Automated detection of motion in breast DCE-MRI to assess study quality and prevent unnecessary call-backs

Poster No.: C-1845
Congress: ECR 2015
Type: Scientific Exhibit
Authors: L. Wang¹, A. Gubern-Mérida², O. Diaz³, Y. Diez³, R. M. Mann², S. Diekmann¹, F. Zöhrer¹, H. Laue¹, J. Schwaab⁴, ¹Bremen/DE, ²Nijmegen/NL, ³Girona/ES, ⁴Heidelberg/DE
Keywords: Breast, MR, Screening, Computer Applications-Detection, diagnosis, Cost-effectiveness, Quality assurance
DOI: 10.1594/ecr2015/C-1845

Any information contained in this pdf file is automatically generated from digital material submitted to EPOS by third parties in the form of scientific presentations. References to any names, marks, products, or services of third parties or hypertext links to third-party sites or information are provided solely as a convenience to you and do not in any way constitute or imply ECR's endorsement, sponsorship or recommendation of the third party, information, product or service. ECR is not responsible for the content of these pages and does not make any representations regarding the content or accuracy of material in this file.

As per copyright regulations, any unauthorised use of the material or parts thereof as well as commercial reproduction or multiple distribution by any traditional or electronically based reproduction/publication method ist strictly prohibited. You agree to defend, indemnify, and hold ECR harmless from and against any and all claims, damages, costs, and expenses, including attorneys' fees, arising from or related to your use of these pages.

Please note: Links to movies, ppt slideshows and any other multimedia files are not available in the pdf version of presentations.

www.myESR.org
Aims and objectives

Dynamic contrast enhanced (DCE) MRI is used for breast cancer screening examinations for women of high risk for developing breast cancer. Motion caused e.g. by muscle relaxation or coughing during image acquisition can reduce the interpretability of breast MRI. For scans with strong motion, it might be necessary to repeat the scan, but in typical screening workflow this is only detected by the radiologist when the woman has already left the clinic. As a consequence, the woman might need to be recalled for a repeated scan. In this work, a fully automated tool based on image-processing is proposed to detect and quantify motion for unambiguous scan quality evaluation before the woman leaves the clinic.
Methods and materials

Data set

The breast MRI data set was provided by Radboud University Nijmegen Medical Centre (RUNMC) and extracted randomly from original breast cancer screening data. In total, 491 screening exams acquired with the current screening MR protocol of RUNMC [1] from 449 patients were supplied. An expert radiologist manually annotated the clinical image data concerning motion artefacts focusing on the high resolution images of time point t0 (pre-contrast) and t1 (first post-contrast), as these were assumed to be representative for the whole data set. The motion artefact was categorised in four classes: no (1), mild (2), moderate (3), or severe (4) motion. For simplification, the classification was transformed into a two class categorisation. Cases were considered "positive" if they showed moderate or severe motion. The rating results are shown in figure 1. Generally, in this text, "positive" refers to an image showing the considered artefact, while "negative" indicates that the image is not affected by this artefact, which means that these two terms do not refer to any diagnostic outcome concerning lesions or tumours.

Automated detection of motion

As the standard DCE breast MRI protocol takes several minutes, it is quite possible that the woman moves slightly during the examination. As shown in figure 2, motion between time 0 (pre-contrast as t0) and the following sequences (post-contrast as t1…tn) leads to blurred maximum intensity projections (MIPs), as they are based on the subtraction of t0. As only the motion between t0 and t1 was rated by the radiologist, we only consider t0 and t1 as a simplification in this first approach.

