Ultrasound in Med. & Biol., Vol. ■, No. ■, pp. 1–12, 2017

0301-5629/$ - see front matter

ELSEVIER

http://dx.doi.org/10.1016/j.ultrasmedbio.2017.04.022

● *Original Contribution*

RANDOM FOREST-BASED BONE SEGMENTATION IN ULTRASOUND

Nora Baka,* Sieger Leenstra,[†] and Theo van Walsum*

*Biomedical Imaging Group Rotterdam, Departments of Radiology & Nuclear Medicine and Medical Informatics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands; and [†]Department of Neurosurgery, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

(*Received* 19 *August* 2016; *revised* 14 *April* 2017; *in final form* 24 *April* 2017)

Abstract—Ultrasound (US) imaging is a safe alternative to radiography for guidance during minimally invasive orthopedic procedures. However, ultrasound is challenging to interpret because of the relatively low signal-to-noise ratio and its inherent speckle pattern that decreases image quality. Here we describe a method for automatic bone segmentation in 2-D ultrasound images using a patch-based random forest classifier and several ultrasound specific features, such as shadowing. We illustrate that existing shadow features are not robust to changes in US acquisition parameters, and propose a novel robust shadow feature. We evaluate the method on several US data sets and report that it favorably compares with existing techniques. We achieve a recall of 0.86 at a precision of 0.82 on a test set of 143 spinal US images. (E-mail: t.vanwalsum@erasmusmc.nl) © 2017 World Federation for Ultrasound in Medicine & Biology.

Key Words: **Ultrasound, Machine learning, Spine, Vertebra, Intra-operative, Ultrasound guidance, Orthopedic procedure.**

INTRODUCTION

Ultrasound (US) imaging is a safe alternative for guidance during minimally invasive orthopedic procedures. Its main advantage compared with X-ray guidance is the lack of ionizing radiation and its cost-effectiveness. However, ultrasound imaging has its own challenges, such as the relatively low signal-to-noise ratio, its inherent speckle pattern, shadowing and several types of artifacts. US guidance is therefore mainly performed by registering the acquired US images to the pre-operative computed tomography (CT) image on which the intervention was planned. Such registration usually requires the bone surface from both modalities, US and CT (Nagpal et al. 2015; Penney et al. 2006). Automatic bone detection algorithms are crucial for such navigation. In this article, we propose and evaluate a method for such automatic bone tissue interface detection from US. We propose learning the appearance of bone interfaces in the images based on annotated training examples and machine learning methods.

Address correspondence to: Theo van Walsum, PO Box 2040, 3000 CA Rotterdam, The Netherlands. E-mail: t.vanwalsum@erasmusmc.nl

There have been several published approaches to solving bone classification from ultrasound, most of them using heuristic functions to calculate bone and non-bone interfaces. The most obvious heuristic is that the bone surface appears bright in the images. An additional intensity correction using the expected depth of the bony structure can suppress other bright interfaces, as proposed in Kowal et al. (2007), to effectively highlight bones. The down side of this method is that the expected depth of the bone must be known; otherwise, noise or soft tissue interfaces will be enhanced, and could be mistaken for bone. Hacihaliloglu et al. (2009) described a different approach, suited for bones at all depths. They proposed using phase symmetry in the frequency domain to find and enhance edges regardless of their brightness. This was determined to be very accurate in finding the bone outline. However, the method enhances all lines in the image, including fat–muscle or other soft-tissue interfaces. A property most widely used to distinguish bone for soft tissue interfaces is shadowing. Because of the large difference in acoustic properties of bone and soft tissue, almost all the sound energy is reflected from the bone surface. The lack of sound traversal through the bone creates the shadow, a region of dark intensities below the bone surface. Karamalis et al. (2012) proposed an algorithm to quantify the chance of sound reaching every image pixel, which can be used as an indicator for shadowing. Quader et al. (2014) combined the two features, phase symmetry and the shadow feature of Karamalis et al., and improved bone detection accuracy. Indeed, multiple properties can be combined to reliably characterize bone in US. Jain and Taylor (2004) proposed a Bayesian framework combining intensity, gradient, shadow, intensity profile along scanline and multiple reflections for bone segmentation; however, how the required conditional probabilities can be obtained is not straightforward. Foroughi et al. (2007) proposed a heuristic combination of intensity, shadow and the Laplacian filtered image to derive a probability map for bones. In a second step, a maximum of one pixel per image column was selected as bone, producing the final segmentation with dynamic programming. This post-processing method became popular in the field, as it could correct for small errors in bone probability images. A similar method was used by Jia et al. (2016) and Cao et al. (2016) with different heuristically calculated feature images. Jia et al. (2016) proposed calculating the bone probability images by multiplying in total seven feature images, including integrated back scattering and local energy. Although these methods exhibited good accuracy on their respective published test data, one might wonder if they are optimal given the manually created cost function.

