CrossMark

ORIGINAL ARTICLE

In vivo validation of spatio-temporal liver motion prediction from motion tracked on MR thermometry images

C. Tanner[1] · Y. Zur[2] · K. French[3] · G. Samei[1] · J. Strehlow[4] · G. Sat[2] ·
H. McLeod[3] · G. Houston[5] · S. Kozerke[6] · G. Székely[1] · A. Melzer[3] · T. Preusser[4,7]

Abstract

Purpose Magnetic resonance-guided focused ultrasound (MRgFUS) of the liver during free-breathing requires spatio-temporal prediction of the liver motion from partial motion observations. The study purpose is to evaluate the prediction accuracy for a realistic MRgFUS therapy scenario, namely for human in vivo data, tracking based on MR images routinely acquired during MRgFUS and in vivo deformations caused by the FUS probe.

Methods In vivo validation of the motion model was based on a 3D breath-hold image and an interleaved acquisition of two MR slices. Prediction accuracy was determined with respect to manually annotated landmarks. A statistical population liver motion model was used for predicting the liver motion for not tracked regions. This model was individualized by mapping it to end-exhale 3D breath-hold images. Spatial correspondence between tracking and model positions was established by affine 3D-to-2D image registration. For spatio-temporal prediction, MR tracking results were temporally extrapolated.

Results Performance was evaluated for 10 volunteers, of which 5 had a dummy FUS probe put on their abdomen. MR tracking had a mean (95 %) accuracy of 1.1 (2.4) mm. The motion of the liver on the evaluation MR slice was spatio-temporally predicted with an accuracy of 1.9 (4.4) mm for a latency of 216 ms. A simple translation model performed similarly (2.1 (4.8) mm) as the two MR slices were relatively close (mean 38 mm). Temporal prediction was important (10 % error reduction), while registration effects could only partially be assessed and showed no benefits. On average, motion magnitude, motion amplitude and breathing frequency increased by 24, 16 and 8 %, respectively, for the cases with FUS probe placement. This motion increase could be reduced by the spatio-temporal prediction.

Conclusion The study shows that tracking liver vessels on MR images, which are also used for MR thermometry, is a viable approach.

**Keywords** Focused ultrasound · Respiration · Motion prediction · Tracking

✉ C. Tanner
tanner@vision.ee.ethz.ch

1 Computer Vision Laboratory, ETH Zürich, Zurich, Switzerland

2 GE Medical Systems Israel, Haifa, Israel

3 Institute for Medical Science and Technology, University of Dundee, Dundee, Scotland, UK

4 Institute for Medical Image Computing, Fraunhofer MEVIS, Bremen, Germany

5 School of Medicine, University of Dundee, Dundee, Scotland, UK

6 Institute for Biomedical Engineering, University and ETH Zürich, Zurich, Switzerland

7 Jacobs University Bremen, Bremen, Germany

Introduction

To enable focused ultrasound (FUS) therapy of the liver during free-breathing, it is important to have accurate information about the position of the structures of interest during therapy. Purely tracking the tumor is not sufficient for FUS, where absorbing and reflecting structures (e.g., bones, gas) can cause thermal injury to neighboring tissue [18,27,32]. Observation of the motion for these structures requires real-time 4D image acquisition and processing, which is currently impossible. Hence prior knowledge about the expected

🖄 Springer

Fig. 1 Main components for respiratory organ motion prediction based on population model, 3D image before and partial observations from tracking during therapy, see "Methods" section
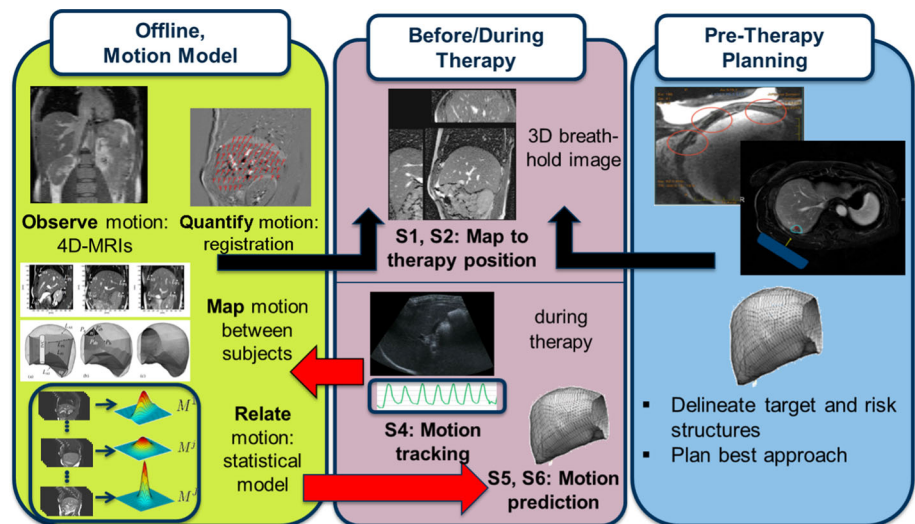
motion is required to complement the partial observations acquired during therapy.

Subject-specific [3,7,11,13,14,16,36] and population-based [2,6,9,12,17,19,20,34] respiratory motion models have been proposed for this purpose [15,29]. Population models are built by gathering 4D data from a number of subjects. Subsequently, the common structures within different subjects are registered and their motion correlates are learned. Figure 1 illustrates the usage of such a population model. The model is individualized by spatially mapping it to the subject using a 3D image. During therapy, the individualized model predicts from partial motion observations (from tracking structures on an MR or US image slice) the motion of unobserved regions. In contrast, a subject-specific model approach adapts to a subject by using 4D data of this subject. For 4D-CTs only very few breathing cycles are observed to avoid excessive radiation. More information can be gathered with 4D-MRI, allowing the observation of variations in breathing states. Yet long 4D-MRIs (>20 min), for observing also non-periodic long-term motion phenomenon (drift) [34], are impractical. Hence subject-specific models based on short 4D-MRIs need combining with a population drift model or regularly updated with 3D observations.