To quantify the motion artefact between t0 and t1 (not within one single acquisition), prominent edges delineating the boundary contours of parenchyma, skin and pectoral muscle were detected in both t0 and t1 studies (see figure 3(a)). Irrelevant motion that occurred in the thorax, such as heart, lung or liver motion, was excluded by a fully automatic breast segmentation, which confines edge detection only to the inside of the breast region. Due to the administration of contrast agent and the consequent enhancement of previously faint structures, normally more edges would be captured in t1 compared to t0 using identical parameter settings for edge detection (see figure 3(b)). We adopted the Canny edge detection algorithm using a Gaussian smoothing kernel with sigma of 1.5 mm (resolution of the images in transversal plane is 0.8x0.8 mm²). To speed up processing, we only focused on the 10 central slices instead of the entire volume, based on the observation that most noticeable motion occurred in these slices and thus, the motion captured here revealed the extent of motion artefact of the entire difference image.
Motion of corresponding breast tissue structures appearing between t0 and t1 results in deformation of their boundaries, thus the deformation of edges can be used to quantify motion strength. In practice, both linear and non-linear deformed edges were demonstrated. Therefore, a simple distance measure between the two sets of detected edges was not suitable to reflect the motion extent, since it was quite difficult to align the two edge sets and find accurate edge pairs for distance calculation (see figure 3(c)).

To cope with the alignment problem, we employed a fast and non-rigid registration method [2] with a volume preservation constraint to register t1 (moving image) onto t0 (fixed image). The obtained deformation field had the same resolution as t0 (see figure 3(d)&(e)). A 3D deformation vector was assigned to each single voxel in t0 representing the direction and magnitude (strength) of the motion occurred in that voxel. The magnitude of the deformation vectors of all the edge voxels detected in t0 was saved to a list, which encoded the motion strength along edges (see figure 3(f)). Then, a histogram was generated from the magnitude list and a set of features was extracted from the histogram as follows:

- Mean
- Standard deviation
- Peak
- Maximum
- Q25: quantile corresponding to of 25 % of the distribution
- Q50: quantile corresponding to of 50 % of the distribution
- Q75: quantile corresponding to of 75 % of the distribution
- Q90: quantile corresponding to of 90 % of the distribution
- Average on top range: the average of magnitudes in range [Q50, Max]

Finally, these features were used for training a Random Forest (RF) classifier [3], which served as a decision maker to predict whether a test case had moderate/severe motion artefact or not. In figures 3 and 4, the motion detection results are shown for the two cases we previously demonstrated in figure 2 (a) with motion and (b) without motion. The average motion strength measured for the case with a motion artefact shown in figure 3 was 1.21 mm, which means on average the edge voxels were deformed by 1.21 mm, leading our classifier to make a positive decision. For the case without a motion artefact, the average motion strength shown in figure 4 was 0.21 mm, resulting in a negative decision by our classifier. In both figures, the areas with strong motion are marked by red squares, where one can observe larger deformation magnitude in the positive case.
(see figure 3 (e)), while for the negative case, the deformation magnitude is much lower (see figure 4 (e)).
Images for this section:

Fig. 1: Incidence and degree of motion artefact in the considered data set as rated by an expert radiologist.

© mediri GmbH - Heidelberg/DE
Fig. 2: Maximum intensity projection images of the difference image $t_1 - t_0$ of the same woman. (a) represents a case without motion and (b) with moderate motion in both breasts.

© mediri GmbH - Heidelberg/DE

Fig. 3: Illustration of motion quantification for the positive case shown in figure 1(a). Its average motion strength is 1.21 mm. Detection scheme: (a) detected edges in $t_0$ (red contours); (b) detected edges in $t_1$ (green contours); (c) detected edges in $t_1$ (green) overlaid with edges in $t_0$ (red); (d) visualization of deformation vectors showing correspondence between edges in $t_0$ and $t_1$; (e) magnified view of the deformation vectors in the red square in (d); (f) colour map of deformation magnitude (displacement in mm) calculated for edge voxels in $t_0$, red correlates with strong motion.

