Ultrasound in Med. & Biol., Vol. ■, No. ■, pp. 1–15, 2014

0301-5629/$ - see front matter

http://dx.doi.org/10.1016/j.ultrasmedbio.2014.06.006

● *Original Contribution*

STANDARD PLANE LOCALIZATION IN ULTRASOUND BY RADIAL COMPONENT MODEL AND SELECTIVE SEARCH

Dong Ni,* Xin Yang,* Xin Chen,* Chien-Ting Chin,* Siping Chen,* Pheng Ann Heng,† Shengli Li,‡ Jing Qin,*§ and Tianfu Wang*

*National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, School of Medicine, Shenzhen University, Shenzhen, China; †Department of Computer Science and Engineering, Chinese University of Hong Kong, Hong Kong, China; ‡Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare Hospital of Nanfang Medical University, Shenzhen, China; and §Center for Human Computer Interaction, Shenzhen Institute of Advanced Integration Technology, Shenzhen, China

(*Received* 25 *January* 2014; *revised* 25 *April* 2014; *in final form* 4 *June* 2014)

Abstract—Acquisition of the standard plane is crucial for medical ultrasound diagnosis. However, this process requires substantial experience and a thorough knowledge of human anatomy. Therefore it is very challenging for novices and even time consuming for experienced examiners. We proposed a hierarchical, supervised learning framework for automatically detecting the standard plane from consecutive 2-D ultrasound images. We tested this technique by developing a system that localizes the fetal abdominal standard plane from ultrasound video by detecting three key anatomical structures: the stomach bubble, umbilical vein and spine. We first proposed a novel radial component-based model to describe the geometric constraints of these key anatomical structures. We then introduced a novel selective search method which exploits the vessel probability algorithm to produce probable locations for the spine and umbilical vein. Next, using component classifiers trained by random forests, we detected the key anatomical structures at their probable locations within the regions constrained by the radial component-based model. Finally, a second-level classifier combined the results from the component detection to identify an ultrasound image as either a "fetal abdominal standard plane" or a "non- fetal abdominal standard plane." Experimental results on 223 fetal abdomen videos showed that the detection accuracy of our method was as high as 85.6% and significantly outperformed both the full abdomen and the separate anatomy detection methods without geometric constraints. The experimental results demonstrated that our system shows great promise for application to clinical practice. (E-mail: lishengli63@126.com and harryqinjingcn@gmail.com or tfwang@szu.edu.cn) © 2014 World Federation for Ultrasound in Medicine & Biology.

Key Words: **Ultrasound, Standard plane, Fetal abdomen, Component, Selective search, Machine learning, Object detection.**

INTRODUCTION

Ultrasound (US) is prevalent in routine clinical examinations and is widely used for pregnancy diagnosis because of its relatively low cost, real-time imaging capability and avoidance of radiation exposure. Generally, the pipeline of US-based diagnosis includes three sequential steps: scanning, selection of the standard plane and diagnosis (Bamber et al. 2012). Of these, acquisition of the standard plane is crucial for the subsequent biometric measurements and diagnosis. However, this step often requires substantial experience and a thorough knowledge of human anatomy. Therefore, it is very challenging for novices and even difficult and time consuming for experienced examiners. It is often cited as one of the inherent disadvantages of US compared with other imaging modalities such as computed tomography and magnetic resonance imaging (Zhang et al. 2012). Hence, the development of automatic methods for localizing the standard planes would enhance the ability of non-experts to operate US devices as well as improve the examination efficiency for experts (Kwitt et al. 2013).

Over the past few decades, a considerable number of research studies have contributed to the automization and computerization of the diagnosis step. Examples include automatic biometric measurement for the evaluation of