Of these various motion models [15,29], only one model was realistically in vivo validated for abdominal organs during respiration [20] and this was based on US tracking. However such an MR/US approach requires an MR-compatible US device, a lengthy setup and synchronization of the US imaging and FUS transmissions. Motion has previously been tracked on MR slices acquired for MR thermometry [21–23,27,28] or on other types of MR slices [5,10]. Yet experiments included mostly periodic, simple motion of phantoms, with the exception of [5,10,21,22]. Few of these studies, which use MR tracking, included temporal predictions [5,27] and none spatio-temporal prediction. This study

evaluates for the first time the in vivo spatio-temporal prediction accuracy of a motion model being driven by tracking structures on the thermometry magnitude MRIs during free-breathing. Furthermore, the sensitivity of the predictions to deformations induced by a FUS probe is analyzed. This is an important step toward translating MRgFUS in the liver into the clinic, which is the aim of the TRANS-FUSIMO project. It builds on the integrated model-based software developed during FUSIMO [26]. The motion model is based on [25,31]. Initial results of this validation, for a spatio-temporal model built from 12 subjects, were included in [26] without providing much details.

Material

Images for validation study

For the validation study, 14 volunteers were scanned in Dundee. The MR sequence was first optimized on 4 volunteers and then fixed for the remaining 10 volunteers (V1–V10). Five of these were scanned with a dummy FUS probe in place. All volunteers were imaged in supine position with the GE Signa 1.5T HDX Echospeed MR scanner, using a 8-channel cardiac array.

– *3D breath-hold MRIs* were used for mapping the population motion model to the subject's therapy position. A 3D FIESTA sequence with 48 sagittal slices of 4 mm thickness was used. Its scan parameters were TE = 1.3 ms, TR = 3.3 ms, flip angle 80°, FOV 40 cm and 256×256 acquisition matrix, see Fig. 2a, c. As the slice acquisition time was 1.75 s, the acquisition was divided into 4 parts of 12 slices to reduce breath-holding time to 21 s.
– *Dynamic two-slice MR sequences* were used for motion tracking and for capturing motion observations for eval-
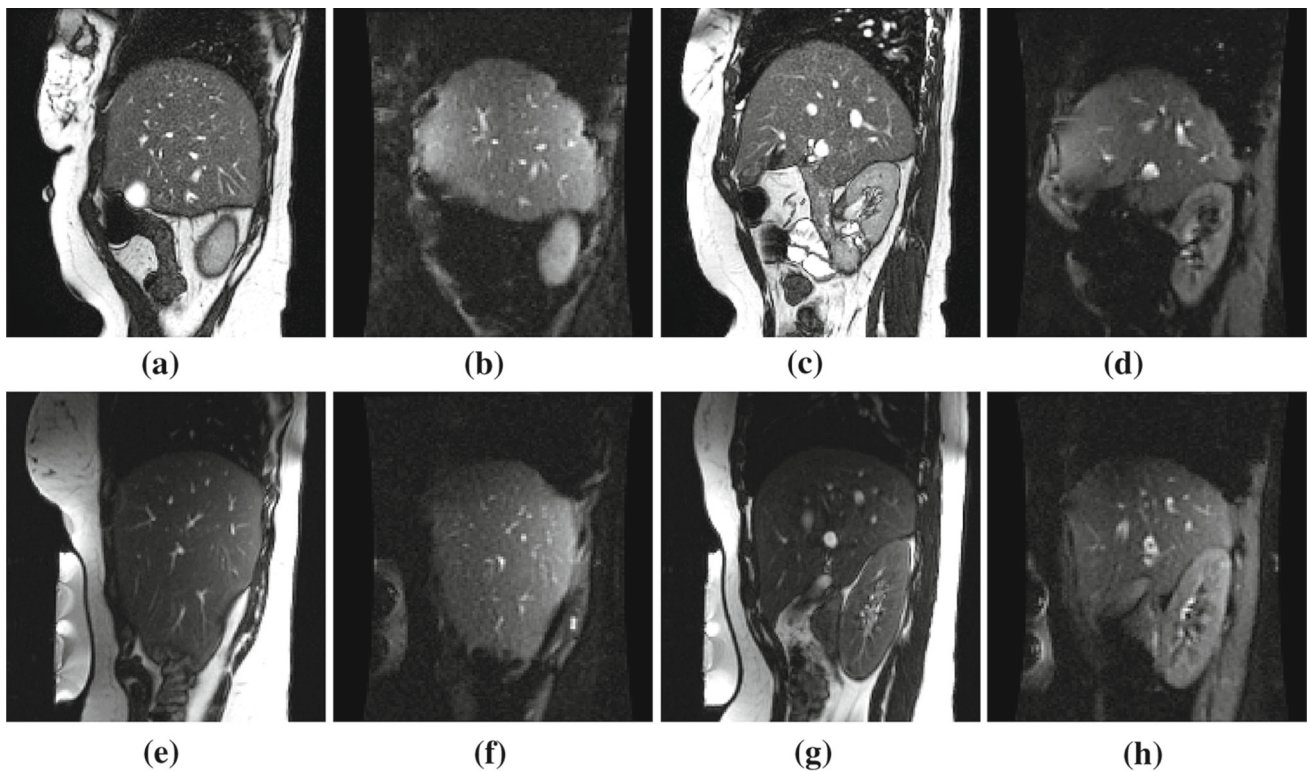
Fig. 2 Example images (**a–d**) without and (**e–h**) with dummy FUS probe. **a**, **c**, **e**, **g** Zoomed in breath-hold slice closest to (**b**, **d**, **f**, **h**) dynamic slice. (**a**, **b**) V3, lateral, (**c**, **d**) V3, medial, (**e**, **f**) V8, lateral, (**g**, **h**) V8, medial

uation. These were based on echo-planar imaging (EPI), since the phase EPI images are used for MR thermometry during FUS therapy, while the magnitude EPI images provide anatomical information for vessel tracking. The sequence was tuned for good vessel contrast (see Fig. 2b, d) and minimum acquisition time (72 ms per slice). Its scan parameters were TE=23.4 ms, TR=144 ms, flip angle 50°, FOV 260 mm and 128×90 acquisition matrix. The 2 sagittal slice locations were planned on a coronal image of the liver, to ensure a reasonable gap between the two slices (range [33.2,44.1] mm) for avoiding interferences and testing spatial prediction ability. The liver vessel appearance on the selected slices was visually inspected. The sequence iterated between the two slices and each slice location was repeated 300 times, giving a scan time of 46 s. This was repeated 6 times, resulting in a total scan time of 4.6 min. The volunteer was told to breathe normally throughout the scan.

