Evaluation of automatic breast MR anatomy segmentation


1Philips Healthcare, Paris, France, 2Philips Healthcare, Best, Netherlands, 3Department of Radiology, The University of Chicago, Chicago, Illinois, United States

Introduction:
MR breast imaging is an emerging technology for detecting and characterizing of breast tumors. During breast MRI examination, all the findings are reported in a standardized way using BI-RADS descriptions. Such a report not only contains the lesions types, but also their positioning relative to the anatomy (breast laterality, o’clock position relative to the nipple, distance from the skin, etc). The reporting automation requires fast automatic detection and segmentation of the breast anatomy, which is particularly difficult given the large variation in anatomical details and MR imaging protocols. In this context of a Breast MR CAD application, we developed a hierarchical approach for automatically detecting breasts regions, segmenting the patient skin, localizing the nipples, and the band of cardiac artifacts when present. This work was focused on automating and facilitating the reporting task. In this paper we will discuss the evaluation of the breast detection algorithm.

Methods:
During Breast MRI examinations, a series of dynamic contrast-enhanced T1-weighted volumes is acquired; typically 4 to 8 volumes are acquired 1-2 minutes apart. Depending on the acquisition protocol, those volumes may vary largely in orientation (axial, sagittal or coronal), and aspect (breast compression, fat tissue saturation, resolution, number of slices, etc.). This variability makes the segmentation task particularly challenging, and requires very robust algorithms.

In our method, first, we analyze image statistics for separating image background from the breast tissues. Then, we exploit anatomical rules for finding breast positions and extensions (volumes of interest), robust to the image acquisition orientation (Figure 1A). Mathematical morphology is applied to extract skin surface (Figure 1B), from which local curvature analysis allows to detect the nipples (Figure 1C), which is visualized schematically (Figure 1D). Depending on the acquisition orientation, a band of cardiac artifacts can be detected and discarded for an improved visualization.

A preliminary clinical validation of the breast detection algorithm has been done by three physicians on 30 data sets, with one physician leaving out one of the sets. The results of the breast segmentation were presented to the physician as regions of interest on coronal, axial and sagittal reformats of the pre-contrast image similar to figure 1A. The physicians could adapt the position and the size of the regions of interest by manual interaction. The values found by the algorithm and the corrected values of the regions of interest for both breasts were stored for the evaluation.

Results:
Our segmentation algorithms have been technically validated on a database of 118-cases proving the algorithms to be stable on the different acquisition protocols. The results of the clinical validation show an overall good performance of the breast segmentation algorithm with a tendency to a slight over-segmentation of the inferior region. The success rate is determined from the overlap of the algorithm found and the user determined regions of interest. The success rate of the algorithm is 93% with a standard deviation of 9%. When comparing this with the inter-observer success rate (93%, 8% standard deviation) we can conclude that the algorithm performs similarly well.

Conclusions and future work:
We have described a method for segmenting various breast anatomy parts from MR dynamic contrast enhanced T1-weighted images of the breast, including volume of interest, skin, nipples and the band of cardiac artifacts. The algorithms were designed to cope with the large variation in anatomical details and MR imaging protocols. The preliminary clinical validation of the volumes of interest has shown a good performance of the breast detection algorithm, also compared to variability of the performance of the different users. This result is encouraging to use the algorithm in the clinical routine.

Reference: