



Using 3D Ultrasound to Improve Reliability in Assessing Developmental Dysplasia of the Hip

Niamul Quader¹, Antony J Hodgson², Kishore Mulpuri³, Anthony Cooper³, Emily Schaeffer³ and Rafeef Abugharbieh¹

¹ Biomedical Signal and Image Computing Lab - Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada

² Department of Mechanical Engineering, University of British Columbia, Vancouver, Canada

³ Department of Orthopaedics, University of British Columbia, Vancouver, Canada

Abstract

Many infants are born with unstable hips which can cause severe mobility issues as they grow older. This condition is known as developmental dysplasia of the hip (DDH), and although it is the most common hip disorder in infants, its diagnosis using 2D ultrasound (US) is prone to significant variability. This paper presents methods to reliably and automatically diagnose DDH in infants using 3D US. Based on 40 infant hip examinations, we demonstrate a 70% reduction in variability of 3D US-based measurements compared to those of standard 2D US-based measurements.

1 Introduction

Developmental dysplasia of the hip (DDH) in infancy refers to a spectrum of hip joint abnormalities in infants. While 2D B-mode ultrasound (US)-based dysplasia metrics are currently used clinically to diagnose DDH, such estimates suffer from high inter-exam variability (or low agreement, e.g. an intra-class correlation coefficient for hip dysplasia class of 0.19-0.43 (Orak, 2015)) and can lead to mistakes in diagnosis. Jaremko (Jaremko, 2014) recently noted that this high variability is primarily due to sensitivity of dysplasia metrics to the positioning of the 2D probe. To reduce this sensitivity of dysplasia metrics on probe positioning, we propose quantifying the acetabular deformity and femoral head irregularities, as recommended by the American College of Radiology (ACR) (ACR, 2012), using two novel dysplasia metrics: α_{3D} for measuring acetabular deformity, and FHC_{3D} for measuring femoral head malposition. We argue that automatically extracting these 3D US-based dysplasia metrics could provide a more reliable diagnosis compared to conventional α_{2D} and FHC_{2D} dysplasia metrics.

2 Methods

As part of routine clinical care, we acquired 3D and 2D B-mode US images from 23 left hips and 17 right hips from 25 infants, all obtained with appropriate institutional review board approvals. Our collaborating clinicians (six surgeons) at BC Children’s Hospital acquired the 2D US and 3D US images along the standard coronal plane. While scanning patients, the acquisition time for acquiring both 2D US and 3D US scans in all the infants was recorded.

Figure 2 presents a block diagram outlining our proposed automatic extraction of α_{3D} , which we define as the angle between the planar approximations of the ilium’s vertical cortex, I , and the acetabular roof, A , and FHC_{3D} , defined as the ratio of the volume of the femoral head portion medial to I to the total volume of the femoral head. Since α_{3D} is computed via detecting the bone boundaries A and I , we first use a machine learning approach (random-forests-based classifier) to automatically identify all coronal plane slices in an US volume $U(x, y, z)$ that contain A and I , which we denote as $U(x, y, z_{A,I})$ (step A1, figure 1). Subsequently, we deploy a local symmetry and attenuation-based technique that we recently developed (Quader 2016 MICCAI) to identify the bone boundaries A and I in $U(x, y, z_{A,I})$ (step A2). This is followed by an M-estimator SAmple Consensus (MSAC) algorithm (Torr, 2000) that finds planar approximations of A and I and then computed α_{3D} as the angle between the approximated planes of A and I (step A3).

To estimate FHC_{3D} , we compute a voxel-wise probability map, $P(x, y, z)$, characterizing the likelihood of a voxel belonging to the femoral head (step B1) and then calculate FHC_{3D} as the ratio of the sum of P values medial to I to the total sum of P values across the entire coordinate space of $U(x, y, z)$ (step B2).