– *Images with dummy FUS probe* To study the influence of the FUS probe positioning on the liver motion during respiration, MRIs were obtained for five volunteers with a dummy probe in place using the same MR protocols, see Fig. 2e–h. Initially the dummy probe was too uncomfortable for the volunteers to tolerate. Therefore the probe was fitted with a new membrane, which allowed for better filling and thus provided almost a cushion effect. Better

support padding was also utilized. The volunteers were able to tolerate the probe, but still found it particularly uncomfortable.

Images for motion model creation

For the motion model, 4D-MRIs were acquired in Zurich for 16 healthy volunteers. An interleaved sagittal 2D sequence was used, where slices covering the liver were alternated with a navigator slice placed at the center of the right liver lobe [33]. After capturing many breathing cycles, the slices were retrospectively sorted based on the liver position on the navigator to form 3D volumes. MRIs were acquired on an 1.5T Philips Achieva whole body MR system using a balanced steady-state free precession sequence, SENSE factor 1.7 and halfscan (flip angle $= 70°$, TR $= 3.1$ ms, TE $= 1.5$ ms). The images had a spatial resolution of $1.33 \times 1.33 \times 4-5$ mm^3 and a temporal resolution of 2.6–2.8 Hz. For the first 12 volunteers, the right liver lobe was imaged and a 4-channel cardiac array coil was used. For the remaining 4 volunteers, the whole liver was imaged with the same MR protocol and a 32-channel cardiac array coil.

The validation images and 25 % of the 4D-MRIs were not used in [2,19,20,33,34].

Methods

The motion prediction concept is illustrated in Fig. 1. It consists of 6 steps as described next, of which steps S4 to S6 are done repeatedly during therapy.

S1 *Individualization of motion model* The off-line-created population motion model [25,31] (for completeness described in "Appendix") is individualized by mapping it to the subject's liver captured on a 3D breath-hold MRI in therapy position, see "Inter-subject correspondences" section.

S2 *Registration of 3D breath-hold to 2D dynamic slices* To relate the MR observations to the motion model, spatial correspondence between the 3D FIESTA breath-hold image and the EPI reference slices needs to be established. The 3D images were registered to the slices by using 3D affine transformations and minimizing the difference in normalized gradient magnitude within the liver to cope with differences in image appearance (Fig. 2). This similarity measure was favoured over mutual information to reduce runtime.

S3 *Automatic detection of liver vessels* Bright blood vessels, which are used as landmarks, were automatically detected on the reference slices with an optimized algorithm based on their shape, size and brightness. Usually 10–15 landmarks are detected.

S4 *MR tracking* The bright blood vessels in the liver were tracked with subpixel resolution on the EPI images using a tracking method inspired by [35]. After the detection of the vessels in step S3, the location of each landmark over time is determined using an autocorrelation algorithm. Due to pulsatile flow in the large arteries, not all vessels can always be tracked reliably. An algorithm detects the reliability of each landmark and reports it by a binary flag ("valid", "invalid"). The motion of the "invalid" landmarks was set to the mean motion of the other landmarks.

S5 *Temporal prediction* To compensate for latency Δ in the therapy system, motion needs to be predicted for a future time point. This was done by first temporally extrapolating the tracking results ($\mathbf{s}_t \rightarrow \mathbf{s}_{t+\Delta}$) and then using $\mathbf{s}_{t+\Delta}$ as input for the spatial prediction model. We evaluated the temporal prediction methods from [30] (adaptive linear (LIN), second-order (POLY2) prediction, support vector regression (SVR), kernel density estimation (KDE), median of these 4 methods (MED)) for an input sampling time (150 ms) and latency Δ (75 ms, 150 ms) similar to the EPI image sequence on an independent dataset. The best-performing method was then selected for the motion model validation.

S6 *Spatio-temporal prediction* The extrapolated tracking results $\mathbf{s}_{t+\Delta}$ are then used as partial observations (sur-

rogates) to predict the liver motion $\Delta \mathbf{p}_{t+\Delta}$ via the population liver motion model [25,31] using Eqs. (3, 4).

Validation strategy

The key for the motion model validation is the acquisition of an interleaved two-slice EPI sequence, such that one slice can be used for MR tracking and the other slice for evaluation of the prediction accuracy.

Liver vessel positions $\mathbf{t}_{j,0}$, $j = 1, \ldots, J$ and $\mathbf{v}_{k,0}$, $k = 1, \ldots, K$ were automatically detected (during step S3) in the reference image of the tracking and validation EPI sequence, respectively. Tracking of $\mathbf{t}_{j,0}$ (S4) provided position $\mathbf{t}_{j,t}$ at time t, and spatio-temporal prediction (S6,7) estimated $\mathbf{v}_{k,t+\Delta}$.

All vessel locations $\mathbf{v}_{k,0}$ were annotated on a randomly selected subset of 5 % of all time frames by one observer (C. T.). Unreliable annotations were marked as "invalid". Landmarks were annotated with subpixel resolution without accessing tracking results and before availability of prediction results. The mean (95 %) intra-observer annotation accuracy, determined by redoing 20 % of the previous annotations, was 0.6 (2.1) mm.

The prediction error per landmark and image frame was quantified by the Euclidean distance $E_{k,t+\Delta} = ||\mathbf{v}_{k,t+\Delta} - \mathbf{g}_{k,t+\Delta}||$ for "valid" manual annotations $\mathbf{g}_{k,t+\Delta}$. The error statistics was summarized by first determining the mean and the 95 % of $E_{k,t+\Delta}$ for all annotated image frames $t + \Delta$ of vessel k (denoted as $\bar{E}_k$, E_k^{95} resp.) and then by calculating the mean of $\bar{E}_k$ and E_k^{95} over all vessels $k = 1, \ldots, K$ and volunteers. The process was repeated after reversing the role of the tracking and validation slice.

Results

Individualization of motion model

The liver in the breath-hold image was manually segmented, anatomical landmarks were selected, and the liver uniformly meshed to achieve inter-subject correspondences, see "Inter-subject correspondences" section. Some difficulties in following structures across slices were experienced for few images when insufficient breath-hold repeatability caused misalignments.

