Physics and Imaging in Radiation Oncology 1 (2017) 14–20

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.elsevier.com/locate/phro

Original Research Article

Prostate fiducial marker detection with the use of multi-parametric magnetic resonance imaging

Catarina Dinis Fernandes [a], Cuong V. Dinh [a], Marcel J. Steggerda [a], Leon C. ter Beek [b], Milena Smolic [a], Laurens D. van Buuren [a], Floris J. Pos [a], Uulke A. van der Heide [a,*]

[a] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[b] Department of Radiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 29 November 2016
Received in revised form 28 February 2017
Accepted 28 February 2017

Keywords:
Prostate cancer
External beam radiotherapy
Fiducial markers
Multi-parametric MRI

ABSTRACT

Background and purpose: The introduction of a magnetic resonance (MR)-only workflow in radiotherapy requires that fiducial markers, used for position verification, can be detected on MR images. Here we evaluate a model for marker detection in prostate cancer patients by combining information from our hospital standard multi-parametric (mp-) MRI protocol (T1-weighted – T1w, T2-weighted – T2w, B_0) with dedicated sequences (balanced steady-state free precession sequence – bTFE, susceptibility weighted imaging – SWI).

Materials and methods: Thirty two patients scheduled for external-beam radiotherapy received a mp-MRI and computed-tomography; the latter was used as ground truth location of the markers. A logistic regression model was implemented for marker detection by combining features from all imaging sequences. The performance of the individual and combined sequences was assessed by determining true and false positive detections.

Results: The combination of different sequences (mp-MRI) resulted in a better performance than the best imaging sequence alone (bTFE). Combining mp-MRI + bTFE resulted in good accuracy and a true positive detection rate of 0.94.

Conclusions: The standard mp-MRI provides valuable information to detect fiducial markers. The combination of different sequences outperforms the use of a single dedicated sequence. We recommend the addition of a bTFE to the standard mp-MRI protocol to improve fiducial marker detection.

1. Introduction

Prostate cancer patients scheduled to receive external-beam radiotherapy (EBRT) are commonly implanted with fiducial markers used for treatment position verification. Currently, the markers are identified using computed-tomography (CT) whereas structures are delineated on magnetic resonance imaging (MRI), requiring registration of MRI to CT. The registration procedure is associated with errors which will propagate throughout the whole treatment chain. The introduction of a magnetic resonance (MR)-only workflow is a possible way to avoid this. However, a precondition is the ability to identify the markers on MRI.

The most commonly used markers are made of gold, usually presenting as a local signal void on MR images. Several studies have investigated the accuracy of specifically optimized MRI sequences for the detection of implantable metallic structures such as fiducial markers. Some suggest the use of optimized spin-echo MRI sequences, improving void visibility [1], or gradient-echo sequences [2,3] sensitive to $T2^*$ decay, using the susceptibility effects to get better marker depiction [4], as well as by combining it with contrast agents [5]. Others visually combine different MRI sequences [6,7], while some groups have developed sequences that allow for positive contrast at the marker [8–11].

Typically, patients receive a multi-parametric (mp-) MRI used for radiotherapy (RT) target delineation as part of their standard clinical treatment. In our hospital this protocol includes anatomical (T1- and T2- weighted) and functional (diffusion-weighted (DWI) and dynamic contrast enhanced (DCE)) sequences as well as a B_0 map acquired for the post-processing of DWI-MRI. Sequence specific parameters result in distinctive marker voids between images, increasing the complexity of reproducible and accurate manual marker localization. We propose the development of a marker

* Corresponding author at: Department of Radiation Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
E-mail address: u.vd.heide@nki.nl (U.A. van der Heide).

http://dx.doi.org/10.1016/j.phro.2017.02.001

detection model using a machine learning framework incorporating information from the clinical mp-MRI protocol. To investigate the performance of some of the most often used dedicated sequences we have incorporated a balanced steady-state free precession sequence (balanced Turbo-Field Echo or bTFE) [2,3] and a susceptibility weighted imaging (SWI) sequence in the scanning protocol. Our goal is to investigate the performance of a model based on the standard mp-MRI protocol and compare it to models based on a single, dedicated sequence.

2. Materials and methods

2.1. Dataset

Thirty two consecutive patients with biopsy proven prostate cancer, scheduled to receive external-beam radiotherapy (EBRT) were treated between October 2015 and January 2016. Up to three gold fiducial markers (RT Cast, 1.0 × 3 mm) were implanted per patient. Approximately 1 week after implantation, and on the same day, patients underwent a mp-MRI and a CT scan used for RT planning purposes.