Learning the combination and importance of features from training data seems a more structured way of characterizing bone interfaces in US. Penney et al. (2006) proposed learning the distribution function of bones versus background using two features, bone intensity and an artifact distance. Any pixel with an intensity <40 was defined as artifact. More recently, Berton et al. (2016) combined the bone probability feature of Foroughi et al. (2007), Hacihaliloglu et al. (2009), Hellier et al. (2010) with local binary pattern (LBP) and Gabor filtering for shadow, bone and soft tissue differentiation. In their work, they used a linear classifier on the already heuristically combined features.

In this article, we propose learning the bone probability map from simple features, using a patch-based classification approach. The contributions of this work are as follows:

- We present a bone segmentation scheme using a patch-based classification approach and perform an extensive evaluation of the method.
- We propose a novel shadow feature and evaluate it in comparison with different shadows and other features.
- We compare the presented method to the standard heuristic methods from the literature.
- We include multiple ultrasound data sets to assess classification robustness.

This work extends our previous work in which we compared linear and non-linear classifiers for bone segmentation from ultrasound images (Baka et al. 2016). This study differed in that it used a simplified classifier, proposed and evaluated a new shadow feature and provided more extensive evaluations between methods, between data sets and for parameters within the method.

METHODS

Our aim is to segment the bone interfaces from US images. For this, we propose learning a classifier from a training set of annotated US images for bone segmentation. Once the classifier is learned, the bone probability map of an unseen image can be computed. The classification is done as follows. First, the image undergoes a preprocessing step. Subsequently, the feature images are computed. Each pixel of the image and the patch around it are then fed into a classifier, which gives the probability of that pixel being part of a bone–soft tissue interface. This step thus results in the bone probability map. If a single interface line segmentation is desired rather than a probability map, a dynamic programming post-processing step for segmentation can be added, as in Foroughi et al. (2007). Below we describe each part of the method in detail.

Random forest classifier

Random forest classifiers are non-linear classifiers consisting of several decision trees, first proposed by Ho (1995). The output of the forest is the average prediction of its trees. To ensure that the trees are sufficiently dissimilar, every tree is trained on a subset of features and on a subset of data. Each tree consists of a series of nodes, that can either branch into two child nodes with a splitting rule or be a leaf node. When learning the tree, at each yet unsplit node, a splitting rule is computed, which best separates the positive and negative samples arriving at that node. At test time, the new sample is passed through each tree according to the splitting rules and ends up in a leaf node. The output probability of the sample from a tree is then equal to the percentage of positive training samples that are in that leaf node.

Random forest classifiers work well with a large number of features, and with selection of the best feature during training for splitting each node, they have an inherent feature selection property. This makes them good candidates for patch-based learning, as we propose for US segmentation. In this work, all the pixels surrounding a US pixel in an $n \times n$ window are taken as feature candidates. Additionally, differential features calculated by downsampling the window to a size of 5×5 and