Registration of 3D breath-hold to 2D dynamic image

Figure 2 shows the closest slice from the 3D FIESTA breath-hold image to the EPI dynamic slice. Large differences in image appearance can be appreciated. The effect of reg-
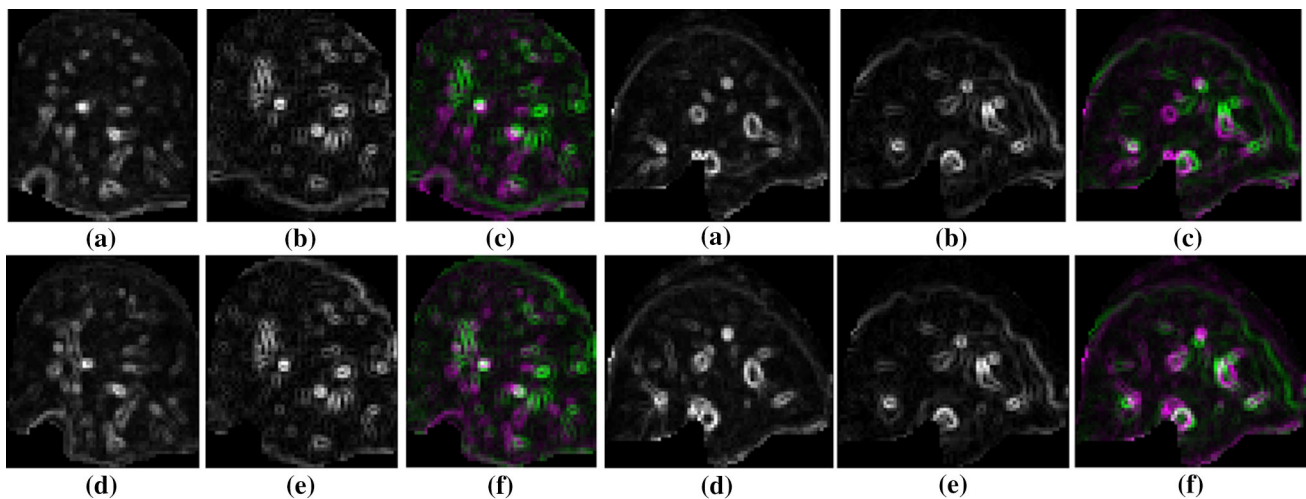
Fig. 3 Example of 3D–2D registration result for (*left*) lateral and (*right*) medial EPI slice of V3. **a**, **d** Gradient magnitude (GM) of slice from 3D image **a** before and **d** after registration. **b**, **e** GM of EPI slice within liver region. **c**, **f** Overlay of **a**, **b** and **d**, **e**
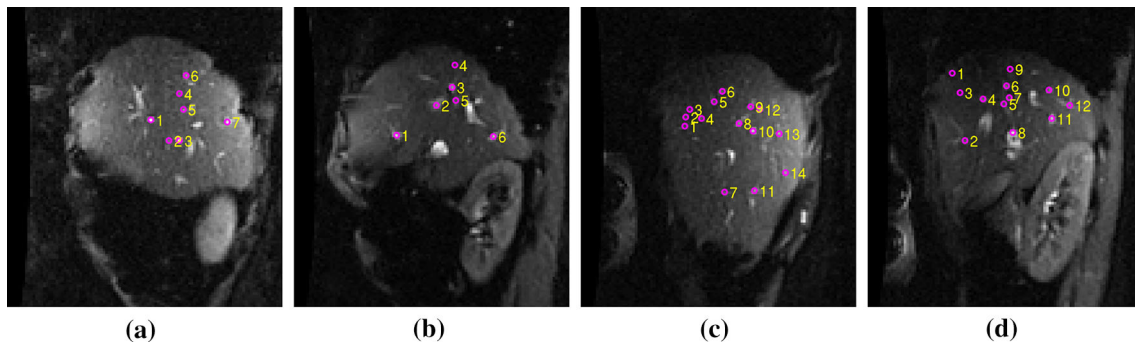

Fig. 4 Detected vessels (*magenta*) on EPI MR reference slice. **a** V3, lateral. **b** V3, medial. **c** V8, lateral. **d** V8, medial

istering the breath-hold image to the dynamic slice when employing a 3D affine transformation can be seen in Fig. 3. An improved alignment of the image feature due to registration can be observed, while some misalignments are still visible. After registration, the reference and tracked vessel positions are spatially transferred to the breath-hold image and hence mapped to the motion model.

Automatic detection of liver vessels

Figure 4 shows example reference slices and the automatically detected vessels. Almost all of these locations are clearly located at vessel cross sections.

MR tracking

Example MR tracking results are shown in Fig. 5. It can be observed that the motion is larger in SI (mean standard deviation (SD): 3.2 mm) than in AP direction (mean SD: 2.0 mm), and that the breathing pattern can be irregular. The motion

patterns of the various vessels are highly correlated per slice ($R > 0.94$ lateral and $R > 0.82$ medial) and across slices.

The manual annotations used for the motion model validation were also used for assessing the MR tracking performance. Overall the mean (95 %) motion is reduced from 4.9 mm (14.2 mm) to 1.1 mm (2.4 mm) by MR tracking. Similar tracking performance was achieved for the two EPI slices [lateral: 1.1 mm (2.3 mm), medial: 1.2 mm (2.5 mm)] and for the volunteers with and w/o the dummy FUS probe [1.1 mm (2.4 mm) for both]. More vessels were marked as "invalid" for the medial than the lateral slice (35 vs. 26 %) and for images with dummy FUS probe than without (33 vs. 26 %). The variation in respiratory motion for the volunteers can be seen in Fig. 6a. MR tracking reduced this to a similar mean accuracy (range 0.9–1.4 mm), see Fig. 6b.