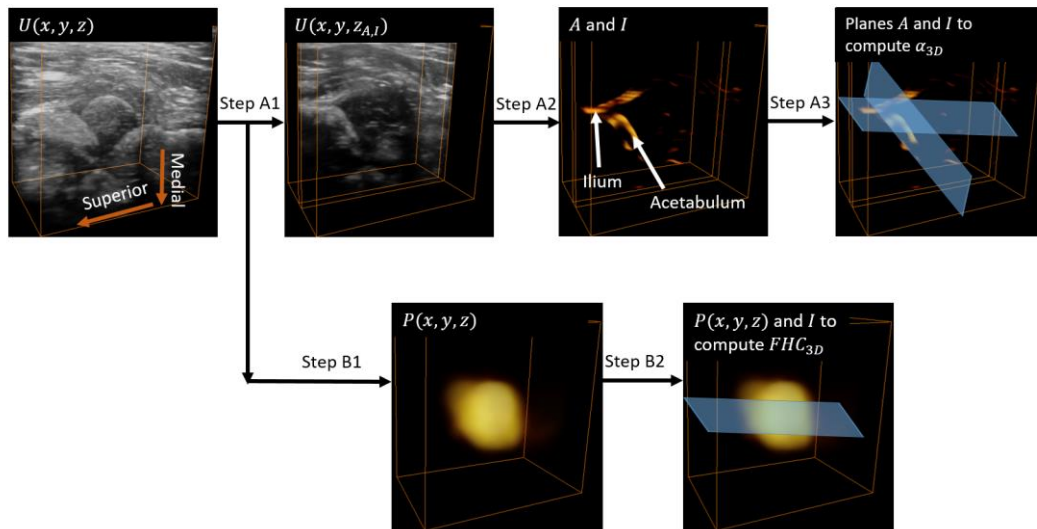


Figure 1: Overview of our proposed automatic extraction of 3D dysplasia metrics, α_{3D} and FHC_{3D} . Steps A1-A3 were used in computing α_{3D} and steps B1-B2 were used in computing FHC_{3D} .

3 Results

Our results show that the 3D US-based dysplasia metrics have significantly lower variability compared to 2D dysplasia metrics as measured by orthopaedic surgeons (figure 2). Intra-rater variability

of FHC_{3D} was 63% lower than FHC_{2D} at 3.19% for FHC_{3D} vs. 8.56% FHC_{2D} with $p < 0.01$), and inter-rater variability of FHC_{3D} was 65% lower than FHC_{2D} at 2.72% for FHC_{3D} vs. 7.79% FHC_{2D} with $p < 0.01$, intra-rater variability of α_{3D} was 73% lower than α_{2D} , at 2.2° for α_{3D} vs. 8.33° for α_{2D} with $p < 0.01$, and inter-rater variability of α_{3D} was 78% lower than α_{2D} at 2.35° for α_{3D} vs. 10.63° for α_{2D} with $p < 0.01$. Agreement between the FHC_{2D} and FHC_{3D} was moderate with the correlation coefficient, r , at 0.4 (95% CI 0.23 and 0.55), with similar agreement between α_{2D} and α_{3D} where $r = 0.41$ (95% CI 0.24 and 0.56).

We found no significant difference between the acquisition time for 3D US images vs. acquisition time for 2D US images, where the average was 23.0 seconds for 3D US data and 25.4 seconds for 2D data ($p = 0.4$).

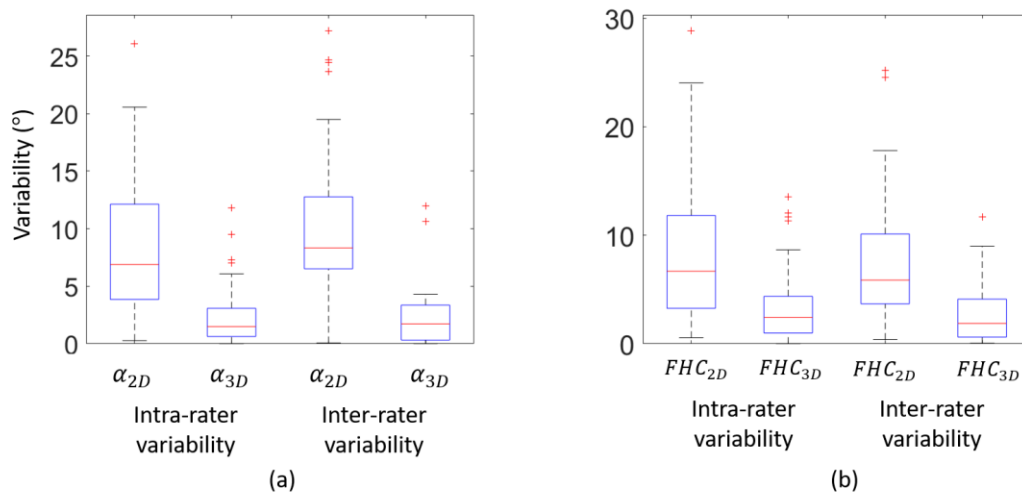


Figure 2: Box-plot of the intra- and inter-rater variability for (a) α_{3D} vs. α_{2D} and (b) FHC_{3D} vs. FHC_{2D} . The box limits indicate the 1st and 3rd quartiles and the whiskers correspond to 99.3% coverage if the data are normally distributed.

4 Discussion and Conclusions

We demonstrated that both the intra- and inter-rater variabilities in the automatic measurements using our proposed 3D US-based dysplasia metrics were significantly lower than those of the standard 2D US-based counterparts with a $\sim 75\%$ reduction for α and $\sim 65\%$ reduction for FHC. Compared to the semi-automatic method proposed by Hareendranathan for computing a 3D-US-based acetabular contact angle metric (ACA) (Hareendranathan, 2016), we found that the intra-rater variability of our proposed α_{3D} is 45% lower than the value reported for the ACA (2.2° for α_{3D} vs. 4° for ACA). This significant reduction in variability comes at no extra cost in terms of US acquisition time in 3D.

We are encouraged by our promising findings; nonetheless, our scans were acquired with a research engineer present in addition to a clinician. We have thus yet to demonstrate that our proposed system can be operated equally efficiently solely by a clinician. Hence, it remains to be studied whether one could recommend clinical adoption of the system in its current form.

References

American College of Radiology. (2012). ACR-AIUM practice guideline for the performance of the ultrasound examination for detection and assessment of developmental dysplasia of the hip (ACR guidelines).

Graf, R., Mohajer, M., & Plattner, F. (2013). Hip sonography update. Quality-management, catastrophes-tips and tricks. *Medical ultrasonography*, 15(4), 299.

Hareendranathan, A. R., Mabee, M., Punithakumar, K., Noga, M., & Jaremko, J. L. (2016). A technique for semiautomatic segmentation of echogenic structures in 3D ultrasound, applied to infant hip dysplasia. *International journal of computer assisted radiology and surgery*, 11(1), 31-42.

Jaremko, J. L., Mabee, M., Swami, V. G., Jamieson, L., Chow, K., & Thompson, R. B. (2014). Potential for change in US diagnosis of hip dysplasia solely caused by changes in probe orientation: patterns of alpha-angle variation revealed by using three-dimensional US. *Radiology*, 273(3), 870-878.

Orak, M. M., Onay, T., Çağırılmaz, T., Elibol, C., Elibol, F. D., & Centel, T. (2015). The reliability of ultrasonography in developmental dysplasia of the hip: How reliable is it in different hands?. *Indian journal of orthopaedics*, 49(6), 610.

Quader, N., Hodgson, A., Mulpuri, K., Cooper, A., & Abugharbieh, R. (2016, October). Towards Reliable Automatic Characterization of Neonatal Hip Dysplasia from 3D Ultrasound Images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 602-609). Springer International Publishing.

Quader, N., Hodgson, A. J., Mulpuri, K., Cooper, A., & Abugharbieh, R. (2017, September). A 3D Femoral Head Coverage Metric for Enhanced Reliability in Diagnosing Hip Dysplasia. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 100-107). Springer, Cham.

Torr, P. H., & Zisserman, A. (2000). MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1), 138-156.