ACA metric ($2.19^\circ$ versus $4.1^\circ$ [13]).

**Computational considerations:** The complete process of extracting $\alpha_{3D}$ from an US volume took approximately 270 seconds, when run on a Xeon(R) 3.40 GHz CPU computer with 12 GB RAM. All processes were executed using MATLAB 2015b. Current practice has a sonographer process the images post-acquisition, so this computation time is not a significant barrier to implementation. Although not critical for clinical use, we plan to work towards optimizing our code to reduce this computation time.

## 4 Conclusions

We presented an automatic 3D dysplasia metric, $\alpha_{3D}$, to characterize hip dysplasia in 3D US images of the neonatal hip. Using the proposed $\alpha_{3D}$ resulted in a statistically significant reduction in variability compared to the currently standard 2D measure, $\alpha_{2D}$. This suggests that this 3D morphology-derived DM could be valuable in improving the reliability in diagnosing DDH, which may lead to a more standardized DDH assessment with better diagnostic accuracy.

Notably, the improvement in reliability associated with the 3D scans was achieved by orthopaedic surgeons, who have limited training in performing US examinations, while the 2D scans and metrics were obtained from radiologists
with explicit training in ultrasound acquisition and analysis. This strongly suggests that we may, in future, be able to train personnel other than radiologists to obtain reliable and reproducible dysplasia metrics using 3D ultrasound machines, potentially reducing the costs associated with screening for DDH.

References