Temporal prediction

The temporal prediction results for 25 resampled US liver vessel motion traces [30] are listed in Table 1 and compared to assuming that no motion has occurred during latency Δ
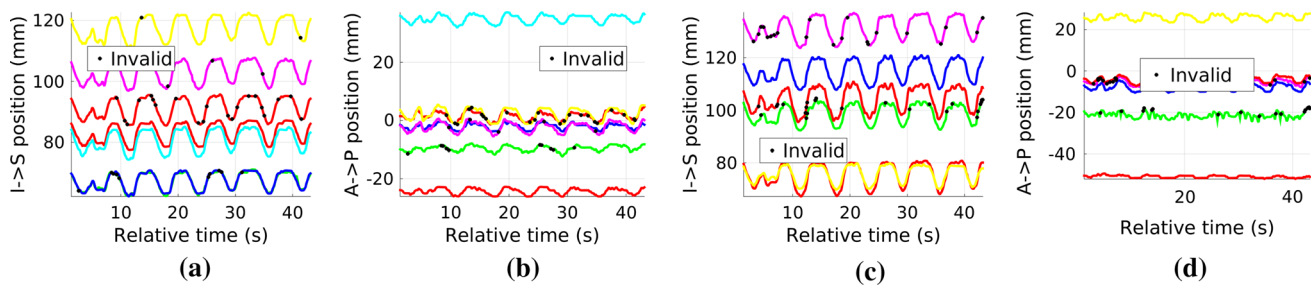
Fig. 5 MR tracking results showing superior–inferior (SI) and anterior–posterior (AP) position of the automatically detected vessels over time. *Black dots* mark low confidence results. **a** V3, lateral, SI. **b** V3, lateral, AP. **c** V3, medial, SI. **d** V3, medial, AP
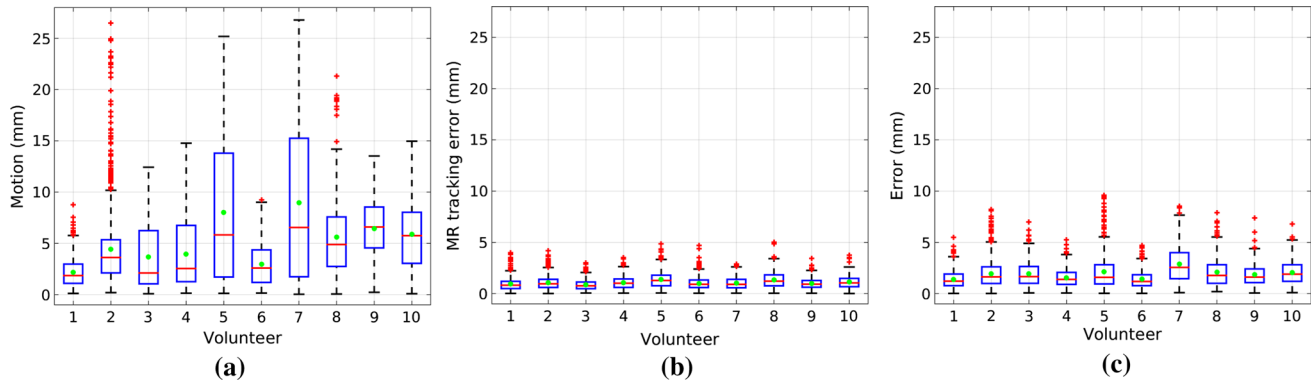

Fig. 6 *Boxplots* showing distribution of initial motion, MR tracking error, spatio-temporal prediction error for 216-ms latency. The mean values are marked by *green stars*. **a** Motion. **b** MR tracking. **c** Spatio-temporal prediction

Table 1 Mean RMS error (in mm) of temporal prediction for sampling rate of 150 ms and latency $\Delta \in \{75, 150\}$ ms

Δ	Prediction method →	NON	LIN	POLY2	SVR	KDE	MED
75	Optimized parameters	1.4	**0.3**	**0.3**	1.0	1.1	0.5
ms	Median parameters	1.4	**0.3**	**0.3**	4.0	1.6	0.8
150	Optimized parameters	1.9	**0.6**	**0.6**	1.1	1.1	0.6
ms	Median parameters	1.9	**0.6**	**0.6**	3.4	1.1	0.7

Method parameters were either subject-specific optimized or set to the median population value. The two lowest values per row are marked in bold

(NON). The linear adaptive filter (LIN) performs best for these relatively short latencies. As before, subject-specific optimization of the parameters (over first 10 breathing cycles) is not required, since the median parameters provide a similar performance for the best two methods when applied to the remaining data. Applying the LIN method to the motion traces extracted by MR tracking for a latency of 72 (216) ms resulted in a mean root-mean-square (RMS) error of 0.4 (1.1) mm, while no prediction (NON) results in 0.6 (1.6) mm. Note, LIN could only be applied after sufficient (180) samples were observed.

Spatio-temporal prediction

In accordance with [31], the model parameters were 99 % cumulative PCA energy, $K = 5$ nearest models, history length $O = 300$ and regularization weight $\eta = \sigma_N^2 / K$ where

$\sigma_N = 1$ mm. Each model was trained on $T = 2000$ time steps regularly sampled from up to 850 breathing cycles. The latency Δ was 72 or 216 ms, as two EPI slices were acquired within 144 ms. Prediction performance was assessed with or w/o 3D affine registration, with or w/o temporal prediction and with mean translation or exemplar model.

Table 2 summarizes the prediction accuracy for all sequences for 10 volunteers. Predictions were compared to MR tracking results (311863 samples) or manual annotations (4934 samples). Results are statistically significantly improved with temporal prediction, but similar with and w/o registration. Using the mean translation was on average 4.0 (8.4) % worse for Δ (72) 216 ms. Runtime states the average time for predicting the liver position excluding MR tracking, while including 3D affine registration. The variation in prediction accuracy across volunteers is shown in Fig. 6. A tendency of increased breathing motion (5.7 vs. 4.6 mm)

Table 2 Mean and 95th percentile error (in mm) w.r.t. reference position over time and then averaged for all vessel locations

Δ	Method	w.r.t. MR tracking		w.r.t. manual annotation		RT (ms)	CC1 (%)	CC2 (%)
		Mean	95 %	Mean	95 %	Mean		
n/a	No prediction (motion)	4.9	15.4	5.0	14.9	n/a	n/a	n/a
72	Translation	1.6	4.0	1.9	4.5	**8**	8	48
	Temporal + translation	1.5	3.9	1.8	4.2	**16**	−6	48
	Spatial	1.5	3.8	1.9	4.3	30	17	46
	Temporal + spatial	**1.5**	**3.7**	**1.7**	**4.0**	49	3	48
	Reg. + spatial	1.5	3.7	1.9	4.3	53	21	47
	Reg. + temporal + spatial	**1.5**	**3.6**	**1.7**	**4.0**	65	8	48
216	Translation	2.0	4.9	2.4	5.9	**8**	32	43
	Temporal + translation	1.7	4.2	2.1	4.8	**16**	−9	39
	Spatial	1.9	4.6	2.4	5.7	34	37	43
	Temporal + spatial	**1.6**	**4.0**	**1.9**	**4.4**	49	−1	43
	Reg. + spatial	1.9	4.6	2.4	5.8	50	39	44
	Reg. + temporal + spatial	**1.6**	**3.9**	**1.9**	**4.4**	65	5	44