Address correspondence to: Shengli Li, Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare Hospital of Nanfang Medical University, Shenzhen, China. Jing Qin and Tianfu Wang, School of Medicine, Shenzhen University, Shenzhen, China. E-mail: lishengli63@126.com and Jing Qin and Tianfu Wang, School of Medicine, Shenzhen University, Shenzhen, China. E-mail: harryqinjingcn@gmail.com or tfwang@szu.edu.cn

fetal growth (Lu et al. 2005; Yu et al. 2008), plaque segmentation for quantitative analysis (Cheng et al. 2013) and image registration for improved visualization of residual brain tumors (Mercier et al. 2013). However, little attention has been paid to the prerequisite step in US diagnosis, that is, acquisition of the standard plane. Recently, some automatic methods have been proposed for localizing standard planes from 3-D US (Carneiro et al. 2008; Lu et al. 2008; Rahmatullah et al. 2011). Although 3-D US is a promising imaging modality for prenatal examination, currently and in the near future, 2-D exams are still being widely used in clinical practice because of better imaging quality, the wider availability of 2-D scanners and the preference of experienced users (Elliott 2008). A few recent studies have investigated methods for localizing standard planes from 2-D US images. Zhang et al. (2012) proposed selecting the standard plane of the gestational sac from US video based on a cascade AdaBoost classifier and local context information. However, the proposed system may not perform well in the detection of standard planes of more complicated anatomical structures, which often have high intra-class variations and typically contain multiple key anatomical structures (KASs) instead of only one KAS (*i.e.,* the gestational sac). Rahmatullah et al. (2012) described a method for detecting anatomical structures from manually extracted abdominal US images by integrating global and local features. This method is semi-automatic, and the detection accuracy reported is still not good enough for clinical use. Kwitt et al. (2013) proposed localizing target structures from US videos by building kernel dynamic texture models, which were evaluated on US videos acquired from three different phantoms. Because actual patient data are much more complex than phantom data, further investigation is required to evaluate the efficacy of their proposed method.

Although previous works have illustrated the promise of automatically localizing standard planes from 2-D US images, this task remains very challenging for several reasons. We take the fetal abdominal standard plane (FASP) as an example to explain these challenges. First, as illustrated in Figure 1, the FASP is determined by the simultaneous presence of three KASs: the stomach bubble (SB), umbilical vein (UV) and spine (SP) (Dudley and Chapman 2002). However, the FASP class often has high intra-class variations because of artifacts and/or deformations, as well as differences in image intensity, gestational age, fetal posture and scanning orientation, all of which make it difficult to detect the FASP as a whole. Second, the simple, separate detection of KASs may not perform well because large numbers of other structures may appear similar to the target KASs. For example, shadows, the abdominal aorta (AO), the inferior vena cava and the hypo-echoic spinal cord are often mistakenly detected as the SB or UV. Third, traditional detection methods, which use sliding window algorithms, may exhaustively scan the entire image (Říha et al. 2013; Zhang et al. 2012) and can thus be extremely time consuming, especially when multiple KASs must be detected.

Recently, component-based methods (Felzenszwalb et al. 2010; Mohan et al. 2001), which search for an object by looking for its identifying components rather than scanning the whole object, have been proposed for detecting objects composed of multiple components and with high intra-class variability (*e.g.*, people) and have proved to be more robust than full-object detection methods. Inspired by this research, we proposed detecting FASPs from consecutive 2-D US images using a hierarchical supervised learning framework. We first designed a novel radial component-based model (RCM) to describe the geometric constraints of the KASs. To improve detection speed and accuracy, we further proposed a novel selective search strategy to produce probable locations for the SP and UV. After each KAS was detected separately by its random forest (RF) (Breiman 2001) classifier (*i.e.*, component detector) based on both the RCM and a selective search method, a second-level classifier combined the results of the component detectors to identify an US image as either a "FASP" or a "non-FASP." The experimental results indicated that the proposed method outperformed existing detection methods in detection accuracy and speed.

METHODS

Figure 2 illustrates the pipeline of the method. We first detected the region containing the abdomen (region of interest [ROI]) using its RF classifier from each frame of the US video to reduce the search range and to exclude regions that appeared similar to the KASs. Next, three distinctive component detectors were trained by the RFs to detect the SB, UV and SP at the probable locations generated by a selective search scheme within the regions constrained by the RCM. During the detection procedure, the component detector iteratively changed its size to find the optimal size of the KAS. Finally, we employed a second-level support vector machine (SVM) classifier to combine the results from the component detectors and identify an US image as either a "FASP" or a "non-FASP."

Training of random forest classifiers

Recently, RFs have emerged as a novel type of classifier that has several advantages, including a non-parametric nature, high classification accuracy, ease of parallelization and the ability to determine variable