The MRI data was acquired using a 3T Achieva dStream MRI scanner (Philips Healthcare, Best, The Netherlands) with the use of an anterior and posterior phased array coil. The mp-MRI protocol included transversal T1-weighted (T1w) T1 THRIVE and T2-weighted (T2w) T2 turbo spin-echo (TSE) images, B_0 map, and the dedicated sequences bTFE and SWI (details of the protocol can be found in the Supplementary Material). Coronal and sagittal T2w scans as well as functional sequences were also acquired as part of the protocol but were not used for further analysis. The CT scans were acquired using a 24-slice CT scanner (Somatom-Sensation-Open, Siemens), with image resolution $0.9 \times 0.9 \times 3\ mm^3$ (x,y,z).

A conservative threshold of 1900 Hounsfield units, resilient to possible streaking artefacts, was used to segment the markers on the CT images. The ground truth (GT) marker coordinates were automatically calculated as the center of mass of the clusters resulting from the segmented CT image. All MR images were acquired within the same exam session. To account for possible within-session motion all images were rigidly registered to the T1w scan with the use of a clipbox around the prostate and periprostatic region [12]. Subsequently the T1w scan was registered to the CT using a similarly defined clipbox; the transformation was then applied to the other MRI sequences. This was achieved by means of rigid-registration (mutual information). Registration accuracy was checked by visual inspection of the prostate boundaries and marker overlap and by assessing the registration error comparing the GT coordinates with a manually defined center of mass of the marker artefact in the bTFE sequence (CM_{mbTFE}). The difference between GT and CM_{mbTFE} coordinates was used as a measure for the target registration error (TRE) between CT and MR.

The prostate gland was delineated using the T2w sequence. The ROI used for model analysis was defined as a 4 pixels (3.6 mm) expansion of the prostate delineation (Fig. 1A–F). Within the ROI the signal intensities (SI) for each image (T1w, T2w, B_0 magnitude, bTFE and SWI images) were normalized to values between 0 and 1 using Min-Max scaling, ensuring the values were comparable between patients. All images were resampled to the grid of voxel size $1 \times 1 \times 1\ mm^3$.

2.2. MRI features and feature extraction

The markers present as signal voids in most MR images and their apparent position depends on their shape and orientation relative to the magnetic field. With exception of T2w, the artefact they induce is usually much larger than the actual physical size and often exhibits a blob-like shape in T1w, bTFE, B_0 magnitude and SWI images – Fig. 1A–F. This same artefact also expands through slices in a cylindrical line-like pattern. In this study we have included multi-scale blobness [13] and line filtering [14,15] of the prostate region as well as intensity based local statistical features such as signal intensity in the voxel *i* and the mean, median, minimum, maximum and standard deviation values from a local window ($3 \times 3 \times 9$ voxels in x,y,z, directions, where z is along the B_0) surrounding this voxel.

Each voxel was described by 42 features in total (6 features from T2w and 9 features from each T1w, B_0 magnitude, bTFE and SWI). Table 1 summarizes all extracted features with detailed description. Single sequence models include all features extracted for that specific imaging sequence.

Performance was investigated for models created using four sequence combination options: 1) mp-MRI only (T1w, T2w and B_0 magnitude); 2) mp-MRI + bTFE; 3) mp-MRI + SWI and 4) mp-MRI + bTFE + SWI. Models for each individual imaging sequence were also created in order to ascertain their individual value.

2.3. Model creation

Ground truth maps (GT_{maps}) were created based on the GT coordinates, incorporating the prostate and margin delineations; these were used to train and validate the model results.

Data analysis was performed using MATLAB R2015a (MathWorks, Natick, MA, USA) and the PRTools toolbox [16]. PRTools is a pattern recognition toolbox that among others supports feature extraction, the use of classifiers and evaluation methods. First, we determined from the GT_{maps} the value of M, representing the average number of voxels labeled as markers. Then, for each of the individual sequences and for the four sequence combinations a model was trained by fitting a logistic regression to the labeled (marker or non-marker) training data. The logistic regression model, applied to an unseen dataset, results in a map of the probability for each voxel of being a marker. The M highest probability voxels were grouped into clusters. We defined the probability of each cluster candidate as the highest probability value of all voxels within it. Prior information regarding the number of markers implanted per patient was used to further post-process the results by selecting the $n(j)$ highest probability clusters, where $n(j)$ is the number of markers implanted in patient *j*. Other than the logistic regression, the process did not involve any optimization of parameters or feature selection.

Each model was evaluated using a leave-one-dataset-out cross-validation method to make sure that the training set – used to build the model – and test set are independent from each other. The different (individual and combination) models were trained and the rate of true positives (TP) and false positives (FP) was used to assess the performance of the model for each patient. Clusters were classified as a FP if the distance between the GT coordinates and the center of mass of the clusters identified by the model (CM_{Model}) was larger than 3 mm in plane or through slice. Model accuracy was assessed for each marker by calculating the distance between the GT coordinates and the CM_{Model} for all TP.

2.4. Statistics

A Shapiro-Wilk test was used to test the normality of the TRE between MRI and CT coordinates. For each model we determined the numbers of TP and FP per patient and combined them into distributions over the population. A non-parametric Wilcoxon signed-rank test was used to evaluate whether the difference in distributions between models was statistically significant