Last columns show runtime (RT), correlation coefficient between mean error and plate distance (CC1) and between mean error and motion (CC2). The two lowest values per column for each latency Δ are marked in bold. Per latency, the median of the best method (temporal + spatial) was statistically significantly different at the 0.005 level (Wilcoxon signed-rank test) to all other methods apart from 'reg. + temporal + spatial'

was observed with FUS dummy probe placement (V6–V10). Motion prediction reduced this difference (2.0 vs. 1.8 mm).

Summary and discussion

This study shows that the liver motion during free-breathing can be predicted with a mean (95 %) accuracy of 1.9 (4.4 mm) for a latency of 216 ms by the population model based on individualization using a 3D breath-hold image and tracking liver vessel motion on 2D MR thermometry images. This performance is similar compared to the other existing realistic in vivo validation, which is based on US tracking (2.4 mm mean 3D accuracy for $\Delta = 200$ ms) [20]. They also use a population-based statistical motion model, but partial observations from US tracking, a neural network for temporal prediction, a single PCA model for spatial prediction, simultaneous 4D-MR and US acquisition for evaluation, and do not investigate the influence of abdominal deformations.

Using the observed mean translation for spatial prediction performed similar due to the closeness of the MR tracking and validation slice. However it will not extrapolate well to greater distances (increase of 95 % by 1 mm for the whole liver in leave-one-subject experiments similar to [31]) and cannot make the most of additional observations. The errors of the individual system components are clearly not additive, with mean errors for MR tracking (1.1 mm) and temporal prediction (0.9 mm) already adding up to 2.0 mm. Application of a FUS probe increased on average breathing frequency and magnitude probably due to discomfort and the anterior motion restriction.

The study was limited by not capturing out-of-plane motion, which should be small, and by having no ground truth for the image registration. Acquiring an EPI slice during the same breath-hold would avoid this registration. The misalignments of the breath-holds had likely a small impact as displacement fields are smooth and the EPI slices lied to 50 % in the same acquisition block.

Main aspects which should improve the prediction include tracking vessels on both MR slices (no need for a validation slice), individualization based on 3D motion observations from breath-holds [31] and automatic registration by acquiring the tracking reference slice within in the same breath-hold. In conclusion, a spatio-temporal model of the liver motion during free-breathing driven by MR motion observations was in vivo validated and provided an encouraging mean accuracy close to the initial clinical requirement of 2 mm.

Compliance with ethical standards

Research involving human participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments

or comparable ethical standards. This study does not contain patient data.

Informed consent Informed consent was obtained from all individual participants included in the study.

Appendix: Population 4D motion model

Liver registration

An intensity-based non-rigid registration [8,24] was used to quantify the liver motion captured by the 4D-MRIs. The registration parameters were iteratively adjusted such that visual inspection showed misregistrations below one pixel [33]. To cope with the sliding boundaries between the liver and the abdominal wall, the B-spline transformation parameters were optimized to maximize normalized cross-correlation only within the reference liver region.

Inter-subject correspondences

We used the approach from [34] to define inter-subject correspondences. First a fine surface mesh was extracted from the liver segmentation. Then sagittal slices were identified, which include the most lateral liver location (slice number S_1), the inferior liver tip (S_2), the bifurcation of the inferior vena cava (IVC) (S_3) or the portal vein (S_4), with $S_1 < S_2 < S_3 < S_4$. Four mechanically relevant landmarks (dorsal and ventral point where the liver joins the ribcage, and most inferior dorsal and ventral point) were manually identified on slices $S_2 + 1$ to $S_3 - 1$. A point on the liver tip replaced the two inferior landmarks for slices S_1 to S_2. An additional landmark was selected on the IVC for slices S_3 to S_4. The landmarks belonging to one location were connected by a B-spline. A prototype of the right liver lobe (46 triangles) was then mapped to the liver by aligning its four edges with the marked delineations. Finally this coarse mesh was gradually refined by a fixed number of regular subdivisions to fit to the fine surface mesh. The resulting liver meshes consist of $N = 290$ points, where $\mathbf{p}_t^n = [p_{SI_t}^n \; p_{AP_t}^n \; p_{LR_t}^n]$ denotes the position of the nth point at time step t in the superior–inferior (SI), anterior–posterior (AP) and left–right (LR) direction. The mesh position at time step $t \in \{1, 2, \ldots, T\}$ was described by $\mathbf{p}_t = [\mathbf{p}_t^1 \; \ldots \; \mathbf{p}_t^N]^T \in \mathbb{R}^{3N \times 1}$ and its motion by $\Delta \mathbf{p}_t = \mathbf{p}_t - \mathbf{p}_{\text{ref}}$ where $\mathbf{p}_{\text{ref}}$ denotes the position in the reference end-exhale image. Mesh interpolation via Barycentric coordinates was used for defining correspondences for any position in the mesh.

Creation of motion model

We used the robust exemplar model [25,31], which has shown improved performance over a principle component analysis (PCA) model for the whole population. It is based on creating subject-specific PCA models and combining their predictions according to their closeness to the tracking results.