© Fraunhofer MEVIS, Fraunhofer MEVIS - Bremen/DE
Fig. 4: Illustration of motion quantification for the negative case shown in figure 1(b). Its average motion strength is 0.21 mm. Detection scheme: (a) detected edges in t0 (red contours); (b) detected edges in t1 (green contours); (c) detected edges in t1 (green) overlaid with edges in t0 (red), yellow indicates perfect alignment of both; (d) visualization of deformation vector showing correspondence between edges in t0 and t1; (e) magnified view of the deformation vectors in the red square in (d); (f) colour map of deformation magnitude (displacement in mm) calculated for edge voxels in t0, blue correlates with weak motion.

© Fraunhofer MEVIS, Fraunhofer MEVIS - Bremen/DE
Results

To evaluate the performance of the motion classifier, 10-fold cross validation at patient level was applied to objectively report prediction accuracy. Test results presented as receiver operating characteristic (ROC) curve and statistical measures are depicted in figure 5. By selecting 0.5 as the classification threshold, it was shown that the classifier obtained a high specificity of 0.965, which means 391 out of 405 cases were classified correctly as showing no-motion. A sensitivity of 0.535 (46 out of 86 moderate and severe cases were detected correctly) with low false positive rate (0.035) was achieved. Moreover, if we intended to further decrease the false positive (FP) rate, elevating the classification threshold to 0.67 would achieve a fairly good (low) FP number as 6, while keeping true positive (TP) number as high as 36. The overall area under the ROC curve was 0.834.

Representative FP and false negative (FN) cases are demonstrated in figures 6 and 7, respectively. The FP case shown in figure 6 was rated by the radiologist as 2 (mild motion), and the mean motion strength evaluated by the method was 2.76 mm. The reason of inconsistence between expert annotation and method prediction might be that the criteria of judging motion artefact is different. Despite of a relatively strong motion in the image, the ultimate judgement made by the radiologist is to investigate whether the motion can impede subsequent diagnosis. The same reason might account for the inverse relation between manual annotation and automatic classification. The FN case shown in figure 7 was rated as 3 (moderate motion) by the radiologist, but the mean motion strength evaluated by the algorithm was only 0.5 mm. The inconsistence might be caused by the fact that the area, where even mild motion was detected, is quite important for diagnosis, which leads to a higher rating by human observer. Further improvements might face this challenge.
**Fig. 5:** Classification accuracy obtained by 10-fold cross validation: ROC curve (left); classification accuracy with classification threshold of 0.51 (middle text block) and 0.67 (right text block).

© Fraunhofer MEVIS, Fraunhofer MEVIS - Bremen/DE

**Fig. 6:** A FP example: (a) detected edges in t0 (red contours); (b) detected edges in t1 (green contours); (c) detected edges in t1 (green) superimposed with edges in t0 (red), and yellow represents completed overlaid contours; (d) colour map of deformation magnitude (deviation in mm) calculated for edge voxels in t0, and red correlates with strong motion.

© Fraunhofer MEVIS, Fraunhofer MEVIS - Bremen/DE
Fig. 7: A FN example: (a) detected edges in t0 (red contours); (b) detected edges in t1 (green contours); (c) detected edges in t1 (green) superimposed with edges in t0 (red), and yellow represents completed overlaid contours; (d) colour map of deformation magnitude (deviation in mm) calculated for edge voxels in t0, and blue correlates with weak motion.

© Fraunhofer MEVIS, Fraunhofer MEVIS - Bremen/DE
Conclusion

In this work, an automated approach for motion detection in DCE breast MR images was presented. Based on the manually annotated data set, the algorithm showed to extract meaningful features that allow a robust automatic classification. Such a system can be used to trigger an alert after acquisition to warn the technician that the study quality is too low. This would allow re-scanning the women shortly after the first scan, thus preventing unnecessary call-backs. In the future, more annotations will be acquired and included so that our performance evaluations will become more objective. Apart from that, a prototype version of this software will be tested in clinical environment to reveal the strengths and weaknesses of the proposed system and, thus, trigger further development in automated image quality assurance.
References

