Visual-Assisted Probe Movement Guidance for Obstetric Ultrasound Scanning using Landmark Retrieval: Supplementary Material

Cheng Zhao[1], Richard Droste[1], Lior Drukker[2], Aris T. Papageorghiou[2], and J. Alison Noble[1]

[1]Institute of Biomedical Engineering, University of Oxford
[2]Nuffield Department of Women's Reproductive Health, University of Oxford
Oxford, United Kingdom
cheng.zhao@eng.ox.ac.uk

1 Network Training Details

The network is implemented using Pytorch, and it is trained on the NVIDIA TITAN RTX GPU (24GB memory) accelerated by CUDA and cuDNN. The network is trained for 40 epochs with a batch size of 1. The number of anchor, positive and negative US images in one batch are set to 1, 20 and 20 respectively for the triplet loss, and that are set to 1, 1 and 40 respectively for the InfoNCE loss. A step learning policy is adopted and the learning rate decay is fixed to 0.1 applied on every 10 epochs. The initial learning rate is 1e-4 and the momentum is fixed to 0.9. The network is trained via Adam optimization, and gradient clipping is adopted during training to increase the training stability. During inference, the required memory footprint is around 3500MB and the inference time is around 0.01s deployed on the NVIDIA TITAN RTX GPU.

2 Landmark Sharing

The training and testing data was captured on different second/third-trimester fetal scans from different women (Sec 3.1). This data shares the same fetal anatomical landmarks which we use in the performance evaluation using image average recall. We achieve ultrasound-probe localization via global descriptor retrieval as the global descriptor is more robust to subject and probe position image appearance variations than direct ultrasound image intensity matching.

3 Fetal Position Dynamics

Landmarks are at fixed 3D anatomical locations on the fetus as shown in Fig.1. If the fetal 3D pose changes, or the probe moves, the retrieved landmark may change. Thus, if the current retrieved landmark is "nose", the probe position is shown on the nose. If the fetus rotates 90 degrees, the probe position doesn't change, but the retrieved landmark will change to "ear". The probe position is then visualized on the ear. We achieve probe localization via landmark searching not 3D probe position prediction in a world co-ordinate system. This alleviates the need to accommodate fetal position dynamics.

4 Motivation against 3D Ultrasound Imaging

Our motivation is to support uptake of the new generation of lower-cost 2D ultrasound (US) probes outside of traditional hospital setting, where users are not highly trained. Assistive guidance tools are potentially useful then. 3D US is typically based in hospitals and performed by highly trained sonographers. This is why we pose the problem as a 2D US one. The proposed method could be adapted to the 3D US case. The only required architecture modification is to replace the 2D CNN encoder with a 3D CNN encoder for 3D data.