Subject-specific model

Assuming that $\Delta \mathbf{p}_t, t \in 1, 2 \ldots T$, belong to a $3N$-dimensional Gaussian distribution $\Delta \mathbf{p}_t \sim \mathcal{N}(\mu, \Sigma)$, the prediction task is to find the most probable vector $\Delta \hat{\mathbf{p}}_t \in \mathbb{R}^{3N \times 1}$, given a subset of its elements $\mathbf{s}_t \in \mathbb{R}^{S \times 1}$, called surrogates. Decomposing all displacements $\Delta \mathbf{p}_t$ and their mean μ, and covariance matrix Σ into the components relating to surrogates $\mathbf{s}_t$ and to the rest of the points ($\mathbf{r}_t$), we get $\Delta \mathbf{p}_t^T = [\mathbf{s}_t^T \; \mathbf{r}_t^T], \mu^T = [\mu_\mathbf{s}^T \; \mu_\mathbf{r}^T]$, and $\Sigma = \begin{bmatrix} \Sigma_\mathbf{ss} & \Sigma_\mathbf{sr} \\ \Sigma_\mathbf{rs} & \Sigma_\mathbf{rr} \end{bmatrix}$. The conditional distribution $\Delta \mathbf{p}_t | \mathbf{s}_t \sim \mathcal{N}\left(\mu_{\Delta \mathbf{p}_t | \mathbf{s}_t}, \Sigma_{\Delta \mathbf{p} | \mathbf{s}}\right)$ [1], with its mean $\mu_{\Delta \mathbf{p}_t | \mathbf{s}_t} = \mu + \begin{bmatrix} \Sigma_\mathbf{ss} \\ \Sigma_\mathbf{rs} \end{bmatrix} \Sigma_\mathbf{ss}^{-1}(\mathbf{s}_t - \mu_\mathbf{s})$ providing the most probable $\Delta \hat{\mathbf{p}}_t$ given $\mathbf{s}_t$.

PCA was employed for dimensionality reduction to the M eigenvectors associated with the highest eigenvalues $\hat{\lambda}^2$ leading to $\Sigma \approx \hat{\mathbf{E}} \hat{\Lambda} \hat{\mathbf{E}}^T$ where $\hat{\mathbf{E}} \in \mathbb{R}^{3N \times M}$ and $\hat{\Lambda} \in \mathbb{R}^{M \times M}$. Finally by extracting the submatrix of eigenvectors of the surrogates $\mathbf{s}$, i.e., $\hat{\mathbf{E}}_\mathbf{s} \in \mathbb{R}^{L \times M}$, the most probable prediction is given by

$$\Delta \hat{\mathbf{p}}_t = \mu + \hat{\mathbf{E}} \hat{\Lambda} \hat{\mathbf{E}}_\mathbf{s}^T (\hat{\mathbf{E}}_\mathbf{s} \hat{\Lambda} \hat{\mathbf{E}}_\mathbf{s}^T)^{-1} (\mathbf{s}_t - \mu_\mathbf{s}). \tag{1}$$

Model regularization

Regularization of the model [4,19] was used to compensate for tracking errors. Equation (1) provides the most probable PCA coefficients $\mathbf{c}_t$ which minimize $||\mathbf{Q}\mathbf{c}_t - (\mathbf{s}_t - \mu_\mathbf{s})||^2$ where $\mathbf{Q} = \hat{\mathbf{E}}_\mathbf{s} diag(\hat{\lambda}_i)$. Using ridge regression, we want to minimize instead

$$\arg \min_{\mathbf{c}_t} ||\mathbf{Q}\mathbf{c}_t - (\mathbf{s}_t - \mu_\mathbf{s})||^2 + \eta ||\mathbf{c}||^2, \tag{2}$$

where η scales the regularization. Equation (2) can be solved by applying singular value decomposition (SVD) to get $\mathbf{Q} = \mathbf{U}_\mathbf{Q} \mathbf{D}_\mathbf{Q} \mathbf{V}_\mathbf{Q}^T$, with $\mathbf{D}_\mathbf{Q} = diag(d_{\mathbf{Q},i})$ and

$$\mathbf{c}_t = \mathbf{V}_\mathbf{Q} diag\left(\frac{d_{\mathbf{Q},i}}{d_{\mathbf{Q},i}^2 + \eta}\right) \mathbf{U}_\mathbf{Q}^T (\mathbf{s}_t - \mu_\mathbf{s})$$
$$\Delta \hat{\mathbf{p}}_t = \mu + \hat{\mathbf{E}} \, diag(\hat{\lambda}_i) \mathbf{Q}\mathbf{c}_t. \tag{3}$$

Robust exemplar model

To create an exemplar model [25,31] for a population, a subject-specific PCA model M^j was built for each subject j. Then the distance d_t^j between surrogates $\mathbf{s}_t$ and model M^j

was calculated by $d_t^j = (\mu_{\mathbf{S}_t} - \mu_{\mathbf{s}}^j)^T \Sigma_{\mathbf{ss}}^{j\,-1} (\mu_{\mathbf{S}_t} - \mu_{\mathbf{s}}^j)$, where $\mu_{\mathbf{S}_t}$ estimates the mean of the surrogate distribution from the last $O \leq t$ observations $\mathbf{S}_t = [\mathbf{s}_{t-O+1}^T \ldots \mathbf{s}_t^T]$ to make d_t^j robust to noise. To predict $\Delta \mathbf{p}_t$ for a new subject given $\mathbf{s}_t$, predictions $\Delta \hat{\mathbf{p}}_t^j$ are obtained for the K closest models by Eq. (3) and combined by

$$\Delta \hat{\mathbf{p}}_t = \sum_j w_t^j \Delta \hat{\mathbf{p}}_t^j, \quad w_t^j = \frac{1/(d_t^j + \epsilon)}{\sum_{k=1}^J 1/(d_t^k + \epsilon)}$$

with ϵ some very small positive value. (4)

References

1. Ahrendt P (2005) The multivariate Gaussian probability distribution. Tech. rep
2. Arnold P, Preiswerk F, Fasel B, Salomir R, Scheffler K, Cattin P (2011) 3D organ motion prediction for MR-guided high intensity focused ultrasound. In: Medical image computing and computer-assisted intervention, pp 623–630
3. Blackall J, Ahmad S, Miquel M, McClelland J, Landau D, Hawkes D (2006) MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. Phys Med Biol 51:4147
4. Blanz V, Vetter T (2002) Reconstructing the complete 3D shape of faces from partial information. Informationstechnik und Technische Informatik 44(6):295–302
5. De Senneville B, Ries M, Moonen C (2013) Real-time anticipation of organ displacement for MR-guidance of interventional procedures. In: IEEE international symposium on biomedical imaging, p 1420
6. Ehrhardt J, Werner R, Schmidt-Richberg A, Handels H (2011) Statistical modeling of 4D respiratory lung motion using diffeomorphic image registration. IEEE Trans Med Imag 30(2):251–265
7. Eom J, Xu X, De S, Shi C (2010) Predictive modeling of lung motion over the entire respiratory cycle using measured pressure–volume data, 4DCT images, and finite-element analysis. Med Phys 37(8):4389–4401
8. Hartkens T, Rueckert D, Schnabel J, Hawkes D, Hill D (2002) VTK CISG registration toolkit: an open source software package for affine and non-rigid registration of single-and multimodal 3D images. In: Bildverarbeitung für die Medizin, p 409
9. He T, Xue Z, Xie W, Wong S (2010) Online 4-D CT estimation for patient-specific respiratory motion based on real-time breathing signals. In: Medical image computing and computer-assisted intervention, p 392
10. Holbrook A, Ghanouni P, Santos J, Dumoulin C, Medan Y, Pauly K (2014) Respiration based steering for high intensity focused ultrasound liver ablation. Magn Reson Med 71(2):797–806
11. King A, Buerger C, Tsoumpas C, Marsden P, Schaeffter T (2012) Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. Med Image Anal 16:252–264
12. Klinder T, Lorenz C, Ostermann J (2009) Free-breathing intra-and intersubject respiratory motion capturing, modeling, and prediction. In: Proceedings of SPIE, vol 7259. International Society for Optics and Photonics, p 72590T
13. Liu X, Oguz I, Pizer S, Mageras G (2010) Shape-correlated deformation statistics for respiratory motion prediction in 4D lung. In: Proceedings SPIE, vol 7625. International Society for Optics and Photonics
14. Low D, Parikh P, Lu W, Dempsey J, Wahab S, Hubenschmidt J, Nystrom M, Handoko M, Bradley J (2005) Novel breathing motion model for radiotherapy. Int J Radiat Oncol Biol Phys 63(3):921–929
15. McClelland J, Hawkes D, Schaeffter T, King A (2013) Respiratory motion models: a review. Med Image Anal 17(1):19–42
16. McClelland J, Hughes S, Modat M, Qureshi A, Ahmad S, Landau D, Ourselin S, Hawkes D (2011) Inter-fraction variations in respiratory motion models. Phys Med Biol 56:251–272
17. Nguyen T, Moseley J, Dawson L, Jaffray D, Brock K (2009) Adapting liver motion models using a navigator channel technique. Med Phys 36(4):1061–1073
18. Pernot M, Tanter M, Fink M (2004) 3-D real-time motion correction in high-intensity focused ultrasound therapy. Ultrasound Med Biol 30(9):1239–1249
19. Preiswerk F, Arnold P, Fasel B, Cattin P (2011) A Bayesian framework for estimating respiratory liver motion from sparse measurements. In: Abdominal imaging, computational and clinical applications, p 207
20. Preiswerk F, De Luca V, Arnold P, Celicanin Z, Petrusca L, Tanner C, Bieri O, Salomir R, Cattin P (2014) Model-guided respiratory organ motion prediction of the liver from 2D ultrasound. Med Image Anal 18(5):740
21. Ross J, Tranquebar R, Shanbhag D (2008) Real-time liver motion compensation for MRgFUS. In: Medical image computing and computer-assisted intervention, p 806
22. Roujol S, Benois-Pineau J, de Senneville B, Ries M, Quesson B, Moonen C (2012) Robust real-time-constrained estimation of respiratory motion for interventional MRI on mobile organs. IEEE Trans Inf Technol B 16(3):365–374
23. Roujol S, Ries M, Moonen C, de Senneville B (2011) Robust real time motion estimation for MR-thermometry. In: IEEE international symposium on biomedical imaging, p 508
24. Rueckert D, Sonoda L, Hayes C, Hill D, Leach M, Hawkes D (1999) Nonrigid registration using free-form deformations: application to breast MR images. IEEE Trans Med Imag 18(8):712
25. Samei G, Tanner C, Székely G (2012) Predicting liver motion using exemplar models. In: Abdominal imaging. Computational and clinical applications, p 147 (2012)
26. Schwenke M, Strehlow J, Haase S, Jenne J, Tanner C, Langø T, Loeve A, Karakitsios I, Xiao X, Levy Y, Sat G, Bezzi M, Braunewell S, Guenther M, Melzer A, Preusser T (2015) An integrated model-based software for fus in moving abdominal organs. Int J Hyperth 31(3):240–250
27. de Senneville B, Mougenot C, Moonen C (2007) Real-time adaptive methods for treatment of mobile organs by MRI-controlled high-intensity focused ultrasound. Magn Reson Med 57(2):319
28. de Senneville BD, Ries M, Maclair G, Moonen C (2011) MR-guided thermotherapy of abdominal organs using a robust PCA-based motion descriptor. IEEE Trans Med Imag 30(11):1987
29. Tanner C, Boye D, Samei G, Székely G (2012) Review on 4D models for organ motion compensation. CR Rev Biom Eng 40(2):135
30. Tanner C, Eppenhof K, Gelderblom J, Székely G (2014) Decision fusion for temporal prediction of respiratory liver motion. In: IEEE international symposium on biomedical imaging, p 698
31. Tanner C, Samei G, Székely G (2015) Robust exemplar model of respiratory liver motion and individualization using an additional breath-hold image. In: IEEE international symposium on biomedical imaging, p 1576 (2015)
32. Tanter M, Pernot M, Aubry JF, Montaldo G, Marquet F, Fink M (2007) Compensating for bone interfaces and respiratory motion in high-intensity focused ultrasound. Int J Hyperth 23(2):141–151
33. Von Siebenthal M, Székely G, Gamper U, Boesiger P, Lomax A, Cattin P (2007) 4D MR imaging of respiratory organ motion and its variability. Phys Med Biol 52:1547

34. Von Siebenthal M, Székely G, Lomax A, Cattin P (2007) Inter-subject modelling of liver deformation during radiation therapy. In: Medical image computing and computer-assisted intervention, p 659

35. Zadicario E, Rudich S, Hamarneh G, Cohen-Or D (2010) Image-based motion detection: using the concept of weighted directional descriptors. IEEE Eng Med Biol 29:87

36. Zhang Q, Pevsner A, Hertanto A, Hu Y, Rosenzweig K, Ling C, Mageras G (2007) A patient-specific respiratory model of anatomical motion for radiation treatment planning. Med Phys 34(12):4772–4781